Lachman/Lieberman's

# The Theory and Practice of Industrial Pharmacy

Fourth Edition

*Editors*

Roop K Khar
SP Vyas
Farhan J Ahmad
Gaurav K Jain

CBS

CBS Publishers & Distributors Pvt Ltd

# Lachman/Lieberman's

# The Theory and Practice of Industrial Pharmacy

## Fourth Edition

# Lachman/Lieberman's

# The Theory and Practice of Industrial Pharmacy

## Fourth Edition

*Editors*


**Roop K Khar** MPharm PhD DBM
Principal and Professor
BS Anangpuria Institute of Pharmacy, Faridabad, Haryana
Former Dean and Head, Faculty of Pharmacy, Jamia Hamdard, New Delhi
Executive Director, Pharmulators Koncept Pvt Ltd, New Delhi

**SP Vyas** MPharm PhD Post Doc
Professor, Department of Pharmaceutical Sciences
Head, Department of Biotechnology
Dr HS Gour University, Sagar, Madhya Pradesh

**Farhan J Ahmad** MPharm PhD
Associate Professor, Department of Pharmaceutics
Faculty of Pharmacy, Jamia Hamdard, New Delhi

**Gaurav K Jain** MPharm PhD
Assistant Professor, Department of Pharmaceutics
Faculty of Pharmacy, Jamia Hamdard, New Delhi


**CBS Publishers & Distributors** Pvt Ltd
New Delhi • Bengaluru • Chennai • Kochi • Mumbai • Kolkata
Hyderabad • Pune • Nagpur • Manipal • Vijayawada • Patna

***Disclaimer***

Science and technology are constantly changing fields. New research and experience broaden the scope of information and knowledge. The authors have tried their best in giving information available to them while preparing the material for this book. Although, all efforts have been made to ensure optimum accuracy of the material, yet it is quite possible some errors might have been left uncorrected. The publisher, the printer and the authors will not be held responsible for any inadvertent errors, omissions or inaccuracies.

eISBN: 978-81-239-2754-1



**First eBook Edition: 2016**



Published by Satish Kumar Jain and produced by Varun Jain for
**CBS Publishers & Distributors Pvt Ltd**

**Head Office:** CBS PLAZA, 4819/XI Prahlad Street, 24 Ansari Road, Daryaganj, New Delhi-110002, India.

Ph: +91-11-23289259, 23266861, 23266867; Fax: 011-23243014; Website: www.cbspd.com; E-mail: delhi@cbspd.com; publishing@cbspd.com; cbspubs@airtelmail.in.

**Corporate Office:** 204 FIE, Industrial Area, Patparganj, New Delhi-110092

Ph: +91-11-49344934; Fax: +91-11-49344935; Website: www.cbspd.com; E-mail: eresources@cbspd.com; admin@cbspd.com

# *Branches*

- **Bengaluru:** Seema House 2975, 17th Cross, K.R. Road, Banasankari 2nd Stage, Bengaluru - 560070, Karnataka

  Ph: +91-80-26771678/79; Fax: +91-80-26771680; E-mail: bangalore@cbspd.com
- **Chennai:** No.7, Subbaraya Street Shenoy Nagar Chennai - 600030, Tamil Nadu

  Ph: +91-44-26680620, 26681266; E-mail: chennai@cbspd.com
- **Kochi:** 36/14 Kalluvilakam, Lissie Hospital Road, Kochi - 682018, Kerala

  Ph: +91-484-4059061-65; Fax: +91-484-4059065; E-mail: kochi@cbspd.com
- **Mumbai:** 83-C, 1st floor, Dr. E. Moses Road, Worli, Mumbai - 400018, Maharashtra

Ph: +91-22-24902340 - 41; Fax: +91-22-24902342; E-mail: mumbai@cbspd.com

- **Kolkata:** No. 6/B, Ground Floor, Rameswar Shaw Road, Kolkata - 700014
  Ph: +91-33-22891126 - 28; E-mail: kolkata@cbspd.com

## *Representatives*

- **Hyderabad**
- **Pune**
- **Nagpur**
- **Manipal**
- **Vijayawada**
- **Patna**

*to*

---

*the pharmaceutical profession*
*and*
*Gen Next of*
*pharmacists and*
*pharmaceutical scientists*

# Foreword

Industrial pharmacy is the foundation for learning and understanding the physicochemical properties of drug molecules, developing products and formulations and establishing stability and shelf-life of medicines. Beside, its knowledge provides the basis of explaining physiological process in human body including drug absorption, distribution, metabolism and elimination. One can also predict the therapeutic behaviour including interactions, adverse drug reactions and contraindications. Its knowledge is useful to understand A to Z of drug and product development.

Before the realization of the importance of industrial pharmacy about 60 years ago, the drug products were empirically prepared without rational basis. Now the formulations have become more efficacious, reliable and reproducible with reduced side effects.

Voluminous literature exists in the form of both text/reference books and research papers dealing with various aspects and facets of industrial pharmacy. The authors of this textbook are teachers at different reputed Indian universities and professionals from various pharmaceutical companies. The chief editor, Prof Roop Kishen Khar, is a renowned pharmacy teacher of more than 35 years of experience of teaching and research in pharmaceutics in general and industrial pharmacy in particular. The present book has a flavour of his in-depth knowledge. Other editor, Prof SP Vyas is an eminent academician and Dr Farhan Jalees Ahmad, Associate Professor of Pharmaceutics with 13 years of teaching with 7 years of industrial experience, has given it a flavour of necessary applications in product development. Dr Gaurav Kumar Jain, another editor is young, budding and enthusiastic assistant professor associated with teaching of the subject for 7 years. He is conversant with the requirement of the students and commonly has very discerningly provided the theoretical basis and practical applications.

The book *The Theory and Practice of Industrial Pharmacy*, being published by CBS Publishers & Distributors, comprises 30 chapters divided

in four sections, viz. • Principles of Pharmaceutical Processing, • Pharmaceutical Dosage Form Design, • Pharmaceutical Dosage Forms, and • Product Processing, Packaging Evaluation and Regulation. The book is well written with necessary details and the multi-author approach has resulted in a unique textbook of industrial pharmacy. This text should be useful to undergraduate and postgraduate students, practitioners in the pharmaceutical and allied health areas, hospital pharmacists, drug patent attorneys, government scientists and regulatory personnel, and other professionals seeking information regarding design, manufacture and control of the pharmaceutical dosage forms.

**Dr SN Sharma**
Emeritus Professor
Jamia Hamdard
New Delhi

# Foreword

The very famous publication by Lachman, Lieberman and Kanig was a standard reference when I practised industrial pharmacy in the USA. Even after my return to India it guided me in my work in pharmaceutical research and development, manufacturing and quality assurance. I am very pleased that this book is being updated and published in enriched format by the Indian experts. I feel privileged to write its Foreword.

*The Theory and Practice of Industrial Pharmacy* by Lachman, Lieberman and Kanig was first published in 1970. This book continues to be a comprehensive resource for students and practitioners of industrial pharmacy, providing information in pharmaceuticals, allied health sciences and others pursuing design, development, manufacture, quality and regulation of pharmaceutical dosage forms. Its last edition was published 26 years ago but since then, intervening period has witnessed many important changes in the field of the industrial pharmacy.

The editors of the fourth edition have achieved the challenge by selecting suitable contributors of the textbook. The fourth edition has been thoroughly revised and updated to keep pace with the changes in pharmacy curriculum, retaining its recognized position in the field of industrial pharmacy by presenting topics in layers—from basic to in-depth discussions. Each section of the book has been critically reviewed to reflect the emerging trends in the field. Hence, this book is expected to familiarize students with the fundamental concepts with practical applications. Furthermore, updated illustrations are summarized with references. This edition has 30 chapters divided into four sections.

**Section I** has been reorganized and simplified so as to focus on imperative topics dealing with pharmaceutical processing. New chapters including pharmaceutical utilities and dissolution have been added, since they are critical part of pharmaceutical processing and quality assurance.

**Section II** includes techniques and methodologies for biopharmaceutical dosage form design and includes pharmaceutical statistics and optimization.

**Section III** is a very useful part of the book in terms of information on various conventional drug delivery systems. Addition of new chapters on excipients and polymers and on novel and target drug delivery broadens the scope. This section brings the reader up-to-date on the latest trends and approaches to dosage form design.

**Section IV** has been significantly improved keeping in mind its use for postgraduate students, as well as the scientists involved in scale up, packaging, evaluation and regulation of drug delivery systems.

The editors and contributors have focused in their skillful review of literature and incorporated their own unique perspective and experience. The multi-author approach has resulted in a uniquely prepared textbook for industrial pharmacy, which is hoped to serve as a standard textbook for students from undergraduate to postgraduate levels.

**Prafull D Sheth**
Former Executive Vice President and
Member of the Board, Ranbaxy Laboratories Ltd and
Vice President, International Pharmaceutical Federation (FIP)

# Preface to the Fourth Edition

For more than 40 years and through the three previous editions (First edition 1970, Second edition 1976, Third edition 1986). *The Theory and Practice of Industrial Pharmacy* has remained the "gold standard" text and reference source covering all the aspects of industrial pharmacy from principles of pharmaceutical processing through dosage forms design and product processing to facility design, management, packaging, quality control and regulation.

The opportunity to edit and expand the fourth edition of this book is invaluable. It is important to take into consideration the contemporary developments and rapid changes that have taken place in this core area of pharmaceutical sciences and provide an up-to-date resource material for the students, teachers and researchers.

The new edition incorporates some new chapters, reorientation, revision and modification of existing chapters and addition of many new photographs and illustrations. The editors have striven to provide the reader with an up-to-date knowledge base for all aspects of industrial pharmacy.

The structure and content of this edition have been changed to reflect modern thinking and current university curricula throughout the world. The contents are designed and written intentionally to serve both novice as well as experts in the field of industrial processing and dosage form design. It covers an umbrella of relevant topics and is divided in four principle sections.

The contents of *Section I* cover the topics of physical pharmacy and unit operation, taught in early years of graduation, and thus the text has been reorganized and simplified to make the material more accessible to students. Three new chapters have been included to widen the comprehensive nature of this text and to cover the entire syllabi of pharmaceutics at different levels.

*Section II* is suitably nurtured and expanded to suffice the requirement of final year undergraduate as well as graduate and postgraduate students. New methodologies for preformulation have been included. Pharmacokinetic and statistical equations have been simplified and are exemplified for better

understanding. Information on optimization and use of software for optimization has also been included in order to be in pace with new FDA guidelines for quality-by-design.

*Section III* is the quintessence of this book and has been modernized to include latest techniques, machinery and developments. The text has been modified to make it usable for both undergraduate and postgraduate students and for formulation scientists working in pharmaceutical industries. Inclusion of new chapters on pharmaceutical excipients and polymers, novel drug delivery and targeted drug delivery further aims to broaden the scope of text and provide elucidation of basic concepts and information relevant for research scientists.

*Section IV* has been appreciably improved keeping in mind its use for postgraduate students as well as for the scientists involved in packaging, evaluation and regulation of drug delivery systems. Text has been significantly improvised to include necessary guidelines and standards. Core information has been carefully selected to provide the sound knowledge base necessary for competent regulatory practice both in the USA and India.

Finally, the *appendix* includes important web portals, solvent miscibility chart, limit of residual solvent and preparation of standard buffer, which provide a comprehensive reference source on modern industrial pharmacy.

The involvement of experienced experts and new generation of experts make the proper blend for imparting knowledge and information to pharmacy undergraduates, graduates, postgraduates and formulation scientists. We are grateful to all section editors and contributors for their skillful review of literature and for incorporating their own unique perspective and experience into their chapters. Without the skillful sharing of their knowledge in the pages of this book, the enormous task of compiling this fourth edition of the textbook would have been impossible to achieve.

We also acknowledge photographs and illustrations supplied by Erweka India Pvt Ltd, Unicure (India) Pvt Ltd, GlaxoSmithKline, Ranbaxy Laboratories Ltd and Arbro Pharmaceuticals Ltd.

We and our contributing authors will be extremely pleased if our efforts would be useful to provide an improved textbook to serve as a teaching and reference source in industrial pharmacy. Any suggestions and changes are solicited.

Finally, I acknowledge the efforts of Mr. SK Jain, Mr. Varun Jain, Mr. YN Arjuna and Mr. Ashish Dixit from CBS Publishers & distributors Pvt. Ltd., New Delhi in publishing this ebook.

**Roop K Khar**
**SP Vyas**
**Farhan J Ahmad**
**Gaurav K Jain**

# About the Book

For more than 40 years and throughout its three editions this popular book has remained the “gold standard” text and reference source in industrial pharmacy covering allaspects ranging from pharmaceutical processing, dosage forms design and product processing through manufacturing processes to facility design and management, packaging, evaluation, quality control and regulation.

This is the thoroughly revised and updated edition, completely rewritten to keep pace with the changes in pharmacy curricula, keeping its distinction in the field of industrial pharmacy. By presenting the topics in layers, from basic to in-depth discussion, the text enables easy conception of A to Z of product development. In addition to the new chapters, this edition retains the flavour and essence of all the previous editions. Wherever necessary chapters have been made concise so as to focus on imperative topics taught in industrial pharmacy and to make the material more accessible to students. The multi-author approach provides contemporary concepts and the knowledge from various experts necessary for dealing with the manifold aspects of modern industrial pharmacy.

## Important Features

- **A balanced and contemporary** view of field, with renowned editors, section editors and contributing authors drawn from different academic institutes and industries.
- **Six new chapters** including • evaporation and distillation • pharmaceutical utility • dissolution • pharmaceutical excipients and polymers • novel drug delivery systems and • targeted drug delivery systems.
- **Completely rewritten chapters** present up-to-date information in a rational and systematic mode embodying salient details and illustrative examples.
- **Appendix** includes web addresses of regulatory bodies, solvent miscibility table, limits of residual solvent and buffer compositions useful for researchers.

# About the Author

**Roop K Khar** M Pharm PhD DBM is Principal of BS Anangpuria Institute of Pharmacy, Faridabad (NCR) and former Dean and Head, Department of Pharmacy, Jamie Hamdard, New Delhi. He is the executive director and founder of Pharmulators Koncept Ltd, providing consultancy services to the pharmaceutical industries and educational institutions. He has contributed more than 255 publications with a cumulative impact factor of 370 and 2830 citations. He is author of 11 text and reference books and has supervised 53 scholars for PhD. He has to his credit several Indian and US patents in the area of dosage form development and has been a popular speaker at various national and international forums.

**Suresh P Vyas** M Pharm PhD Post Doc is a Professor, Department of Pharmaceutical Sciences, and Head, Department of Biotechnology, Dr HS Gour University, Sagar. A pioneer scientist in the field of targeted drug delivery, none-biotechnology and immunology, he has supervised 47 scholars for PhD and more than 120 M Pharm students. He has fo his credit over 300 publications, 13 reference books and contributed more than 20 chapters to several international books. He has been on the editorial boards of various infernafional journals.

**Farhan J Ahmad** M Pharm PhD is Associate Professor, Jamie Hamdard, New Delhi. He has a rich experience in industrial research through his fruilful association with Ranbaxy Laboratories Ltd, for six long years. He is working in the area of nanomedicine for the last ten years and two of his nanoproducts are in clinical stage. He has supervised more than 25 scholars for PhD and contributed to more than 200 publications, including 25 patents and several textbooks. He is currently President, Indian Pharmaceutical Association-Delhi Branch.

**Gaurav K Jain** M Pharm PhD is Assistant Professor, Jamie

Hamdard, New Delhi, teaching concepts of physical and industrial pharmacy along with development of nano drug delivery systems. He has published over 30 publications, including textbooks, and holds several patents. He provides consulfancy fo pharmaceutical companies and has done projects for Ranbaxy Laboratories Ltd, GlaxoSmilhKiine, Galderma India and Arbro Pharmaceuticals. He is also Joint Secretary, Indian Pharmaceutical Association-Delhi Branch, and General-Secretary, Good Society for Ethical Research.

# List of Contributors

# EDITORS

**Roop K Khar** MPharm PhD DBM
Principal and Professor
BS Anangpuria Institute of Pharmacy
Faridabad, Haryana, India
Former Dean and Head, Faculty of Pharmacy
Jamia Hamdard, New Delhi, India
Executive Director, Pharmulators Koncept Pvt Ltd
New Delhi, India

**SP Vyas** MPharm PhD Post Doc
Professor, Department of Pharmaceutical Sciences
Head, Department of Biotechnology
Dr HS Gour University
Sagar, Madhya Pradesh, India

**Farhan J Ahmad** MPharm PhD
Associate Professor, Department of Pharmaceutics
Faculty of Pharmacy, Jamia Hamdard
New Delhi, India

**Gaurav K Jain** MPharm PhD Assistant Professor, Department of Pharmaceutics Faculty of Pharmacy, Jamia Hamdard New Delhi, India

## SECTION EDITORS

**Nitin Jain** MPharm PhD
Research Associate, Nanomedicine Laboratory
Faculty of Pharmacy, Jamia Hamdard
New Delhi, India

**Musarrat H Warsi** MPharm PhD
Sr. Research Fellow, Department of Pharmaceutics
Faculty of Pharmacy, Jamia Hamdard
New Delhi, India

**Vaidehi Jain** MPharm
Research Scholar, Department of Quality Assurance
Faculty of Pharmacy, Jamia Hamdard
New Delhi, India

**Sohail Hasan** MPharm PhD
Manager, Formulation Development
Pharmulators Koncept Pvt Ltd, New Delhi, India

# CONTRIBUTORS

**Akash Chaurasiya** MPharm PhD
Sr. Research Scientist, Formulation Development
Fresenius Kabi Oncology Ltd, Gurgaon, Haryana, India

**Anirudh Vashishth** MPharm
Associate Manager, Global Regulatory CMC
Novartis Health Care Pvt Ltd Hyderabad, India

**Bhuvaneshwar Vaidya** MPharm PhD
Associate Professor, Pharmaceutics
ISF College of Pharmacy, Moga, Punjab

**Devyani Dubey** MPharm PhD
Associate Professor, Pharmaceutics
ISF College of Pharmacy, Moga, Punjab

**Harshveer Singh** MPharm
Research Scientist, Packaging Department
Jubilant Life Sciences Ltd, Noida
Uttar Pradesh, India

**Mayank Singhal** MPharm
Research Scientist, School of Pharmaceutical Sciences
University of Geneva, Geneva, Switzerland

**Mohammed Anwar** MPharm
Sr. Research Fellow, Department of Pharmaceutics
Faculty of Pharmacy, Jamia Hamdard, New Delhi, India

**Neeraj Kumar** MPharm PhD
Consultant
Institute of Nuclear Medicine and Allied Sciences
DRDO, New Delhi, India

**Neha Mallick** MPharm

Sr. Research Fellow, Department of Pharmaceutics
Faculty of Pharmacy, Jamia Hamdard
New Delhi, India

**Jayabalan Nirmal** MPharm PhD
Research Postdoctoral Scientist
Oakland University William Beaumont School of Medicine
Royal Oak
Michigan, US

**Nilu Jain** MPharm PhD
Formulation Scientist
Vyome Biosciences, Delhi, India

**Pankaj Musyuni** MPharm
Senior Executive Regulatory Affairs
Parijat Industries (India) Pvt Ltd, New Delhi, India

**Rishi Paliwal** MPharm PhD
Assistant Professor, Pharmaceutics
Columbia Institute of Pharmacy, Raipur, Chhattisgarh

**Ritu Hans** M Pharm
Scientific Writer
Novartis Health Care Pvt Ltd, Hyderabad, India

**Saurabh Arora** MPharm PhD
Managing Director, Auriga Research Ltd
Executive Director, Arbro Pharmaceuticals Ltd
New Delhi, India

**Shadab A Pathan** MPharm PhD
Deputy Manager, New Product Development
GlaxoSmithKline Consumer Healthcare Ltd
Research and Development Centre
Gurgaon, Haryana, India

**Shilu D Thomas** MPharm

Research Scholar, Department of Pharmacology
Faculty of Pharmacy, Jamia Hamdard
New Delhi, India

**Shivani R Paliwal** MPharm PhD
Assistant Professor, Pharmaceutics
SLT Institute of Pharmaceutical Sciences
Guru Ghasidas Vishwavidyalaya
Bilaspur, Chhattisgarh

**Sreejith C Kurup** MPharm
Associate Head, Drug Regulatory Affairs
Jubilant Life Sciences Ltd, Noida, Uttar Pradesh, India

**Suruchi Suri** MPharm PhD
Sr. Research Fellow, Department of Biotechnology
Faculty of Pharmacy, Jamia Hamdard, New Delhi, India

**Tripta Bansal** MPharm PhD
Scientist B
National Dope Testing Laboratory
MYAS Govt of India, New Delhi, India

## LINGUISTIC SUPPORT

**Ayesha Baig** MPharm
Research Scholar, Department of Pharmaceutics
Faculty of Pharmacy, Jamia Hamdard
New Delhi, India

**Ayesha Pasha** MPharm
Research Scholar, Department of Pharmaceutics
Faculty of Pharmacy, Jamia Hamdard
New Delhi, India

# Contents

RK Khar, N Mallick, GK Jain, JH Wood*

7. **Pharmaceutical Utility**

FJ Ahmad, A Chaurasiya

8. **Dissolution**

GK Jain, V Jain

## Section II
Pharmaceutical Dosage Form Design

9. **Preformulation**

GK Jain, FJ Ahmad, RK Khar, EF Fiese*, TA Hagen*

10. **Biopharmaceutics and Pharmacokinetics**

N Jain, T Bansal, J Nirmal, RK Khar, KC Kwan*, MR Dobrinska*, JD Rogers*, AE Till*, KC Yeh*

11. **Pharmaceutical Statistics and Optimization**

N Jain, J Nirmal, RK Khar, S Bolton*

## Section III
Pharmaceutical Dosage Forms

12. **Pharmaceutical Excipients and Polymers**

SP Vyas, RK Khar, GK Jain, N Jain

13. **Tablets**

Part I: Tablet Design

Part II: Tablet

Coating Part III: Tablet Punch Forces

GK Jain, N Jain, RK Khar, GS Banker*, NR Anderson*, K

# References

* Original Authors of third edition of the book

## Section I:

# Principles of Pharmaceutical Processing

# 1:
# Mixing

Almost every pharmaceutical product contains more than one component, and this necessitates mixing or blending stages in their manufacturing process. Perry and Chilton (1973) defined *mixing* as a process "in which two or more ingredients in separate or roughly mixed condition are treated so that each particle of any one ingredient is as nearly as possible adjacent to a particle of each of the other ingredients." The term *blending* is synonymous with mixing, and *demixing* or *segregation* is the opposite.

Mixing tends to result in a randomization of dissimilar particles within a system. This is to be distinguished from an ordered system in which the particles are arranged according to some iterative rule and thus follow a repetitive pattern. It is possible to consider the mixing of particles differing only by some vector quantity, such as spatial orientation or translational velocity.

Mixing is a fundamental step in most process sequences, and is normally carried out:

1. To control heat and mass transfer
2. To secure uniformity of composition so that small samples withdrawn from a bulk material represent the overall composition of the mixture
3. To improve single phase and multi-phase systems
4. To promote physical and chemical reactions, such as dissolution, in which natural diffusion is supplemented by agitation

Mixing can be classified as positive, negative, or neutral. Positive mixing applies to the systems where spontaneous, irreversible and complete mixing would take place, by diffusion, without the expenditure of energy, provided time is unlimited, although the input of energy by using mixing apparatus will shorten the time required to obtain the desired degree of mixing. In general, positive mixtures, such as a mixture of two gases or two miscible liquids do not present any problems during mixing. Negative mixing is

demonstrated by biphasic systems, in which the phases differ in density. Any two-phase systems such as suspensions of solids in liquids, emulsions and creams tend to separate out quickly, unless energy is continually expended on them. Negative mixtures are generally more difficult to form and maintain, and require a higher degree of mixing as compared to positive mixtures. Neutral mixing occurs when neither mixing nor demixing takes place unless the system is acted upon by an external energy input. Neutral mixtures are static in behavior, have no tendency to mix spontaneously or segregate spontaneously and include mixture of powders, pastes and ointments. The following text deals with the fundamental concepts and equipment employed in the pharmaceutical industries to obtain satisfactory mixing and focuses on practical considerations involved in the evaluation of mixing efficiency. The equipments used for mixing liquids, semisolids and solids are depicted in Fig. 1.1.

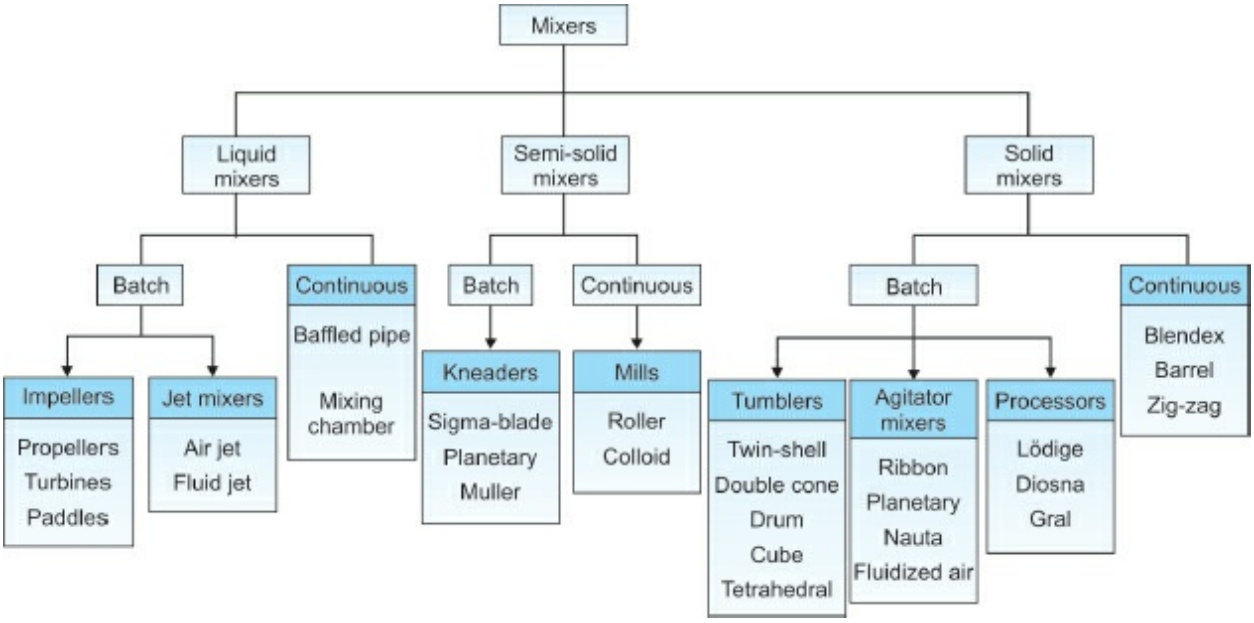



**Fig. 1.1:** Classification tree of mixing equipments

# LIQUID MIXING

## Mixing Mechanisms

Mixing mechanisms for fluids fall essentially into four categories: bulk transport, turbulent flow, laminar flow, and molecular diffusion. Usually, more than one of these mechanisms is operative in practical mixing situations.

### *Bulk Transport*

The movement of a relatively large portion of the material being mixed from one location in a system to another constitutes bulk transport. A simple circulation of material in a mixer, however, does not necessarily result in efficient mixing. For bulk transport to be effective, it must result in a rearrangement or permutation of the various portions of the materials to be mixed. This is usually accomplished by means of paddles, revolving blades, or other devices, within the mixer, arranged so as to move adjacent volumes of the fluid in different directions, thereby shuffling the system in three dimensions.

### *Turbulent Mixing*

The phenomenon of turbulent mixing is the direct result of turbulent fluid flow, which is characterized by a random fluctuation of the fluid velocity at any given point within the system. The fluid velocity at any given instant may be expressed as the vector sum of its components in the x, y, and z directions. With turbulence, these directional components fluctuate randomly about their individual mean values, as does the velocity itself.

In general, with turbulence, the fluid has different instantaneous velocities at different locations and such velocity differences within the body of fluid produce a randomization of the fluid molecules. For this reason, turbulence is a highly effective mechanism of mixing.

Turbulent flow can be conveniently visualized as a composite of eddies of various sizes. An eddy is defined as a portion of a fluid moving as a unit in a direction often contrary to that of the general flow. Large eddies tend to break up, forming eddies of smaller size until they are no longer distinguishable. The size distribution of eddies within a turbulent region is referred to as the

*scale of turbulence*. Thus, when small eddies are predominant, the *scale of turbulence* is low.

An additional characteristic of turbulent flow is its *intensity*, which is related to the velocities with which the eddies move. A composite picture of eddy size versus the velocity distribution of each size eddy may be described as a complex spectrum. Such a spectrum is the characteristic of a turbulent flow and is used in its analysis.

## *Laminar Mixing*

Streamline or laminar flow is frequently encountered when highly viscous fluids are being processed. It can also occur if stirring is relatively gentle and may exist adjacent to stationary surfaces in the vessels in which the flow is predominantly turbulent. When two dissimilar liquids are mixed through laminar flow, the shear that is generated stretches the interface between them. If the mixer employed folds the layers back upon themselves, the number of layers, and hence the interfacial area between them, increases exponentially with time. This relationship is observed because the rate of increase in interfacial area with time is proportional to the instantaneous interfacial area.

Mixers may also operate by simply stretching the fluid layers without any significant folding action. This mechanism does not have the stretch compounding effect produced by folding, but may be satisfactory for some purposes in which only a moderate reduction in *mixing scale* (defined in detail later) is required. It should be pointed out, however, that by this process alone, an exceedingly long time is required for the layers of the different fluids to reach molecular dimensions. Therefore, good mixing at the molecular level requires a significant contribution by molecular diffusion after the layers have been reduced to a reasonable thickness (several hundred molecules) by laminar flow.

## *Molecular Diffusion*

The primary mechanism responsible for mixing at the molecular level is diffusion, resulting from the thermal motion of the molecules. When it occurs in conjunction with laminar flow, molecular diffusion tends to reduce the sharp discontinuities at the interface between the fluid layers, and if allowed to proceed for sufficient time, results in complete mixing.

The process is described quantitatively in terms of Fick’s first law of

diffusion:

$$\frac{dm}{dt} = -DA\frac{dc}{dx} \qquad \ldots (1)$$

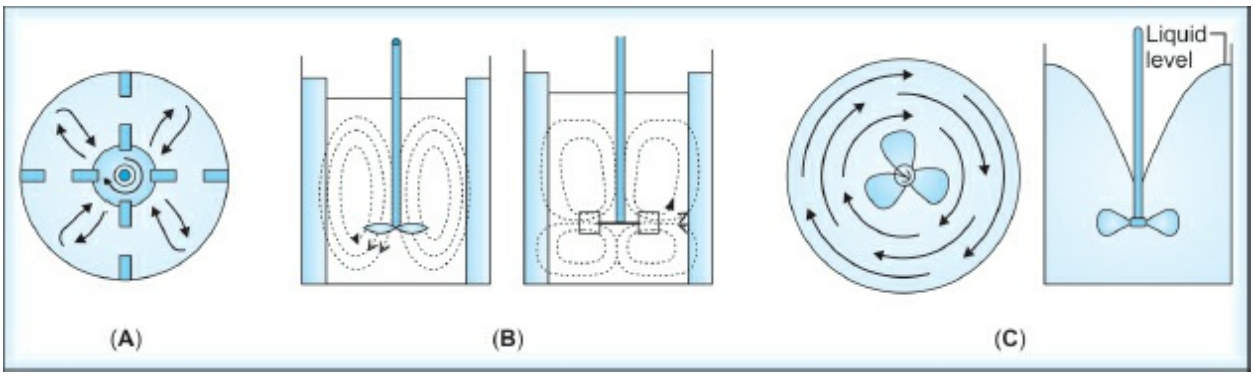
where the rate of transport of mass, *dm/dt*, across an interface of area, *A*, is proportional to the concentration gradient, *dc/dx*, across the interface. The rate of intermingling is governed also by the diffusion coefficient, D, which is a function of the variables, i.e. fluid viscosity and size of the diffusing molecules. The sharp interface between dissimilar fluids, which has been generated by laminar flow, may be rather quickly eliminated by the resulting diffusion. Considerable time may be required, however, for the entire system to become homogeneous.

## Equipments

A system for liquid mixing commonly consists of two primary components: (1) a tank or other container suitable for holding the material being mixed, and (2) a means of supplying energy to the system so as to bring about reasonably rapid mixing. Power may be supplied to the fluid mass by means of an impeller, air stream, or liquid jet. Besides supplying power, these also serve to direct the flow of material within the vessel. Baffles, vanes, and ducts are also used to direct the bulk movement of material in such mixers, thereby increasing their efficiency. When the material to be mixed is limited in volume so that it may be conveniently contained in a suitable mixer, *batch mixing* is usually more feasible, however, for larger volumes *continuous mixing* is preferred.

### *Impellers*

Liquids are most commonly mixed by impellers rotating in tanks. These impellers are classified as (i) propellers, (ii) turbines and (iii) paddles. The distinction between impeller types is often made on the basis of the type of flow pattern they produce, or on the basis of the shape and pitch of the blades. The turbulent flow imposed by impeller causes mixing by projecting eddies into, and entraining liquid from the neighboring zone, thereby preventing the formation of dead zones. The flow pattern may be analyzed in terms of three components: radial (perpendicular to the impeller shaft), axial or longitudinal (parallel to the impeller shaft), and tangential (tangential to the circle of rotation around the impeller shaft). These may occur singly or in various combinations. Figure 1.2 illustrates these patterns as they occur in vertical cylindrical tanks.

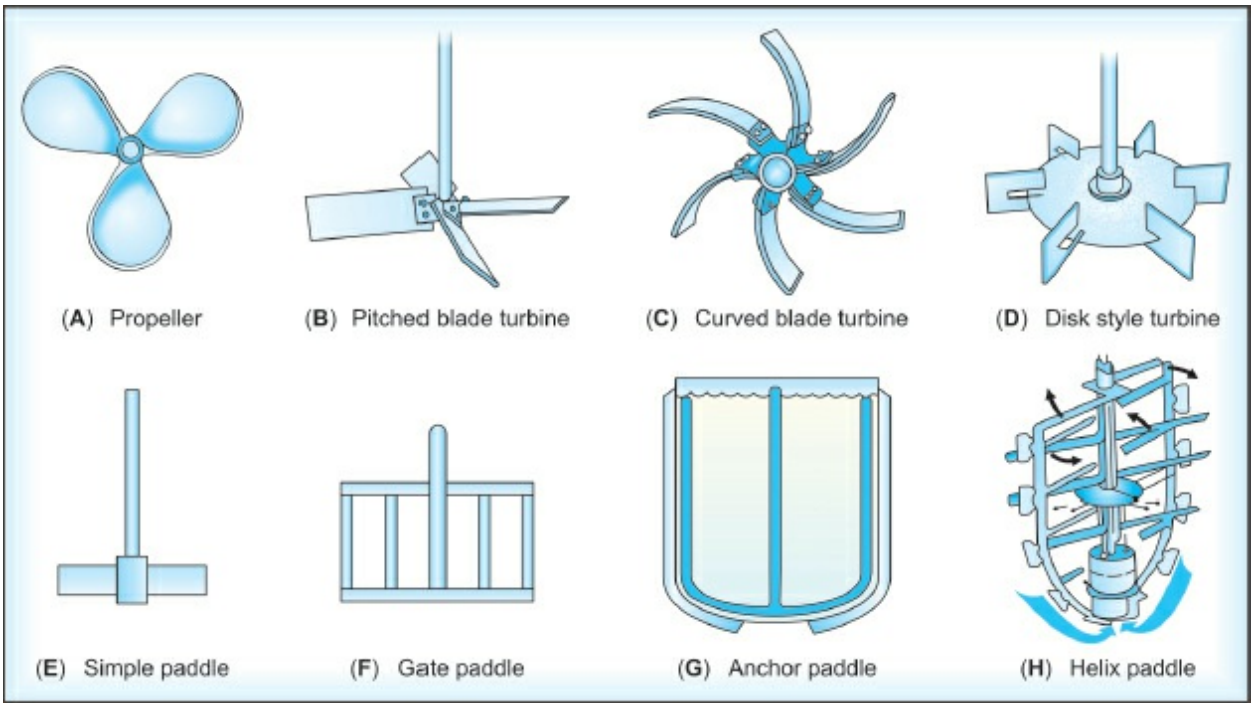



**Figs 1.2A to C:** Diagrammatic representation of flow patterns induced by

impellers: (A) Radial flow, (B) Axial flow, (C) Tangential flow

## *Propellers*

Propellers of various types and forms are used, but all are essentially a segment of a multithreaded screw, that is, a screw with as many threads as the propeller blades (Fig. 1.3A). Also, like the machine screws, propellers may be either right- or left-handed depending on the direction of slant of their blades. As with screws, the propeller pitch is defined as the distance of axial movement per revolution, if no slippage occurs. Although any number of blades may be used, the three-blade design is most commonly used with fluids. The blades may be set at any angle or pitch, but for most applications, the pitch is approximately equal to the propeller diameter. Propellers are most efficient when they run at high speeds in liquids of relatively low viscosity. Although some tangential flow does occur, the primary effect of a propeller is due to axial flow. Also, intense turbulence usually occurs in the immediate vicinity of the propeller. Consider, for example, a downdraft propeller vertically mounted midway to the bottom of the tank. Moderate radial and tangential flow occurs above and below the blade, which acts in conjunction with the axial flow near the shaft, brings portions of fluid together from all regions of the tank, and passes them through the intense turbulence near the blades.




**Figs 1.3A to H:** Impeller blade types: (A) Propeller; (B-D) Turbines; (E-H) Paddles

## *Turbines*

They are usually distinguished from propellers in that the blades of the latter do not have a constant pitch throughout their length. When radial-tangential flow is desired, turbines with blades set at a 90-degree angle to their shaft are employed. With these type of impellers, a radial flow is induced by the centrifugal action of the revolving blades. The drag of the blades on the liquid also results in tangential flow, which in many cases is undesirable. Turbines having tilted blades produce an axial discharge quite similar to that of propellers. Because they lend themselves to a simple and rugged design, these turbines can be operated satisfactorily in fluids 1000 times more viscous than fluids in which a propeller of comparable size can be used. Various types of turbines are depicted in Figs 1.3B to D.

## *Paddles*

Paddles are also employed as impellers and are normally operated at low speeds of 50 rpm or less. Their blades have a large surface area as compared to the tank in which they are employed, a feature that permits them to pass close to the tank walls and effectively mix viscous liquids and semisolids,

which tend to cling to these surfaces. Circulation is primarily tangential, and consequently, concentration gradients in the axial and radial directions may persist in this type of mixer even after prolonged operation. Operating procedures should take these characteristics into account, so as to minimize their undesirable effects. With such mixers, for example, ingredients should not be layered when they are added to the mixing tank. Such vertical stratification can persist even after very long mixing times. The mainstays of high viscosity mixing systems have been gate paddle, anchor paddle and helix (Figs 1.3E to H). These continue to rule the roost, albeit with some interesting hybridization. One of the promising hybrids of paddle mixer is dispertron (Fig. 1.4). It has coaxial blades one for macro mixing and the other for micromixing. Counter rotation of macro-and micromixing elements with variable speed is useful for mixing extremely viscous materials, materials with high solid content and for emulsification and homogenization. The equipment generates maximum shear without vortex formation and minimal air entrapment.

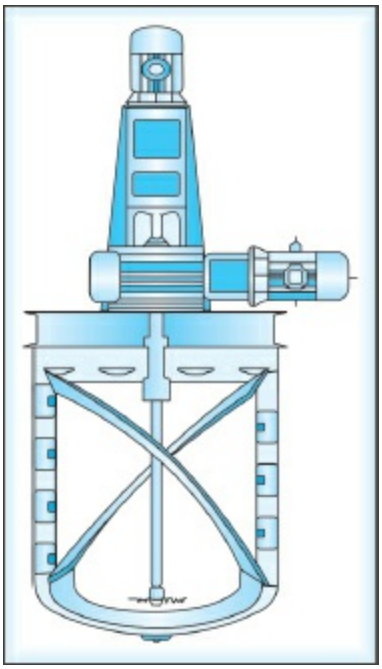


**Fig. 1.4:** Dispertron: Hybrid of paddle mixer

## *Jet Mixers*

### *Air Jets*

Air jet devices involve sub-surface jets of air, or less commonly of some other gas, for effective mixing of certain liquids. Of necessity and for obvious reasons, the liquids must be of low viscosity, nonfoaming, nonreactive with the gas employed, and reasonably nonvolatile. The jets are usually arranged so that the buoyancy of the bubbles lifts liquids from the bottom to the top of the mixing vessel. This is often accomplished with the aid of draft tubes (Fig. 1.5). These serve to confine the expanding bubbles and entrained liquids, resulting in a more efficient lifting action by the bubbles, and inducing an upward fluid flow in the tube. This flow tends to circulate fluid in the tank, bringing it into the turbulent region in the vicinity of the jet. The overall circulation in the mixing vessel brings fluid from all parts of the tank to the region of the jet itself. Here, the intense turbulence generated by the jet produces intimate mixing.

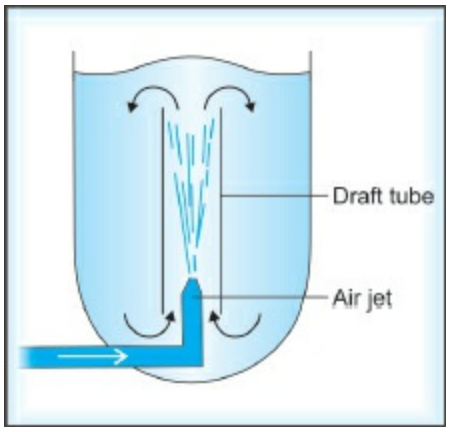



**Fig. 1.5:** Vertical tank with centrally located air jet and draft tube

## *Fluid Jets*

They utilize liquids pumped at high pressure into a tank for mixing. The power required for pumping can often be used to accomplish the mixing operation, either partially or completely. In such a case, the fluids are pumped through nozzles arranged to permit a good circulation of material throughout the tank (Fig. 1.6). In operation, fluid jets behave somewhat like propellers and they generate turbulent flow axially. However, they do not themselves generate tangential flow, like propellers. Jets also may be operated simply by pumping liquid from the tank through the jet back into the tank.

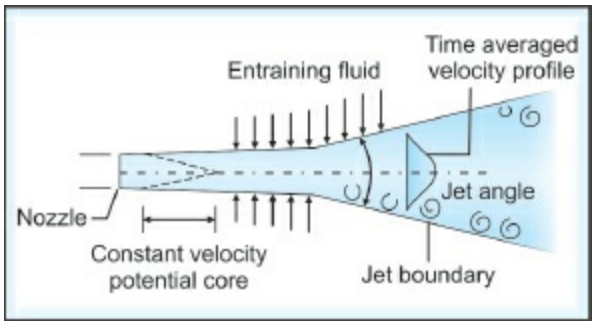
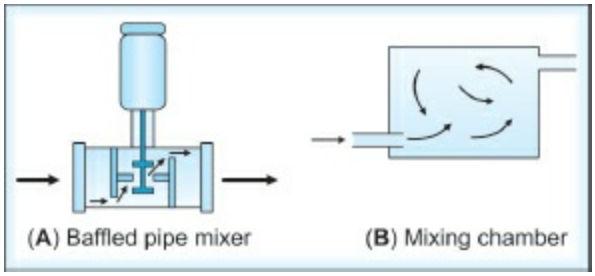



**Fig. 1.6:** Schematic representation of the fluid jet entering a mass of stationary liquid

## *Continuous or In-line Mixers*

The process of continuous mixing produces an uninterrupted supply of freshly mixed material, and is often desirable when very large volumes of materials are to be handled. It can be accomplished essentially in two ways: in a *tube or pipe* through which the material flows and in which there is very little back flow or recirculation, or in a *mixing chamber* in which a considerable amount of holdup and recirculation occur (Fig. 1.7). To ensure good mixing efficiency, devices such as vanes, baffles, screws, grids, or combinations of these are placed in the mixing tube. As illustrated in Fig. 1.7A, mixing takes place mainly through mass transport in a direction normal to that of the primary flow. Mixing in such systems requires careful control of the feed rate of raw materials if a mixture of uniform composition is to be obtained. The requirement of an exact metering in such a device results from the lack of recirculation, which would otherwise tend to average out concentration gradients along the pipe. Where suitable metering devices are available, this method of mixing is very efficient. Little additional power input over that required for simple transfer through a pipe is necessary to accomplish mixing.




**Figs 1.7A and B:** Continuous fluid mixing devices: (A) Baffled pipe mixer;

(B) Mixing chamber

When input rate is difficult to control and fluctuations in the ratio of added ingredients are unavoidable, continuous mixing equipment of the tank-type is preferred. Fluctuations in composition of the mixture are greatly reduced by the dilution effect of the material contained in the tank. For example, consider a tank of volume *V*, which is stirred so as to be perfectly mixed at all times, as illustrated in Fig. 1.8. If each increment of added material is instantaneously distributed evenly throughout the vessel, and the concentrations of equal volumes entering and leaving the mixer are designated as *C*, and $C_o$ respectively, conservation of mass requires that:

$$V\frac{dC_o}{d_t} = (C_i - C_o)\frac{dv}{dt} \qquad \ldots (2)$$

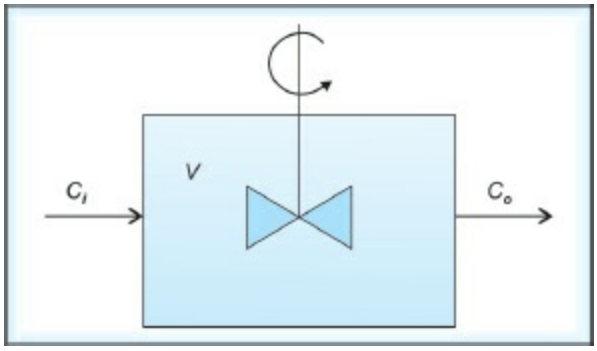



**Fig. 1.8:** Diagrammatic representation of a perfectly mixed tank in a flow stream with a flow rate, *dv/dt*. $C_i$ and $C_o$ represent the concentrations entering and leaving the tank at any given instant, respectively

where, *dv/dt* is the rate of flow of material through the tank. For a given concentration difference ($C_i - C_o$) and flow rate *dv/dt*, the rate of change of concentration of the effluent with time, $dC_o/dt$, is inversely proportional to the tank volume. Two tanks in series, each having a volume *V*/2 or half that of the single tank just discussed, would be even more effective in reducing concentration fluctuations while having the same hold-up. This is true when random fluctuations in concentration occur over small volume increments compared with the tank volumes. This is essentially a serial dilution effect.

***Example:*** When integrated, equation (2) yields the expression:

$$C_0 = C_i\,(1 - e^{-kt}) \ldots (3)$$

where, $k = d\upsilon/Vdt$. When two identical tanks, each of volume $V/2$, are connected in series, the relationship between input and output concentrations becomes:

$$C_o = C_i\,(1 - e^{-2kt} - 2kte^{-2kt}) \ldots (4)$$

For comparison purposes, let us set $k$ equal to 0.1 $min^{-1}$, and examine the ratio of $C_o$ to $C_i$, after 5 min of operation, with the mixing tank(s) at an initial concentration of $C_o$ and with a constant inlet concentration of $C_i$,. When a single tank is used, $C_o/C_i$ equals 0.393, whereas with two tanks in series, each having one-half of the volume, $C_o/C_i$, is 0.264. This effect appears more pronounced at shorter times and less so over longer periods in relation to $k$, when $C_o$ closely approaches $C_i$,.

$$\frac{C_o}{C_i} = 1 - e^{-1/2} = 0.393$$

$$\frac{C_o}{C_i} = 1 - e^{-1} - e^{-1} = 0.264$$

An effect similar to that obtained with two tanks can be observed with a turbine-agitated tank having vertical side-wall baffles. If the turbine impeller is located near the middle of the tank, two regions of mixing occur above and below the impeller as shown in Fig. 1.9. Mass transport between these zones is relatively slow. This has the effect of two areas of rapid mixing, and the net effect of such a mixer is analogous to that obtained by the operation of two tanks, of the type shown in Fig. 1.8, in series. Complex arrays of interconnected tanks, both in series and parallel, can be used for special mixing situations. The differential equations that arise from such systems may be solved by a variety of methods depending on their form. The reader is referred to mathematical texts for the appropriate techniques. The great variety of agitation systems that may be used for continuous mixing in tanks has been discussed in connection with batch mixing.

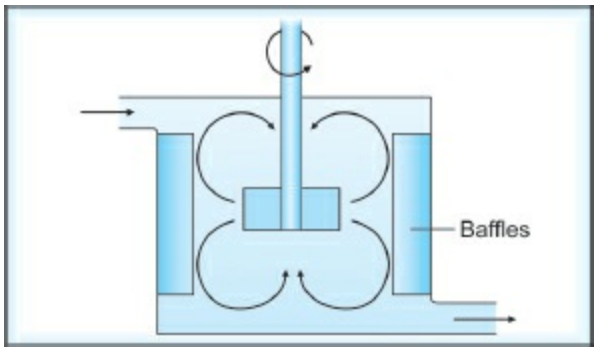



**Fig. 1.9:** Diagrammatic representation of a turbine-agitated, continuous mixing tank with vertical side-wall baffles

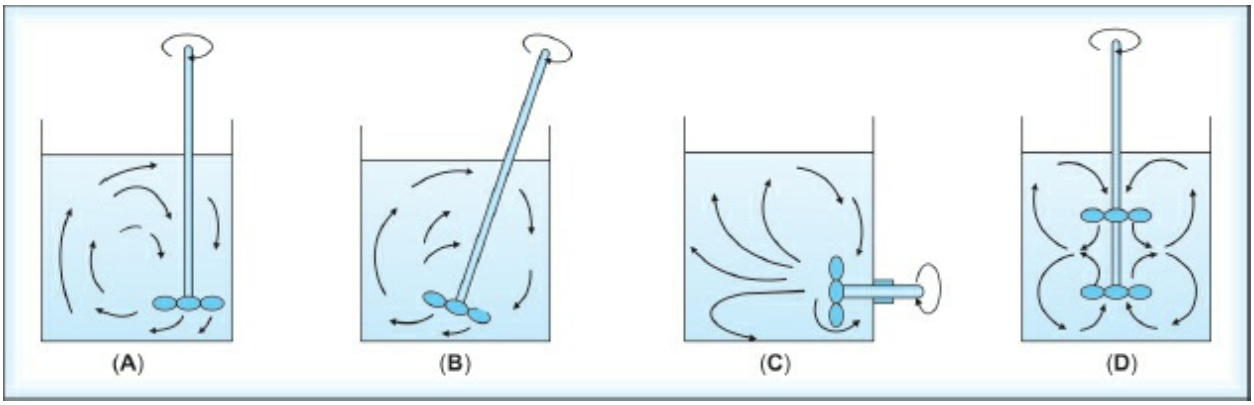
## Practical Considerations

### *Vortexing*

A vortex develops at the center of the vessel when liquids are mixed by a centrally-mounted vertical-shaft impeller. This particularly is characteristic of turbine with blades arranged perpendicular to the impeller shaft. These impellers tend to induce tangen-tial flow, which does not itself produce any mixing, except possibly near the tank walls where shear forces exist, instead, swirl and the vortex formation. This is true except at very low impeller speeds or at very high liquid viscosities (>20,000 cps), neither of which is normally encountered in practice in the pharmaceutical industry. When a vortex is formed, air is drawn into the impeller and is dispersed into the liquid, which is undesirable, as it may lead to foaming, especially if surfactants are present, and also because the full power of the impeller is not imparted to the liquid. The entrapped air also causes oxidation of the substances in certain cases and reduces the mixing intensity by reducing the velocity of the impeller relative to the surrounding fluid.

Apart from the above two problems, vortexing makes it difficult to scale-up the process, as it is impossible to achieve kinematic and geometric similitudes (kinematic similarity is achieved when flow patterns in small-and large-scale vessels are similar, whereas geometric similarity is achieved when the corresponding dimensions of the small-and large-scale vessels are in the same ratio). Vortices may be avoided by (i) changing arrangement of the impeller, (ii) changing the tank geometry, (iii) using a push-pull propeller, (iv) using baffles and (v) using diffuser ring.

Mounting an impeller either to an off-center position (Fig. 1.10A), inclined position (Fig. 1.10B) or side-entering position (Fig. 1.10C), thus destroying mixer symmetry, is the preferred method in the industry.



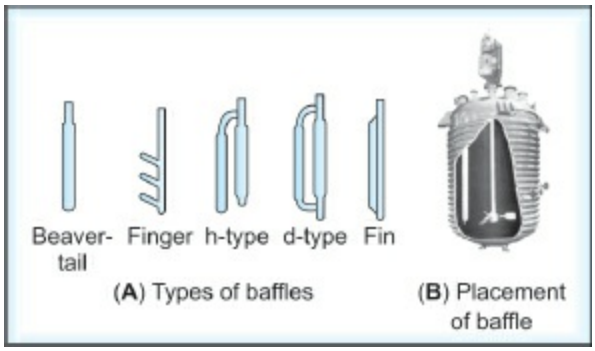

**Figs 1.10A to D:** Different arrangements of impellers in a vessel with flow pattern to prevent vortex: (A) Off-centre; (B) Inclined; (C) Side-entering; (D) Push-pull propeller

Side-entering propellers are often effectively employed. Swirl is seldom a problem with such an arrangement, as the tank geometry relative to the impeller provides a baffling effect and results in circulation of material from top to bottom in the vessel. A major drawback of such a system is the difficulty in sealing the propeller entry port. The packing around the shaft must assure a positive seal but must allow reasonably free rotation. Such a seal is also a source of contamination and may be difficult to clean. An asymmetric or angular tank geometry relative to the impeller may be used to produce an effect similar to that of baffles. Such a technique is useful in swirl prevention, but in many cases necessitates a longer mixing time than that required with a properly baffled tank of equivalent size.

This is due to the presence of regions within such tanks in which circulation is poor. Vortexing may also be avoided by the use of a push-pull propeller in which two propellers of opposite pitch are mounted on the same shaft so that the rotatory effects are in the opposite direction and thus cancel each other (Fig. 1.10D).

Alternatively, baffle plates (baffles), auxiliary devices that convert tangential flow into axial flow, may be used. Various types of baffles are commercially available with their placement depending largely upon the type of agitator used (Figs 1.11A and B). Side-wall baffles, when vertically mounted in cylindrical tanks, are effective in eliminating excessive swirl and further aid the overall mixing process by inducing turbulence in their proximity. For these reasons, the power that can be efficiently applied by the impeller is significantly increased by the use of such baffles. Vertical

movement of the fluid along the walls of the tank can be produced by arranging baffles in a steep spiral down the tank sides. It should be pointed out that if an elaborate baffle system seems necessary, the situation is best corrected by a change in the impeller design, so as to provide the desired general flow pattern. Baffles are always used in turbulent flow systems with a gap between the baffle and container wall to prevent stagnation behind the baffles. Difficulty of cleaning baffled vessels is an impediment in their use for preparing sterile products. A diffuser/stator ring that fits around the impeller can be used. This arrangement gives a high degree of turbulence which may be desirable, especially if emulsification is required. Vortexing can be avoided by operating a closed, airtight vessel to its full capacity. This method is ideal for achieving geometric and kinematic similarities. High-velocity mixing results in splashing of the liquids onto the sides of the tank above the liquid level and on the cover. Vortexing problem may result in poor reproducibility of the batches. This annoyance can be ameliorated by intermittent spraying of the continuous phase through nozzles located inside the tank.




**Figs 1.11A and B:** Various types of commercially available baffles: (A) Types of baffles; (B) Placement of baffle

# SEMISOLID MIXING

## Mixing Mechanism

The mechanism involved in mixing semisolids depends on the characteristic of the material which may show a considerable variation. When a powder and a liquid are mixed, at first they are likely to resemble closely the mixing of powders, however, at later stages the mixing mechanisms of liquids become predominant.

To the initial powder (*powder state*) when a small amount of liquid is added, the powder balls up and forms a pellet (*pellet state*), until eventually all the material is in this state. At this stage the mass has a coarse granular appearance, and the rate of attainment of homogenization is low. As the liquid content is increased further, the granular appearance is lost and the mixture becomes homogenous (*plastic state*). Plastic properties are shown and the mixture becomes difficult to shear but homogenization can be achieved rapidly. Mixing of the plastic material is facilitated by the application of shear forces. Continual incorporation of liquid causes the mixture to a attain paste like appearance (*sticky state*). The mass flows easily, even under low stresses but homogeneity is attained only slowly. Further addition of the liquid results in a decrease in consistency until a fluid state is reached (*liquid state*). In this state, the rate of homogenization is rapid and the behaviour of the mixture is described by the theory of liquid mixing as discussed previously.

## Equipments

### *Kneaders*

### *Sigma-Blade Mixer*

Sigma-blade mixer has counter-rotating blades or heavy arms that work the plastic mass. The blades rotate tangentially with a speed ratio of about 2:1. The shape and difference in rotational speed of the blades facilitate lateral pulling of the material and impart kneading and rolling action on the material (Fig. 1.12). Shear forces are also generated by the high viscosity of the mass and are thus effective in deaggregation as well as distribution of solids in the fluid vehicle.

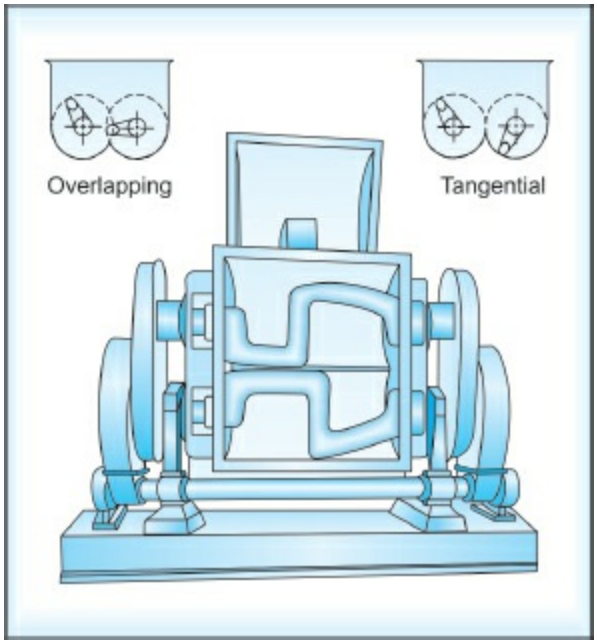



**Fig. 1.12:** Schematic representation of a top-loading sigma-blade mixer with overlapping blades

### *Planetary Mixer*

It imparts planetary mixing action, whereby the mixing element rotates round the circumference of the mixer's container, while simultaneously rotating about its own axis. The double rotation of the mixing element and its offset position reduces the dead zones and avoids vortex formation. The schematic diagram of the planetary mixer is shown in Fig. 1.13.

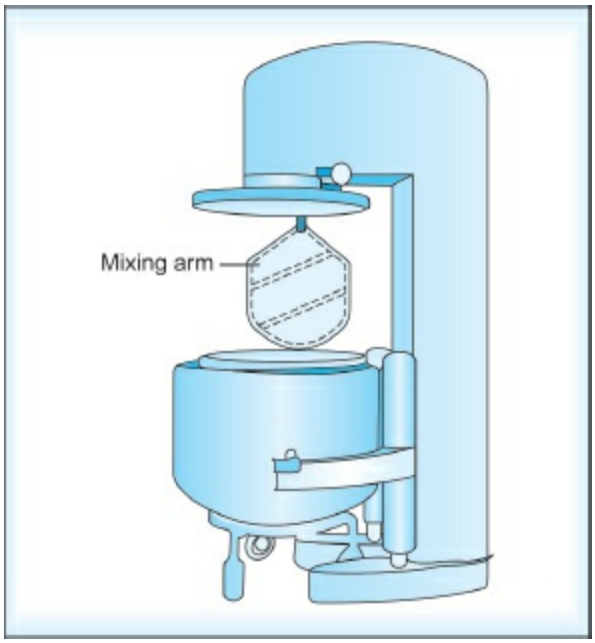



**Fig. 1.13:** Schematic representation of a planetary mixer

## *Mulling Mixers*

Mulling mixers provide forces that incorporate kneading, shearing, smearing, and blending of materials for a total uniform consistency. This process produces just enough pressure to move, intermingle and push particles into place without crushing, grinding, or distorting the ingredients. The result is a final mixture of truly uniform consistency in both physical and chemical structure. Mulling is an extension of mixing resulting from the intensification of work forces (Fig. 1.14). The work forces are applied via the tread of weighted mulling wheels. The weight, and thereby the mixing efficiency, is controlled through a spring suspension arrangement on the wheel that is fully adjustable, and allows the user to increase or decrease the amount of work that is applied to the mixture via the mulling wheel. This extension of mixing has proven to be a successful method in a wide range of applications. Mulling mixers are efficient in deaggregation of solids, but are typically inefficient in distributing the particles uniformly throughout the entire mass. These devices are suitable for mixing previously mixed material of uniform composition, but containing aggregates of solid particles. In the event of segregation during mulling, a final remixing may be necessary.

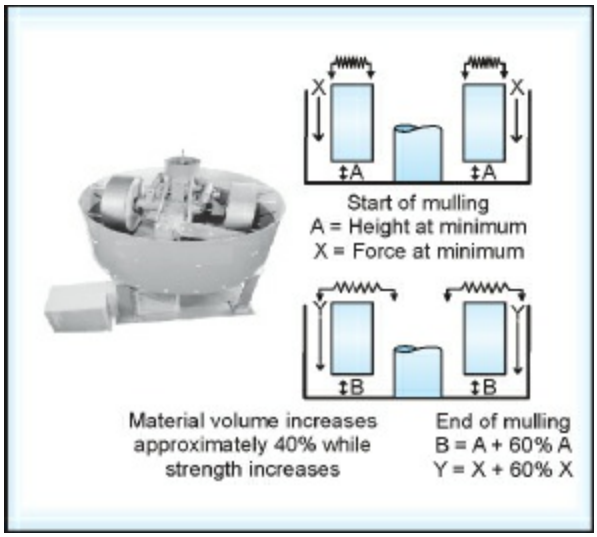



**Fig. 1.14:** Schematic representation of the mulling mixer and mulling pressure

## *Mills*

## *Roller Mills*

Roller mills consist of one or more rollers and are commonly used. Of these, the three-roller types are preferred (Fig. 1.15). In operation, rollers composed of a hard, abrasion-resistant material, and arranged to come into close proximity to each other are rotated at different rates. Depending on the gap, the material that comes between the rollers is crushed, and also sheared by the difference in rates of movement of the two surfaces. In Fig. 1.15 the material passes from the hopper, A, between rolls B and C, and is reduced in size in the process. The gap between rolls C and D, which is usually less than that between B and C, further crushes and smoothens the mixture, which adheres to the roll C. A scraper, E, is arranged so as to continuously remove the mixed material from the roller D. The arrangement is such that no material that has not passed between both the sets of rolls can reach the scraper.

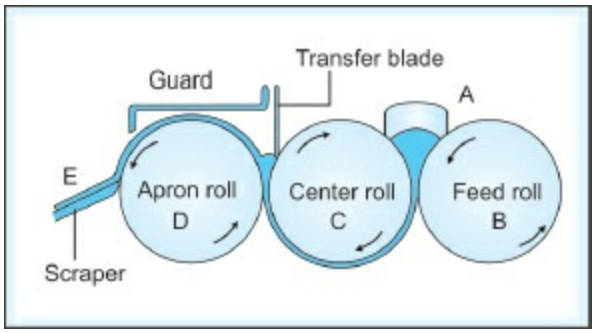



**Fig. 1.15:** Cross section of a three-roller mill

In extreme cases of solid-liquid mixing, a small volume of liquid is to be mixed with a large quantity of solids. This process is essentially one of coating the solid particles with liquid and subsequent transfer of liquid from one particle to another. In this type of mixing, the liquid is added slowly to reduce the tendency of the particles to form a lump. However, the process is not for fluids mixing, but for solids mixing. When the particles tend to stick together because of the surface tension of the coating liquid, the equipment used is the same as that for pastes. If the solids remain essentially free flowing, the equipment is the same as that used for solids mixing, which is discussed later in this chapter.

## *Colloid Mill*

A colloid mill consists of a high-speed rotor (3,000 to 20,000 rpm) and a stator with conical milling surfaces between which an adjustable clearance ranging from 0.002 to 0.03 inches is present, as indicated by the schematic diagram in Fig. 1.16. The material to be grounded should be pre-milled as finely as possible to prevent damage to the colloid mill.

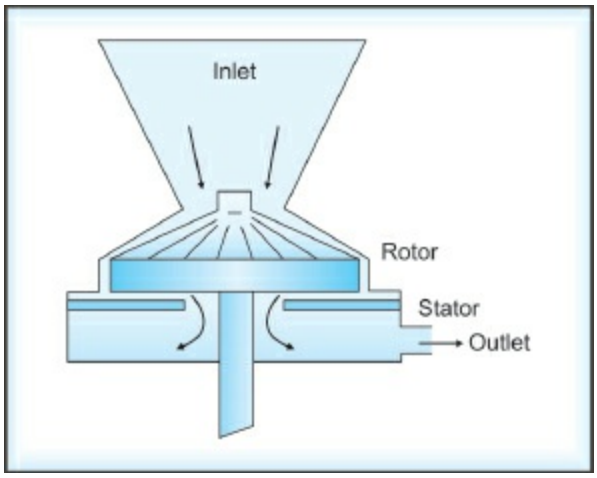



**Fig. 1.16:** Schematic representation of a colloid mill

In pharmacy, the colloid mill is used to process semisolids and not dry materials. The pre-milled solids are mixed with the liquid vehicle before being introduced into the colloid mill. Interfacial tension causes a part of the material to adhere to, and rotate with, the rotor. Centrifugal force throws a part of the material across the rotor, and onto the stator. At a point between the rotor and stator, motion imparted by the rotor ceases, and hydraulic shearing forces exceed the particle-particle attractive forces which hold the individual particles in an aggregate. The particle size of milled particles may be smaller than the clearance, because the high shear is the dispersing force. For example, in emulsification, a clearance of 75 μm may produce dispersion with an average particle size of 3 μm. The milled liquid is discharged through an outlet in the periphery of the housing and may be recycled.

Rotors and stators may be either smoothsurfaced, or rough-surfaced. With smoothsurfaced rotors and stators, there is a thin, uniform film of material between them which is subjected to maximum amount of shear. Rough-surfaced mills add intense eddy currents, turbulence, and impaction to the shearing action. Rough-surfaced mills are useful with fibrous materials because fibers tend to interlock and clog smooth-surfaced mills.

## Mixer Selection

One of the first and often most important considerations in any mixing problem is equipment selection. Factors that must be taken into consideration for appropriate mixer selection include (1) the physical properties of the materials to be mixed, such as density, viscosity, and miscibility, (2) economic considerations regarding processing, for example, the time required for mixing and power expenditure necessary and (3) cost and maintenance of the equipment. In any given case, one or more of these factors may be taken into consideration, however, the selection of equipment depends primarily upon the viscosity of the liquids, and is made according to the mechanism by which intense shearing forces can best be generated.

### *Low Viscosity Systems*

Monophasic systems of low viscosity are classified as positive mixtures, and if given time, mix completely without external agitation. Agitation reduces the time required for mixing, allowing a fast decay in the intensity of segregation. In general, for low viscosity liquids no great problems are encountered unless the operational scale is very large. The viscous character and density of the fluid(s) to be mixed determine, to a large extent, the type of flow that can be produced, and therefore, also the nature of the mixing mechanisms involved. Fluids of relatively low viscosity are best mixed by methods that generate a high degree of turbulence, and at the same time circulate the entire mass of material. These requirements are satisfied by air jets, fluid jets, and the high-speed propellers discussed earlier. A viscosity of approximately 10 poises may be considered as a practical upper limit for the application of these devices.

### *Intermediate Viscosity Systems*

The mixing of systems composed of immiscible liquids (emulsions) or finely divided solids with a liquid of low viscosity (suspensions) depends on the subdivision or deaggregation of one or more of these phases, with subsequent dispersal throughout the mass of the material to be mixed. These processes are often carried out in a single mixing operation, provided that shear forces of sufficient intensity to disrupt aggregates can be generated. At low solid-disperse phase concentrations the flow properties are Newtonian and mixing by propellers is satisfactory as long as the dispersed components oppose

settling. Under such conditions it may be desirable to increase the impeller size and decrease its speed. Emulsions and suspensions are of such viscosity that it is difficult, if not impossible to generate turbulence within their bulk, and laminar mixing, and molecular diffusion must be relied upon. Mixing of such fluids may be accomplished with a turbine of flat blade design. A characteristic feature of such impellers is the relative insensitivity of their power consumption to the density and/or viscosity of the material. For this reason, they offer a particularly good choice when emulsification or dispersion of added solids may affect these properties of the material significantly during the mixing operation. This property of turbines is due to the mechanisms by which they produce their characteristic radial flow, viz. (1) density-and viscosity-dependent fluid entrainment into the area of the blades and (2) centrifugal displacement in the axial direction. The effects of density and viscosity tend to cancel out, since they contribute in both a positive and negative way to the circulation.

When compared with a propeller of similar size, flat blade turbines of the radial flow type have a significantly lower pumping capacity, which makes them less suitable for mixing in large tanks. In case of suspensions, when deaggregation is to be carried out following a general mixing step, the high-speed turbines, frequently fitted with stators to produce increased shearing action, are often employed. Preparation of fine emulsions, whereby large globules are successively broken down into smaller ones, is accomplished by the process termed as ‘Homogenization’, as described in Chapter 2.

## *High Viscosity Systems*

Viscous ointments are efficiently mixed by the shearing action of two surfaces in close proximity, and moving at different velocities with respect to each other. This is achieved in paddle mixers, in which the blades clear the container walls by a small tolerance. Such mixers are relatively efficient, since they not only generate sufficient shear to reduce globule size, but if properly constructed, also induce sufficient circulation of the material to ensure a uniform dispersion throughout the complete mixture. The comparative mixing characteristics of the various types of impellers are shown in Table 1.1.

**Table 1.1:** Comparative mixing characteristics of various types of impellers

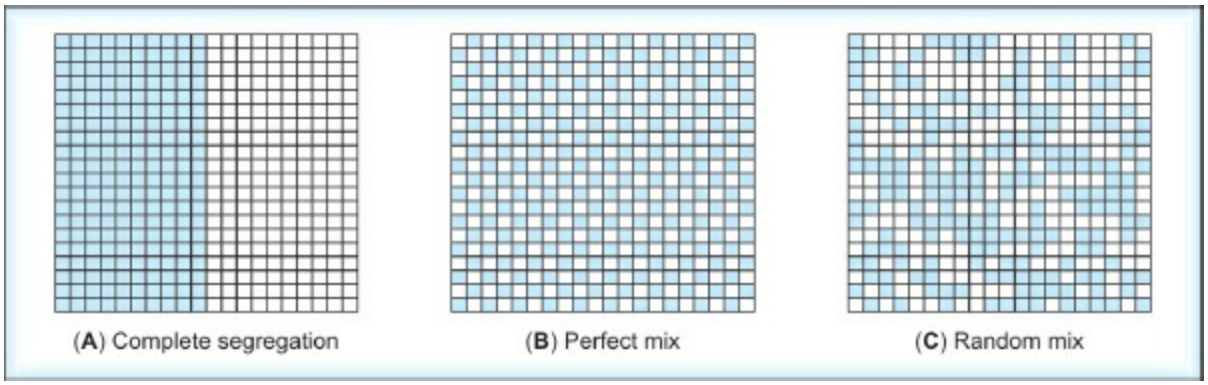
| Parameter | Propellers | Turbines | Paddles |
|---|---|---|---|
| Primary flow pattern | Axial | Radial and tangential, axial (with pitched-blade turbines) | Tangential |
| Operational speed | High, up to 8,000 rpm | Low, 50–200 rpm | Low, < 100 rpm |
| Ratio of container-to-impeller blade | Large, ~ 20 | Intermediate, ~ 2 to 3 | Small, ~ 1.1 |
| Pitch | Not constant | Constant | Constant |
| Vortex formation | Intensive | Moderate | Does not occur |
| Application | Positive mixtures: Solutions, elixirs | Negative mixtures: Suspensions, emulsions | Neutral mixtures: Gels, pastes, ointments |
| Limitations | Suitable only for liquids with low viscosity, < 5 pascal.seconds slurry of 10% solids | Suitable for liquids with moderate viscosity, ~700 pascal.seconds slurry of 60% solids | Suitable for liquids with comparative higher viscosity, > 700 pascal.seconds |

As the percentage of solids is increased, or if highly viscous fluids are employed, the solid-liquid system takes on the consistency of a paste or dough. For thicker pastes and plastic masses, a kneading, stretching and folding action is employed. The forces required to induce shear are considerable, and the equipment used is of heavy design. In such cases, sigma-blade mixer and muller mixer are the commonly used mixers. Considerable variation in rheological properties may occur during mixing, and thus a robust mixer construction is essential. The differential speed of the rolls of a roller mill induces high shear rates in the material suitable for paste mixing. With more fluid dispersions, the colloid mill may be used. Many of the mixing characteristics attributed to the various impellers, jets and other mixing equipment can be considerably altered, often unfavorably, by changes in the relative size, shape, or speed of their component parts. Although the methods of scale-up are usually considered in relation to the problem of going from laboratory scale to pilot plant to production scale.

## SOLID MIXING

The theory of solid mixing has not advanced much beyond the most elementary concepts and, consequently, is far behind the one which has been developed for fluids. This lag can be attributed primarily to an incomplete understanding of the ways in which particulate variables influence such systems and to the complexity of the problem itself. When viewed superficially, such multi-particulate solids such as pharmaceutical bulk powders or tablet granulations are seen to behave somewhat like fluids. That is, to the casual observer, they appear to exhibit fluid-like flow when they are poured from one container to another, and seem to occupy a more or less constant bulk volume. Dissimilar powders can be intimately mixed at the particulate level much like miscible liquids, at least in principle. Contrary to these similarities with fluids, however, the mixing of solids presents problems that are quite different from those associated with miscible liquids. For example, the latter, once mixed, do not readily separate and can be poured, pumped, and otherwise subjected to normal handling without the concern for demixing. In addition, they can be perfectly mixed in any standard equipment, with the primary concerns being the power efficiency and time required. In contrast, well-mixed powders are often observed to undergo substantial segregation during routine handling following the mixing operation. Such segregation of particulate solids can occur during mixing as well, and is perhaps the central problem associated with the mixing and handling of these materials.

To understand the principle of solid mixing, consider the example of mixing of two different coloured powdered components of the same size, shape and density. Fig. 1.17A depicts the two-dimensional representation of the initial unmixed or completely segregated state. From the definition of mixing, the ideal situation (Fig. 1.17B) or “perfect mix” in this case would be produced when each particle lies as closely as possible in contact with a particle of the other component. Although a perfect mixture would offer a point uniformity, such arrangement is virtually impossible to get in practice by any mixing equipment. Powder mixing, however, is a “chance” process and in practice the best type of mix likely to be obtained is a “random mix” (Fig. 1.17C), where the probability of finding one type of particle at any point in the mixture is equal to its proportion in the mixture.




**Figs 1.17A to C:** Two-dimensional representation of powder mixing: (A) Complete segregation; (B) Perfect mix; (C) Random mix

## Mixing Mechanism

Flow in tumbling mixers is two-dimensional and has been described qualitatively in terms of regimes such as slipping, avalanching, rolling, cascading, cataracting and centrifuging. The slipping regime occurs when a granular bed undergoes solid body rotation, and then slides, usually against the rotating tumbler walls. A second regime seen at a slow tumbling speed is the avalanching or slumping. In this regime, a group of powder travels down the free surface, and comes to rest before a new group is released from the above. At higher tumbling speeds, discrete avalanches give way to continuous flow at the surface of the blend. Powder beneath the cascading layer rotates nearly as a solid body with the blender until it reaches the surface. For larger tumblers, the flowing layer is thin and nearly uniform in speed and thickness, and has been modeled as depth-averaged, plug-like flow. As the rotational speed of the tumbler is increased, the surface becomes increasingly sigmoidal until the powder become airborne, and at higher speeds yet, the powder centrifuges against the tumbler wall. These regimes are termed as cascading, cataracting and centrifuging.

To model, flow and blending in complicated geometries, particle-dynamic simulations have been implemented. It has been generally accepted that in three-dimensional mixers, solid mixing proceeds by a combination of one or more mechanisms given below:

### *Convective Mixing*

This mechanism may be regarded as analogous to bulk transport as discussed in connection with fluid mixing. Depending on the type of mixer employed, convective mixing can occur by an inversion of the powder bed, by means of blades, paddles, a revolving screw, or by any other method of moving a relatively large mass of material from one part of the powder bed to another. Although convection can effectively intersperse powder in a tumbler within tens to hundreds of revolutions and is by far the fastest and most efficient mixing mechanism, it suffers from the same mixing limitations known for fluids: convective flows can, and very often do, possess barriers to mixing that prevent interaction with surrounding material.

### *Shear Mixing*

As a result of forces within the particulate mass, slip planes are set up. Depending on the flow characteristics of the powder, these can occur singly or in such a way as to give rise to laminar flow. When shear occurs between regions of different composition and parallel to their interface, it reduces the scale of segregation by thinning the dissimilar layers. The shear occurring in a direction normal to the interface of such layers is also effective, since it too reduces the scale of segregation. In addition, large or irregular grains of powder tend to be expelled from regions of high shear through a mechanism known as “shear-induced migration.” Finally, arrangements of adjacent powder grains become distorted through the influence of external strain, and this is the most classic form of shear mixing.

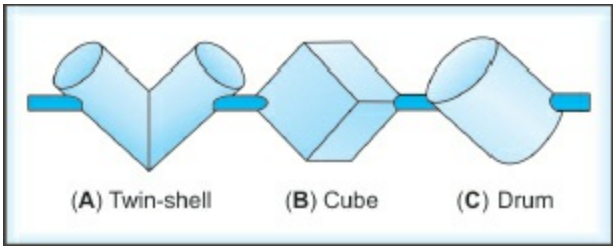
## *Diffusive Mixing*

Diffusion, also referred to as dispersion, is said to occur when random motion of particles within a powder bed causes them to change position relative to one another. Such an exchange of positions by single particles results in a reduction of the intensity of segregation. Diffusive mixing occurs at the interfaces of dissimilar regions that are undergoing shear, and can take hundreds to thousands of revolutions to act. Various stratagems, including the use of baffles, asymmetric cross-flow designs, irregular rotation protocols, or axial rocking, have been introduced to mitigate this limitation. Not withstanding these improvements, diffusion is the rate-limiting mechanism for mixing, and there is much potential for improvement of diffusive mixing. These mechanisms will be considered further in connection with the various types of mixers in common use.

## Equipments

### *Tumblers/Blenders*

A tumbling mixer consists of a container of the one of several geometric forms, which is mounted so that it can be rotated about an axis to cause movement of the material in all planes, which is necessary for rapid overall mixing. The mild forces employed, which preclude mixing materials which aggregate strongly, allow friable materials to be handled satisfactorily. The resulting tumbling motion is accentuated by means of baffles, lifter blades or simply by virtue of the shape of the container. In operation, the asymmetric geometry results in a sideways movement of material in addition to the tumbling action of the mixers. Various types of tumblers with different geometric shapes such as twin-shell, double-cone, drum, cube and tetrahedral blenders are commercially available (Figs 1.18A to C) which may be rotated about almost any axis depending on the manufacturer.




**Figs 1.18A to C:** Schematic representation of the tumbling mixers: (A) Twin-shell; (B) Cube; (C) Drum

The popular twin-shell blender is of this type and takes the form of a cylinder that has been cut in half at approximately a 45-degree angle with its long axis, and then rejoined to form a "*V*" shape. This is rotated so that the material is alternately collected at the bottom of the "*V*" and then split into two portions when the "*V*" is inverted. This is quite effective because the bulk transport and shear, which occur in tumbling mixers, generally, are accentuated by this design. A bar containing blades that rotate in a direction opposite to that of the twin-shell is often used to improve agitation of the powder bed, and may be replaced by a hollow tube for the injection of liquids.

The efficiency of tumbling mixers is highly dependent on the speed of

rotation. Rotation that is too slow neither produces the desired intense tumbling or cascading motion, nor does it generate rapid shear rates. On the other hand, rotation that is too rapid tends to produce centrifugal force sufficient to hold the powder to the sides of the mixer, and thereby, reduces efficiency. The optimum rate of rotation depends on the size and shape of the tumbler and also on the type of material being mixed, but commonly ranges from 30 to 100 rpm.

## *Agitator Mixers*

Agitator mixers employ a stationary container to hold the material and bring about mixing by means of moving screws, paddles, or blades. Since these mixers do not depend entirely on gravity as do the tumblers, they are useful in mixing solids that have been wetted, and are therefore in a sticky or plastic state. The high shear forces that are set up are effective in breaking up lumps or aggregates. Well-known mixers of this type include the following:

## *Ribbon Mixer/Blender*

It consists of a horizontal cylindrical tank usually opening at the top and fitted with helical blades or ribbons (Fig. 1.19). The blades are mounted on the horizontal axle by struts, and are rotated to circulate the material to be mixed. The helical blades are wound (in most cases) in the opposite directions to provide for the movement of material in both directions along the axis of the tank. Although little axial mixing in the vicinity of the shaft occurs, mixtures with high homogeneity can be produced by prolonged mixing even when the components differ in particle size, shape, or density, or there is some tendency to aggregate.

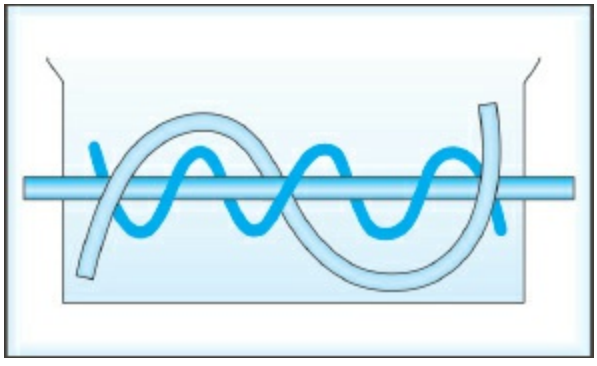


**Fig. 1.19:** Side view of a top-loading ribbon blender

## *Planetary Mixers*

They are also used for solid-solid blending although mostly in a step prior to the introduction of liquids. The detail of the equipment is discussed under semisolid mixing.

## *Nauta Mixer*

It is a vertical screw mixer which imparts three-dimensional mixing (Fig. 1.20). A screw assembly is mounted in a conical chamber, with the screw moving in a planetary motion, and also lifting the powder to be blended from the bottom to the top. The screw assembly orbits around the conical chamber wall to ensure more uniform mixing. The nauta mixer was originally designed as a powder and semisolid mixer but now-a-days also used as a mixer-granulator. The basic operation following power mixing includes the incorporation of the liquid-granulating agent, wet massing, and drying as hot, dry air is passed through the wet material. The hot air moves up through the material, which is kept in a state of motion by the orbiting screw assembly. It dries the granulation and exits through the top of the processor. If additional help is required for particular drying needs, the Nauta can be constructed to utilize vacuum drying. Accessory equipment designed to monitor and control the processor operation includes a lump breaker, which may be attached at the bottom of the conical chamber; a temperature monitor; a nuclear, non-contact density gauge; an ammeter or wattmeter; an infrared moisture analyzer; and a sampling system. Additionally, the vertical screw mixer has the advantage of requiring little floor space.

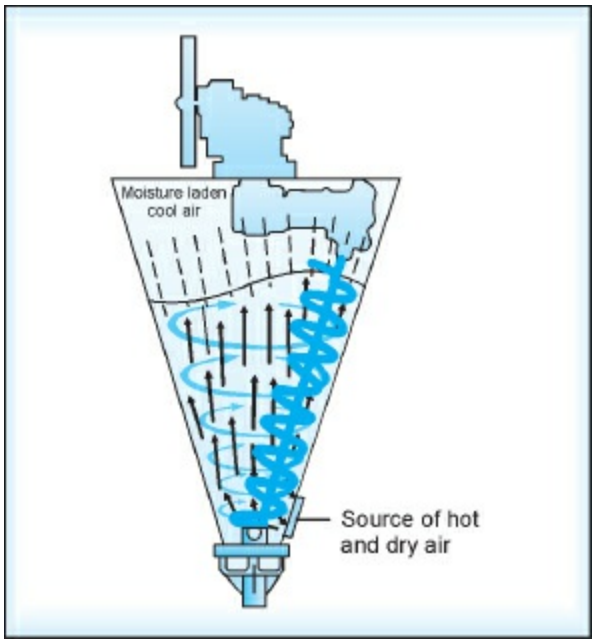



**Fig. 1.20:** Schematic representation of the Nauta mixer

## *Fluidized Air Mixer*

It is a modification of the vertical impeller. The impeller is replaced by a rapidly moving stream of air fed into the bottom of the shell. The body of powder is fluidized, and mixing is accomplished by circulation and over tumbling in the bed.

## *Processors or Rapid Mixer-Granulator*

Newer equipment has been developed that can accomplish both dry mixing and wet granulation efficiently, and in lesser time. These new mixers are classified as rapid speed mixer-granulators.

## *Lödige Mixer*

The Lödige mixer was one of the first high-shear powder blenders capable of rapidly blending pharmaceutical powders and wet massing within the same equipment. With some formulations, the equipment may also be capable of producing agglomerated granular particles that are ready for fluidized bed or other drying methods without further processing. Figure 1.21 illustrates a conventional Lödige mixer consists of a horizontal cylindrical shell equipped with a series of plow-shaped mixing tools, and one or more high-speed blending chopper assemblies mounted at the rear of the mixer. In operation,

the plow-shaped mixing tools may be revolved at variable speeds to maintain the contents of the mixer in an essentially fluidized condition, and provide a high-volume rate of transfer of material back and forth across the blender. The chopper blades are used to effectively disperse dry lumps of powder. Using this type of high-shear powder mixing equipment, complete mixing may be obtained in as little as 30 to 60.

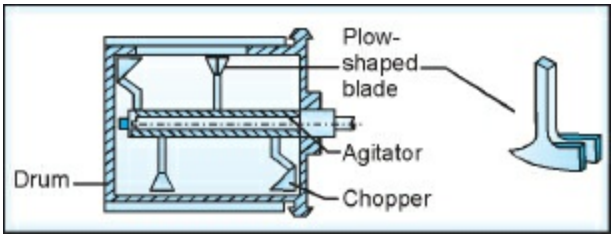



**Fig. 1.21:** Diagrammatic representation of the Lödige mixer

## *Diosna Mixer-Granulator*

It is another type of high-speed powder mixer and processor (Fig. 1.22). The mixer utilizes a bowl mounted in the vertical position, a highspeed mixer blade that revolves around the bottom of the bowl and a high-speed chopper blade, which functions as a lump and agglomerate breaker.

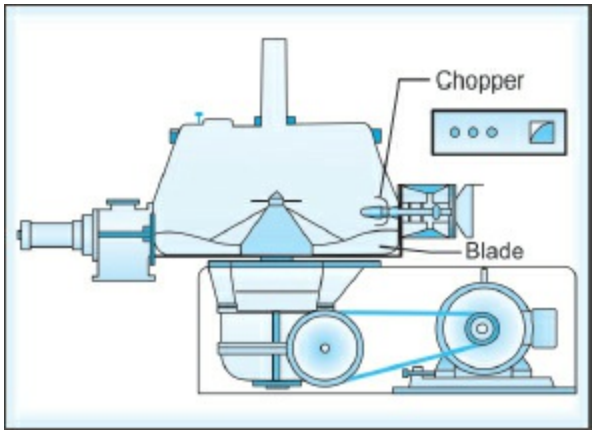



**Fig. 1.22:** Diagrammatic representation of the Diosna mixer

## *Gral Mixer-Granulator*

Figure 1.23 illustrates the Gral mixer-granulator. This equipment is a modification of the industrial planetary mixers. The difference between the Gral mixer-granulator and a standard planetary mixer is that the new unit

contains two mixing devices. A large mixing arm is shaped to the rounded configuration of the bowl and provides the large-scale mixing motion to the powder. A smaller chopper blade enters off-center from the mixing arm and is located above it. The larger mixing blade and a secondary chopper blade system is therefore similar to the Lödige and Diosna units previously described. The advantage of the unit is that the main mixing blade is not a part of the bowl, thus making clean-up easier.

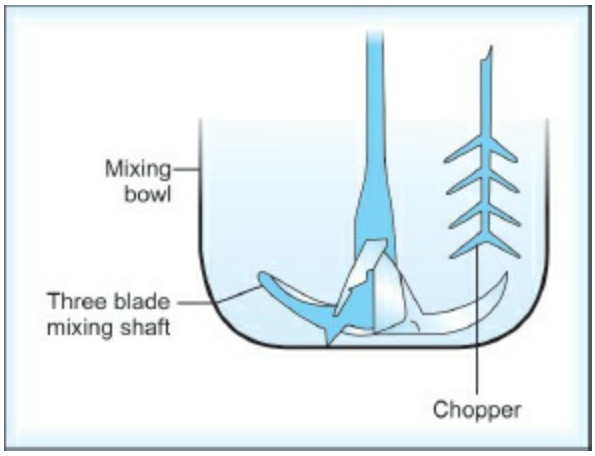



**Fig. 1.23:** Diagrammatic representation of the Gral mixer-granulator

## *Continuous Mixers*

A characteristic of solids mixing equipment is that all else being equal, mixtures produced by large mixers have greater variations in composition than those produced by small mixers. This is an important consideration when relatively small portions of the mixture are required to fall consistently within a narrow composition range. The production of tablets and capsules are examples of pharmaceutical processes in which composition uniformity is critical. The effective volume of a solid blender may be reduced considerably by the use of continuous mixing equipment. Continuous mixing processes are somewhat analogous to those discussed under fluid mixing. Metered quantities of the powders or granules are passed through a device that reduces both the scale and intensity of segregation, usually by impact or shearing action.

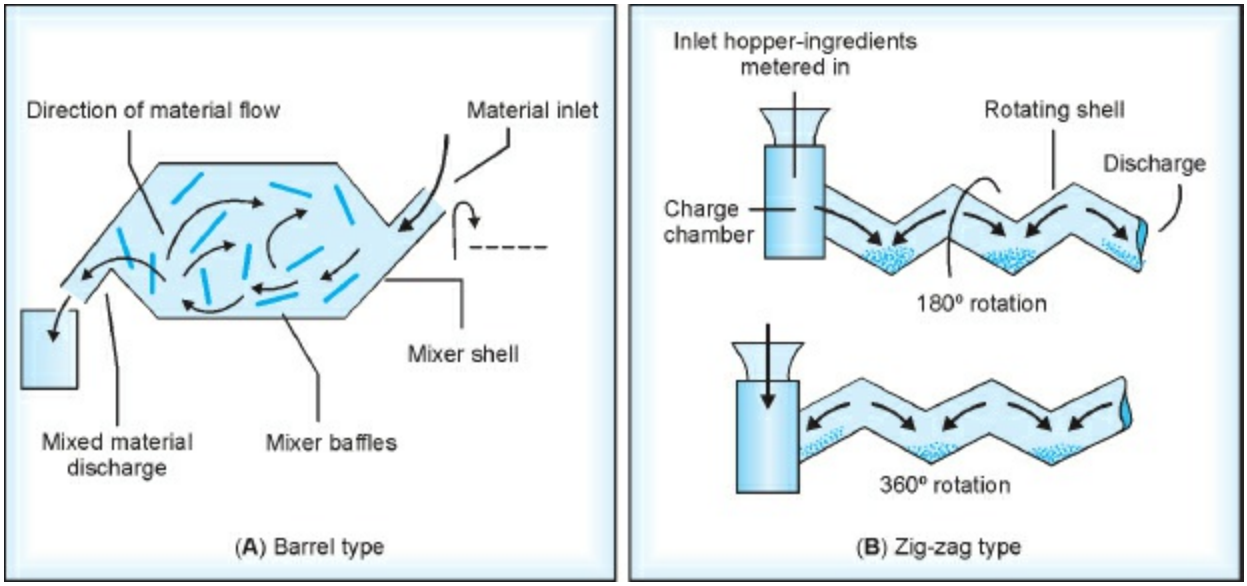
## *Blendex*

It is an in-line continuous processing device with no moving parts. It consists

of a series of fixed, flow-twisting or flow-splitting elements designed for blending of free-flowing solids. Four pipes interconnected with successive tetrahedral chambers are constructed to operate in a vertical plane. The number of chambers needed depends on the quality of mix desired. The powders fall freely through the mixer from overhead hopper and are mixed by interfacial surface generation. This type provides efficient mixing for a variety of solids without heat generation or particle size reduction.

## *Barrel Type Continuous Mixer*

In this mixer, the material is mixed under tumbling motion (Fig. 1.24A). The presence of baffles further enhances the mixing. When the material approaches the midpoint of the shell, a set of baffles causes a part of the material to move backwards. Such a mechanism provides intense mixing of ingredients. The movement of the material in this type of mixer is shown in Fig. 1.24A.




**Figs 1.24A and B:** Schematic representation of continuous mixers: (A) Barrel type; (B) Zig-zag type

## *Zig-zag Continuous Blender*

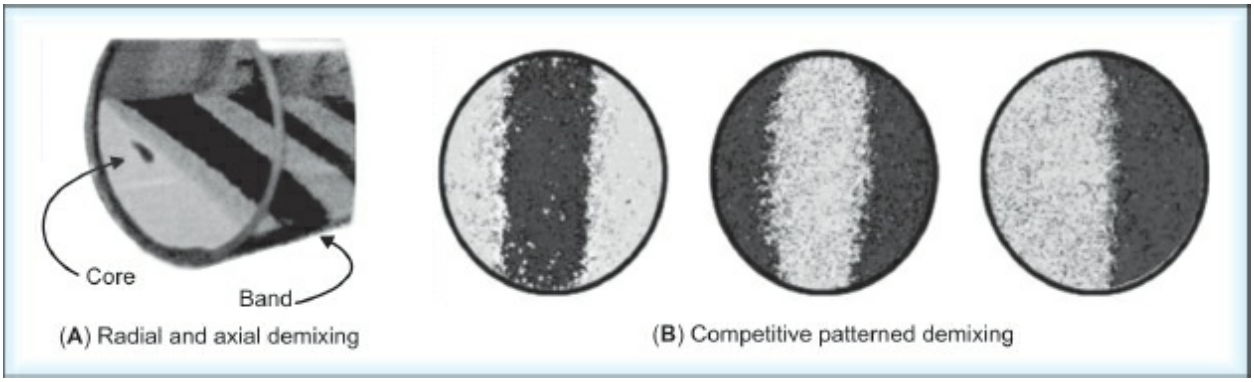
It consists of several "V"-shaped blenders connected in series (Fig. 1.24B). When the blender is inverted, the material splits into two portions, one-half of the material moves backwards, while the other moves forward. In each

rotation, a part of the material moves towards the discharge end.

## Practical Considerations

### *Segregation or Demixing*

Processing the blends of dissimilar powder grains almost invariably promotes segregation, also referred to as demixing, which is characterized by the spontaneous emergence of regions of non-uniform composition. Solids tend to segregate by virtue of differences in the size, density, shape, and other properties of the particles of which they are composed. The second requirement for segregation can be met by the Earth's gravitational field, or by a centrifugal, electrical, magnetic field generated in the course of processing. Even in the absence of such fields, this requirement can be satisfied by a gradient in shear rate within the powder bed. The process of segregation occurs during mixing as well as during subsequent handling of the completed mix. Powders that are not free-flowing or that exhibit high force of cohesion or adhesion are less susceptible to segregation. In practice, the problem of segregation is most severe when one is working with free-flowing, cohesionless, or nearly cohesionless particulate matter. Segregation has been attributed to various types of mixers: those that generate principally convective motion have been classified as "nonsegregating," while those that produce shear or diffusive mixing are termed "segregating." In practice, segregation manifests itself in granular mixing that characteristically improves over a brief period (while convection generates large-scale mixing) and then degrades, often dramatically (as the slower segregational fluxes take over). Demixing should not be confused with the phenomenon of overblending, which is also frequently encountered in blending applications. Overblending is associated with physical degradation of material properties, as in the case when a waxy lubricant is excessively deformed causing it to coat pharmaceutical grains and reduce their bioavailability, or when coated granules are damaged through abrasion or fracture. Three distinct types of demixing, i.e. radial demixing, axial demixing, and competitive patterned demixing are moderately well characterized in tumblers. *Radial segregation* is associated with migration of fine grains towards the center surrounded by larger grains segregating radially as identified in Fig. 1.25A.




**Figs 1.25A and B:** Typical segregation pattern seen during powder mixing: (A) Radial and axial demixing; (B) Competitive patterned demixing

A second stage of segregation occurs as powder grains in the core migrate along the tumbling axis. The result of this *axial migration* is the formation of a series of bands as shown in Fig. 1.25A. In this final state, two pure phases of material are formed, divided by sharp boundaries with very little intermixing.

*Competitive patterned demixing* has been observed in more complex, and more common, tumbler geometries. These distinct segregation patterns as depicted in Fig. 1.25B are believed to arise from a competition between surface segregation of coarse grains flowing over a radially segregated core of fine grains and interactions with the boundaries of the tumbler.

## *Factors Affecting Demixing*

### *Particle Size and Size Distribution*

In practice, a difference in the particle sizes of components of a formulation is the main cause of segregation in powder mixes. Smaller particles tend to fall through the voids between larger particles, and thus move to the bottom of the mass. This is known as *percolation segregation*. During mixing, the larger particles will tend to have greater kinetic energy (owing to their larger mass), and therefore move greater distances than smaller particles before they come to rest. This may result in separation of particles of different sizes, and is referred to as *trajectory segregation*. This effect, along with percolation segregation, accounts for the occurrence of the larger particles at the edge of a powder heap when it is poured from a container. During mixing, or when a material is discharged from a container, very small particles (dust) in a mix may tend to be ‘blown’ upwards by turbulent air currents as the mass 

tumbles, and remain suspended in the air during mixing. When the mixer is stopped or material discharge is complete, these particles will sediment and subsequently form a layer on top of the coarser particles. This is called *elutriation segregation* or *fluidization segregation* or *dusting out*. Particle size and particle size distribution are important, since they largely determine the magnitude of forces, gravitational and inertial, that can cause inter-particulate movement relative to surface forces, which resist such motion. As a consequence of high inter-particulate forces, as compared with gravitational forces, few powders of less than 100 μm mean particle size are free-flowing. It should be remembered that the particle size distribution and particle shape may change during processing (due to attrition, aggregation etc.), and thus the tendency to segregate. Segregation problem due to particle size difference can be rectified by:

- Selection of a particular sized fractions (e.g. by sieving to remove fines or lumps) to achieve drug and excipients of the same narrow particle size range.
- Milling of components to either reduce the particle size range or to ensure all that particles are below approximately 30 μm, at which size segregation does not tend to cause serious problems.
- Granulation of the powder mix (size enlargement) so that large numbers of different particles are evenly distributed in each segregating unit/granule.

## *Particle Density*

If components are of different densities, the denser particles will have a tendency to move downwards even if their particle sizes are similar. Trajectory segregation may also occur with particles of the same size but different densities due to their difference in mass. Often, the materials used in pharmaceutical formulations have similar densities and density effects are not generally too important.

- Selection of excipients which have densities similar to the active component(s) reduces the chances of demixing.
- Reduction of particle size of denser components nullifies the effect of density difference.

## *Particle Shape*

In moving from one location to another, relative to its neighbours, a particle must surmount certain potential energy barriers which arise from forces resisting movement insofar as neighboring particles must be displaced. This effect is a function of both particle size and shape, and is most pronounced when high packing densities occur. Particle shape is important because as the shape of a particle deviates more significantly from a spherical form, the free movement it experiences along its major axis also decreases. Spherical particles exhibit the greatest flowability, and are therefore more easily mixed, but they also segregate more easily than non-spherical particles. Irregular or needle-shaped particles may become interlocked, decreasing the tendency to segregate once mixing has occurred. Non-spherical particles also have a greater surface area-to-weight ratio (specific surface area), which will tend to decrease segregation by increasing any cohesive effects (greater contact surface area) but will increase the likelihood of 'dusting out'. Particle shape is perhaps the most difficult variable to describe and is commonly expressed by scalar quantities known as shape factors. When applied to solid mixing, shape factors provide a number index to which mixing rate, flow rate, segregation rate, angle of repose, and other static or dynamic characteristics can be related. As scalar quantities, shape factors serve as proportionality constants between mean particle diameters, and particle surface area and volume. They also serve to relate results of experimental particle size measurements by different methods. A typical example of a surface shape factor, a, is defined by the expression:

$$\alpha_s = \frac{S}{\sum n_i d_i^2} \qquad \ldots (5)$$

where, $S$ is the total surface area of the powder having $n_i$ particles of projected diameter $d_i$. Powders whose particles are highly irregular in shape generally exhibit large values of $\alpha_s$.

- Controlled crystallization during production of the drug/excipients to give components of a particular crystal shape or size range reduces the tendency to segregate.

## *Particle Charge*

The mixing of particles whose surfaces are nonconducting (electrically) often results in the generation of surface charges, as evidenced by a tendency of the

powder to clump following a period of agitation. Surface charging of particles during mixing is undesirable, for it tends to decrease the process of inter-particulate “diffusion.”

Unfortunately, surface charges in powder beds are not readily measurable. If the beds were electrically insulated during agitation, its net charge would be zero, whereas the intensity of charge on individual particles could be quite high. In such a system, a given particle may be singly-charged positively or negatively, multi-charged with like charges, or multiply charged with either an equal or unequal number of positive and negative charges. The net charge of a powder can be determined and is often taken as a measure of the tendency of the particles to undergo charge separation. Charge separation can be prevented by:

- Adding small amounts of surfactants to the powder, thereby increasing the conductivity of the surface.
- Mixing under conditions of increased humidity (above 40%).

According to theory, it is possible to prevent segregation by eliminating any one of the necessary conditions for its existence. Total avoidance of undesirable environmental conditions during the course of mixing and processing is virtually impossible. If a mixture gives persistent trouble regarding homogeneity, it is usually best to try to improve the characteristics of the mixture rather than the mixer. With free-flowing materials, the goal is to make all components as alike as possible in size, shape, and density (in that order).

It is sometimes possible to select pharmacologically inert excipients that have a selective affinity for an active mixture component. The particle-to-particle binding between drug and inert carrier, which results in such mixtures, can greatly improve homogeneity and stability towards separation of components. This technique, termed as ‘ordered mixing’, is most valuable when potent drugs are to be mixed in relatively low percentages. When the drug is added as a fine powder, it can be made to coat carrier particles uniformly and as a consequence, to be mixed uniformly throughout the batch. Usually, this is best accomplished by selecting an excipient that has a polarity similar to that of the drug. For example, a steroid would adhere well to lipid-like surfaces. In this case, however, the inclusion of significant amounts of waxy or fatty materials in a tablet formulation may cause disintegration or dissolution difficulties. When stability is not a problem, it is often more

practical to place the drug in a relatively dilute solution and spray it onto an inert excipient. After drying, this drug excipient mixture can be mixed with the remainder of the formulation. Appropriate selection of (i) suitable mixer volume, that allows sufficient space for dilation of the bed, (ii) mixing mechanism, that applies suitable shear force and (iii) mixing time, to attain suitable degree of mixing, is critical. Geometric dilution, wherein the first component with lesser amount is mixed with an equal quantity of second component and then more amount of second component is added in 50:50 proportions, and so on until all has been added, ensures effective mixing. Production of an 'ordered mix', also referred to as *adhesive* or *interactive mixing* and the extent to which the powder mass is subjected to vibration or movement after mixing, should be taken care of.

## *Mixing Measures*

A prerequisite to meaningful evaluation and interpretation of mixing is the use and understanding of a reliable measure of mixing. Though this concept may seem straight forward, some care needs to be exercised in its implementation.

## *Scale and Intensity of Segregation*

The quality of mixtures must ultimately be judged upon the basis of some measure of the random distribution of their components. Such an evaluation depends on the selection of a quantitative method of expressing the quality of randomness or "goodness of mixing." Danckwerts has suggested two criteria that are statistically defined and may be applied to mixtures of mutually soluble liquids, fine powders, or gases. Bulk transport, turbulent flow, and laminar flow all result in the intermingling of "lumps" of the liquids to be mixed. The shape and size of these lumps largely depend on the relative contribution of each of these mechanisms to the overall process and on the time over which mixing is carried out. Unless molecular diffusion occurs, however, the composition of the lumps varies discontinuously from one to the next. This can be altered only if molecular diffusion in the case of liquids and gases, or inter-particulate motion in the case of powders, tends to eliminate concentration gradients between adjacent lumps. On this basis, Danckwerts defined two quantities to describe the degree of mixing—namely "the *scale of segregation* and the *intensity of segregation*."

The scale of segregation is defined in a manner analogous to the scale of turbulence discussed earlier, and may be expressed in two ways: as a linear scale or as a volume scale. The linear scale may be considered to represent an average value of the diameter of the lumps present, whereas the volume scale roughly corresponds to the average lump volume. The intensity of segregation is a measure of the variation in composition among the various portions of the mixture. When mixing is complete, the intensity of segregation is zero. As the scale of segregation is reduced, with a resulting increase in interfacial area, molecular diffusion becomes significant. As pointed out earlier, diffusion is necessary for the effective reduction of the intensity of segregation to zero, at which time mixing is complete.

## *Scale of Scrutiny*

When a mixture is examined closely, regions of segregation are often found. The smallest region that can measure imperfections in a mix is the "*scale of scrutiny*" and can be measured as length, area, volume, or weight. The samples chosen must be large enough to contain sufficient particles to represent accurately the region from which they were taken, yet not so large as to obscure important small-scale variations in composition. The selection of a scale of scrutiny also depends on the ultimate use of the mixture. If the product is a tablet, then the scale of scrutiny is the weight of a tablet. Analysis of multiple samples of this size would allow prediction of tablet-to-tablet variations due to imperfect mixing. The lower limit is set by particle size. When the scale of scrutiny is large, or when the particle size is small, a large number of particles are present in a sample and the mix appears uniform.

In the simple case of a binary mixture of equal-sized particles of two different components, the statistics follow the binomial distribution having mean, $\mu$, and standard deviation, $\sigma$. Thus, the following equations apply:

$$\mu = np \qquad \ldots (6)$$

$$\sigma = \sqrt{np(1-p)} \qquad \ldots (7)$$

where, $n$ is the number of particles in the sample, and $p$ is the number fraction of particles of the component of interest in the mixture.

***Example:*** For the purpose of illustration, consider a capsule formulation consisting of a mixture of equal-sized pellets of two different compositions,

A and B. The pellets are mixed in the number ratio of 3 parts of A to 7 parts of B. The problem is to predict the variation in content of capsules containing 500 pellets each, assuming random mixing with no systematic segregation. In such a system, collecting a sample of 500 pellets at a time is equivalent to picking out 500 pellets one at a time randomly. Thus, the selection of pellets follows the binomial distribution, in which the expected composition of the sample and its expected variability in composition are given by equations (6) and (7), respectively:

$$\mu = np = (500)\ (0.3) = 150$$

$$\sigma = \sqrt{np(1-p)} = [(500)(0.3)(0.7)]^{1/2} = 10.2$$

One sees that on an average, a capsule will contain 150 type A pellets, the remaining 350 being type B. The number standard deviation calculated from equation (7) is 10.2. In a normal distribution, 68% of the measurements (in this case, the number of type A pellets per capsule) lie within plus or minus one standard deviation (10.2 pellets) of the mean (150 pellets per capsule). This means that even with perfect random mixing, approximately only 68% of the capsules would contain 150 ± 10.2, or approximately 140 to 160 pellets of type A among 500 pellets each.

It can be inferred from this example that as the number of particles in a sample is increased, the percentage variation in composition from sample-to-sample decreases, all else being equal. In evaluating the cause of problems with content uniformity related to tablets and capsules, the statistics of the sample should be considered. Calculations involving multi-component mixtures are more complicated, as has previously been mentioned, and are covered in references listed at the end of this chapter.

## *Sampling Techniques*

More important than the choice of the degree of mixing is the method of sampling emplo yed. Unless samples that accurately represent the system are taken, the most elaborate statistical analysis is worthless. To produce reliable samples, a correctly designed sampler is required. Samples may be withdrawn periodically during discharge of the mixture or may be taken directly from the mixer by a sampling “thief.” This tool is composed of two concentric tubes, one enclosing the other and can be used for free-flowing powders. The outer tube is pointed, with holes cut in corresponding positions in inner and outer tubes. The holes are opened or closed through the rotation

of the inner tube to capture the material (Fig. 1.26). A review of sampling techniques is shown in Table 1.2.

Techniques have also been developed for visualizing the interior of granular beds. These include:

- Diffusing wave spectroscopy, used to measure statistics of fluctuations.
- Positron emission tomography, in which a single powder grain is tracked during flow.
- Magnetic resonance imaging, in which magnetic moments of hydrogenated powder grains are tracked for short periods of time.
- X-ray tomography, in which a population of radio-opaque grains is tracked in the flow of interest.

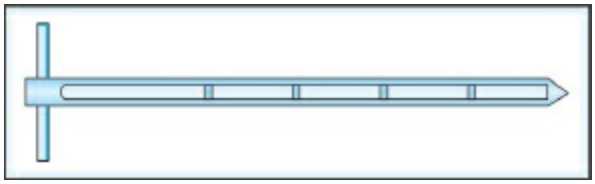


**Fig. 1.26:** Diagrammatic representation of a sampling thief

**Table 1.2:** Techniques of powder sampling

| Apparatus | Method | Advantage | Disadvantage |
|---|---|---|---|
| Sample thief | Tubular steel retains a core sample when inserted into powder bed | Good for free-flowing powder | Can be hard to push into powder, particle can fracture, segregation may occur, fines may lodge between tubes |
| Hand scoop | Cross-sectional sample | Simple, cheap | Biased results, overfilling can lead to excess of fines |
| Shovel | Pits are dug in the powder bed and samples taken from bottom and sides | Simple, cheap | Cannot be used with particles of more than 5 cm diameter |
| Auger sampler | Slotted tube rotates and worn screw carries out material | Easy to use | Biased results |

## Mixer Selection

### *Mixer Property*

An ideal mixer should produce a complete blend rapidly with as gentle mixing action as possible to avoid product damage. It should be dust-tight, cleaned and discharged easily, and require low maintenance and low power consumption. All these assets are generally not found in any single piece of equipment, thus requiring some compromise in the selection of a mixer.

Rotating shell mixers suffer from poor cross-flow along the axis. The addition of baffles or inclining the drum on the axis increases cross-flow and improves the mixing action. In cubical and polyhedron-shaped blenders, due to their flat surfaces, the powder is subjected more to a sliding than a rolling action, a motion that is not conductive to efficient mixing. In double cone blenders the mixing pattern provides a good cross-flow with a rolling rather than sliding motion. The uneven length of each shell in twin-shell blender provides additional mixing action when the powder bed recombines during each revolution of the blender. Twin-shell and double-cone blenders are recommended for precision blending.

The shearing action that develops between moving blades and trough in agitator mixers serves to breakdown powder agglomerates. Ribbon mixers are not precision blenders and also suffer from the disadvantage of being more difficult to clean than the tumblers and having a higher power requirement. The mechanical heat build-up and the relatively higher power requirement are the drawbacks also associated with sigma blade and planetary mixers. However, the shorter time interval necessary to achieve a satisfactory blend may offset these factors. Blendex provides efficient batch and continuous mixing for a wide variety of solids without particle size reduction and heat generation. Units are available to mix quantities ranging from 100 to 5000 lb/hour.

### *Material Property*

Powders that are not free-flowing or that exhibit high forces of cohesion or adhesion between particles of similar or dissimilar composition are often difficult to mix owing to agglomeration. A granular bed consisting of cohesive materials (e.g. non-tacky grains in the size range 50–300 μm

exhibits stick-slip motion so that flow becomes intermittent rather than continuous. As the size diminishes or as interparticle cohesion grows, stick-slip flow transforms mixing interfaces from smooth, regular pattern to a complex, irregular pattern. The clumps of particles can be broken down in such cases by the use of mixers that generate high shear forces or that subject the powder to impact. The use of agitators preferably planetary and sigma blade are recommended for such powders.

For particles smaller than about 100 μm, cohesive forces between particles become comparable to particle weights, and small particles can stick to one another in relatively rigid aggregates. Unless such aggregates are destroyed, the system will behave as if it had an effective particle size much larger than the primary particle size. For strongly cohesive materials, it is typically necessary to fragment agglomerates through the introduction of high shear, “intensification,” devices such as agitators or mills that energetically deform grains on the finest scale. Intensification is commonly performed in the early pre-blending stage using a fraction of the total desired excipient to avoid overblending of the final product.

## *Power Requirement*

Unlike fluids mixing, the requirements for power of a given solids mixing operation cannot be readily predicted. This is not a problem, however, since efficiency of power utilization parallels operating conditions for optimum mixing. Consequently, minimum power is that required to operate the mixer for the time necessary to reach a satisfactory steady state. Unlike most liquids mixers, solids mixers cannot be made to produce good mixtures, when they are operated incorrectly, simply by mixing for a long period of time. The process of solids mixing is accompanied by the process of segregation, as pointed out earlier, in which particles having different characteristics preferentially concentrate in various regions of the mixer. Because of this, the mixture reaches an equilibrium state of mix that is a function of speed of operation of the mixer.

The statement is often made that solid mixers result in demixing if mixing is continued for an excessive length of time. Such observations are a result of improper operation of the mixer or the use of the wrong mixer, or both. Such a mixer produces an equilibrium mixture having a significant degree of segregation. When the material is loaded into the mixer to cause an

intermingling of the various solids as they migrate towards their steady state locations, the apparent mixing-demixing phenomenon is seen.

A second cause of apparent demixing after prolonged mixing operation is the milling that inadvertently occurs because of abrasion of the particles. This frequently occurs during scaling up to production scale from small laboratory mixers. The often substantial fill weights of production mixers can generate high shear forces between particles sliding past each other under a heavy load of material above. As a consequence, the particle size distribution after mixing may bear only slight resemblance to the original distribution. An expected and common effect is the generation of fine particulate matter (fines), which can dilute lubricants and otherwise modify formulation properties.

## *Mixer Loading*

Correct loading of ingredients in a mixer is also important. The ingredients loaded in horizontal layers can be mixed relatively rapidly, while the ingredients layered side by side (either intentionally or inadvertently) will typically mix enormously more slowly.

## SCALE-UP OF MIXING

The ultimate goal of mixing research is to enable processing to be scaled-up from laboratory scale to pilot plant to production scale. An empirical approach, involving comparison of the system under study with systems of known performance, is employed for prediction of the desired operational conditions. The method is based upon dimensionless groups that characterize the mixing systems. These groups consist of combinations of the physical and geometric quantities that affect the fluid dynamics and hence, also affect the mixing performance of a given piece of equipment.

Scale-up methods are well developed in fluid mixing operations. The dimensionless group associated with fluid flow is Reynolds number, $R_e$ and is commonly defined by the expression:

$$R_e = \frac{\upsilon L \zeta}{\eta} \qquad \text{... (8)}$$

where, $\upsilon$ is the velocity of the fluid relative to the surfaces of the equipment involved. The density and dynamic viscosity are denoted by $\zeta$ and $\eta$, respectively and the dimension of length, $L$, the impeller diameter. The subgroup $\upsilon\ L\ \zeta$ is indicative of inertial forces in the system, and the Reynolds number indicates the ratio between these and the viscous forces. At high Reynolds numbers, the former predominate and the flow is turbulent, whereas at low values of $R_e$, laminar flow occurs. A transition range is known to exist since the transition from laminar to turbulent flow is not abrupt. Dimensions of the mixer and its mechanical components as well as their location within the mixer are of prime importance. Included also are impeller speed or jet pumping rate, fluid density, fluid viscosity, and height of fill of the mixer. In short, any factor that can possibly influence the behavior of the materials as they are mixed is potentially important.

In powder flow systems, in which gravitational effects occur, the Froude number, $F_r$, should be taken into account. This group is defined by the equation:

$$F_r = \frac{v_2}{gL} \qquad \text{... (9)}$$

where, $\upsilon$ and $L$ are the terms previously defined and $g$ is the acceleration due to gravity. In the case of high Froude numbers, the inertial forces

predominate over those due to gravity. This group is important whenever there is an interaction between gravitational and inertial forces.

The power that may be dissipated in a mixer by an impeller or other device is related to the power number, $P_n$:

$$P_n = \frac{p'}{v^2 \zeta} \qquad \text{... (10)}$$

where, $p$' is the pressure increment responsible for flow. The power number is thus the ratio between the forces producing flow and the inertial forces that resist it.

The power number can also be written as:

$$P_n = \frac{Pg}{\zeta \omega^3 d^5} \qquad \text{... (11)}$$

where, $P$ is the power input, $d$ is the impeller diameter, and $\omega$ is its rotational velocity.

Although a complete correlation function must take into account all the variables in a given system, satisfactory results may be obtained if only the most significant variables are considered. Therefore, while the general dimensionless equation for correlating power input contains several dimensionless groups in addition to the Reynolds and Froude numbers, these latter two quantities are usually sufficient for correlations with the power number if geometrically similar systems are investigated. The power number is thus commonly written as a function of $R_e$ and $F_r$ in the exponential form:

$$P_n = G R_e^a F_r^b \qquad \text{... (12)}$$

The exponents $a$ and $b$ and the constant $G$ must be determined experimentally. $G$ is not a universal constant since it may take on different values for different ranges in the magnitude of the associated dimensionless groups. It is reasonably constant, however, over ranges in which no gross changes in flow character occur. The exponents $a$ and $b$, which should be considered as empirical quantities, also remain remarkably constant over considerable ranges of the operating conditions.

Consider, for example, a propeller operating at a speed at which the flow is predominantly laminar. In such a system, the exponent, $a$, of the Reynolds number is found to be −1, and $b$ is zero, since vortex formation does not occur under these conditions. The correlation equation can then be written as:

$$\frac{Pg}{\zeta\omega^3 d^5} = \frac{G\eta}{\omega d^2 \zeta} \quad \text{... (13)}$$

Upon rearrangement, the functional relationship of power input to the several variables is apparent:

$$P = Gg^{-1}\,\eta\omega^2\,d^3 \text{ ... (14)}$$

Thus, power input is proportional to viscosity, and dependent on the second and third powers of the propeller velocity and diameter, respectively. The density of the fluid is not a factor under these conditions of operation.

The literature indicates that the same mixer operating under completely baffled conditions with turbulent flow can be expected to exhibit power numbers that are independent of $R_e$ and $F_r$, that is, the coefficients $a$ and $b$ in the correlation equation will both be zero. The power number is thus equal to the experimentally determined constant G, and the power input may be expressed by the equation:

$$P = Gg^{-1}\,\zeta\omega^3\,d^5 \text{ ... (15)}$$

Here, the power required for a given flow is independent of the viscosity but linearly dependent on density, in contrast to laminar flow. Also, in this case, power input is more sensitive to changes in the rotational speed and the diameter of the propeller than with laminar flow.

***Example:*** Consider a 500 L, baffled mixing vat agitated at 1,750 rpm by means of a centrally mounted propeller of 15 cm diameter. To provide more rapid mixing, the propeller rpm is doubled, and its diameter increased to 23 cm. This design change requires a more powerful drive motor, but in order to make an estimate of the increase needed, several variables must be considered. Given that the viscosity of the fluid is 1.5 poise and the specific gravity is 1.05, the Reynolds number can be estimated. If the propeller pitch is 0.8 diameters per revolution, it will pump liquid at a velocity determined by the product of the pitch and rpm. Thus, fluid velocity is given by the following:

$$v = \frac{(0.8)(15)(1750)}{(60)} = 350 \text{ cm}$$

The Reynolds number, from equation (8), is equal to:

$$R_e = \frac{(350)(15)(1.05)}{(1.5)} = 3{,}675$$

Since this is within the turbulent range, equation (15) applies. Therefore, the power required under the new conditions, $P_n$, as compared with that needed previously, $P_o$, is given by:

$$P_n = \left(\frac{\omega_n}{\omega_o}\right)^3 \left(\frac{d_n}{d_o}\right)^5 = (2)^3 \left(\frac{23}{15}\right)^5 = 67.8$$

On the basis of these calculations, it may be decided that the increased speed of mixing resulting from this design change does not warrant the additional power required.

The foregoing conclusions are valid only if geometric similarity is maintained in the mixing systems. Also, the value of $G$, while reasonably constant over the ranges of laminar flow and fully developed turbulence, is not numerically the same in these two regions.

These examples illustrate the usefulness of dimensionless groups in predicting and calculating the influence of systematic variables on the mixing process. The same general technique is also useful in correlations involving more complex systems, which require additional groups for satisfactory calculations.

# 2:
# Milling

Few materials used in pharmaceuticals exist in the optimum size, and most must be comminuted at some stage or the other during the production of a dosage form. Milling is the mechanical process of reducing the particle size of solids. Various terms (comminution, crushing, disintegration, dispersion, grinding, and pulverization) have been used synonymously with milling depending on the product, equipment, and the process. Although fine particles can be produced directly by controlled precipitation, crystallization or drying a fine spray of solution, in many cases the material is powdered in some kind of mill. Milling equipment is usually classified as coarse, intermediate, or fine according to the size of the milled product. Size is conventionally expressed in terms of mesh (number of openings per linear inch of a screen). As an arbitrary classification for the consideration of pharmaceuticals, coarse milling produces particles larger than 20-mesh, intermediate milling produces particles from 200-to 20-mesh (74 to 840 μm), and fine milling produces particles smaller than 200-mesh. A given mill may operate successfully in more than one class: a hammer mill may be used to prepare a 16-mesh granulation and to mill a crystalline material to a 120-mesh powder.

Numerous examples have been quoted to stress the importance of fine particles in pharmacy and milling or grinding offers a method by which these particles can be produced. The surface area per unit weight, which is known as the *specific surface*, is increased by size reduction. In general, a 10-fold increase in surface area has been given by a 10-fold decrease in particle size. This increased surface area affects:

**Dissolution and therapeutic efficacy:** Dissolution and therapeutic efficiency of medicinal compounds that possess low solubility in body fluids are increased due to increase in the area of contact between the solid and the dissolving fluid. The control of fineness of griseofulvin led to an oral dosage regimen half that of the originally marketed product. In inhalational products, the size of particles determines their position and retention in the

bronchopulmonary system. Transdermal delivery is also facilitated by particle size reduction and an increased antiseptic action has been demonstrated for calomel ointment when the particle size of calomel was reduced. The rectal absorption of aspirin from a theobroma oil suppository is also related to particle size. In addition to absorption, particle size may affect the texture, taste, and rheology of oral suspensions. Control of particle size and specific surface influences the duration of adequate serum concentration, rheology, and product syringeability as in the case of intramuscular suspension of procaine penicillin G.

**Extraction:** Extraction or leaching from animal glands (liver and pancreas) and crude vegetable drugs is facilitated by comminution. The time required for extraction is shortened by the increased area of contact between the solvent and the solid and the reduced distance the solvent has to penetrate into the material. The control of particle size in the extraction process provides for more complete extraction and a rapid filtration rate when the solution is filtered through the marc.

**Drying:** The drying of wet masses may be facilitated by milling, which increases the surface area and reduces the distance that the moisture must travel within the particle to reach the outer surface. In the manufacture of compressed tablets by wet granulation process, the sieving of the wet mass is done to ensure more rapid and uniform drying. Micronization and subsequent drying also increases the stability because the occluded solvent is removed.

**Flowability:** The flow property of powders and granules is affected by particle size and size distribution. The freely flowing powders and granules in high-speed filling equipment and tablet presses produce a uniform product. For suspensions of high disperse phase concentration, reduction in particle size leads to increase in viscosity.

**Mixing or blending:** The mixing or blending of several solid ingredients of a pharmaceutical is easier and more uniform if the ingredients are of approximately the same size. This provides a greater uniformity of dose. The larger the number of particles in a sample, the closer the sample will represent the overall proportions of the mixture and therefore, the sampling accuracy of mixing could be increased by increasing the number of particles, i.e. by reducing the particle size of the mix components. Solid pharmaceuticals that are artificially coloured are often milled to distribute the colouring agent to ensure that the mixture is not mottled and uniform from

batch-to-batch. Even the size of a pigment affects its colour.

**Formulation:** Lubricants used in compressed tablets and capsules function by virtue of their ability to coat the surface of the granulation or powder. A fine particle size is essential if the lubricant is to function properly. The milling of ointments, creams, and pastes provides a smooth texture and better appearance in addition to improved physical stability. Also, the sedimentation rate of suspensions and emulsions is a function of particle size and is reduced by milling.

## THEORY OF COMMINUTION

At present, there is meager basic understanding of the mechanism and quantitative aspects of milling. The mechanical behavior of solids, which under stress are strained and deformed, is shown in the stress-strain curve in Fig. 2.1. The initial linear portion of the curve is defined by Hooke's law (stress is directly proportional to strain), and Young's modulus (slope of the linear portion) expresses the stiffness or softness of a solid in dynes per square centimeter. The stress-strain curve becomes nonlinear at the yield point, which is a measure of the resistance to permanent deformation. With still greater stress, the region of irreversible plastic deformation is reached. The area under the curve represents the energy of fracture and is an approximate measure of the impact strength of the material.

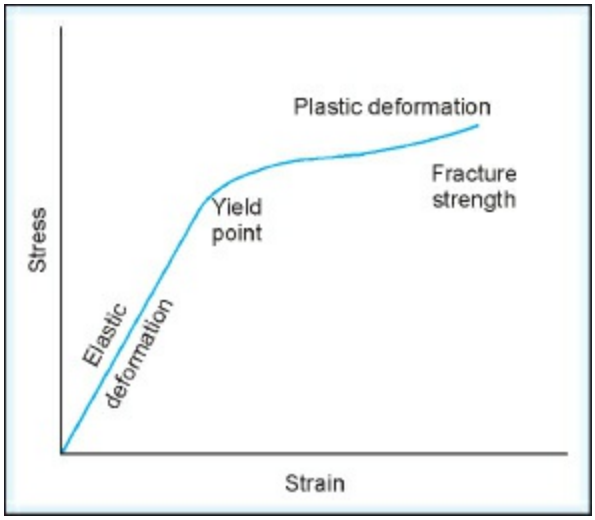



**Fig. 2.1:** Stress-strain diagram for a solid

In all milling processes, it is a random matter if and when a given particle will be fractured. If a single particle is subjected to a sudden impact and fractured, it yields a few relatively large particles and a number of fine particles, with relatively few particles of intermediate size. If the energy of the impact is increased, the larger particles are of a smaller size and greater number, and although the number of fine particles is increased appreciably, their size is not greatly changed. It seems that the size of the finer particles is related to the internal structure of the material, and the size of the larger particles is more closely related to the process by which comminution is accomplished.

Size reduction begins with the opening of any small cracks that were initially present. Thus, larger particles with numerous cracks fracture more readily than smaller particles with fewer cracks. In general, fine grinding requires more energy, not only because of the increased new surface, but also because more energy is needed to initiate cracks.

For any particle, there is a minimum energy that will fracture it, however, conditions are so haphazard that many particles receive impacts that are not sufficient to fracture them and are eventually fractured by some excessively forceful blow. As a result, the most efficient mills utilize less than 1% of the energy input to fracture particles and create new surfaces. The rest of the energy is dissipated in (1) elastic deformation of unfractured particles, (2) transport of material within the milling chamber, (3) friction between particles, (4) friction between particles and mill, (5) heat, (6) vibration and noise and (7) inefficiency of transmission and motor.

If the force of impact does not exceed the elastic limit, the material is reversibly deformed or stressed. When the force is removed, the particle returns to its original form, and the mechanical energy of stress in the deformed particle appears as heat. For polymeric materials, hysteresis is frequent. When a force is applied and then removed, an elastic loop, or hysteresis, occurs in the stress-strain cycle of the polymeric material. The area of the loop represents the dissipation of stress energy (usually as heat).

A force that exceeds the elastic limit fractures the particle. Usually, the surfaces of particles are irregular, so that the force is initially taken on the high portion of the surface, with the result that high stresses and temperatures may be set up locally in the material. As fracture occurs, the points of application of the force are shifted. The energy for the new surfaces is partially supplied by the release of stress energy. Crystalline materials fracture along crystal cleavage planes, and non-crystalline materials fracture at random. If an ideal crystal were pressed with an increasing force, the force would be distributed uniformly throughout its structure until the crystal disintegrated into its individual units. A real crystal fractures under much less force into a few relatively large particles and several fine particles, with relatively few particles of intermediate size. Crystals of pure substances have internal weaknesses due to missing atoms or ions in their lattice structures and flaws arising from mechanical or thermal stress.

A flaw in a particle is any structural weakness that may develop into a

crack under strain. It has been proposed that any force of milling produces a small flaw in the particle. The useful work in milling is proportional to the length of new cracks produced. A particle absorbs strain energy and is deformed under shear or compression until the energy exceeds the weakest flaw and causes fracture or cracking of the particle. The strain energy required for fracture is proportional to the length of the crack formed, since the additional energy required to extend the crack to fracture is supplied by the flow of the surrounding residual strain energy to the crack.

***Griffith theory:*** The Griffith theory of cracks and flaws assumes that all solids contain flaws and microscopic cracks, which increase the applied force according to the crack length and focus the stress at the atomic bond of the crack apex. The Griffith theory may be expressed as:

$$T = \sqrt{\frac{Y \epsilon}{c}} \qquad \ldots (1)$$

where, $T$ is the tensile stress, $Y$ is the Young's modulus, ε is the surface energy of the wall of the crack, and $c$ is the critical crack depth required for fracture. A linear relationship between the square of tensile strength of minerals and the critical height for drop weight impact suggests that the square of tensile strength is a useful criterion for impact fracture.

Thermodynamic treatment of the milling process has been attempted, but there is confusion about the meaning of surface tension, surface stress, and surface energy of solids. In addition, there is some questions as to whether a reversible path may be devised for a milling process. Thermodynamics have shown that the work to fracture a particle depends on surface energy, and that the yield stress depends on the rate of strain and temperature of the fluid filling the particle pore. Fracture is predicted to be more efficient at an elevated temperature.

The weakest flaw in a particle determines its fracture strength; and controls the number of particles produced by fracture. Particles with the weakest flaws fracture most easily and produce the largest particles; however, they are not necessarily easier to mill to a given size, as they may require several more stages of fracture than particles of the same size whose weakest flaw is stronger.

The immediate objective of milling is to form cracks that spread through the deformed particles at the expense of strain energy and produce fracture.

The useful work is directly proportional to the new surface area. Since the crack length is proportional to the square root of the new surface area produced, the useful work is inversely proportional to the square root of the product diameter minus the feed diameter. The energy *E'* expended in producing a new surface is:

$$E' = E\left(\frac{\sqrt{D_1}}{\sqrt{D_1} - \sqrt{D_2}}\right) \qquad \text{... (2)}$$

where, $D_1$ is the diameter of the material fed to the mill, $D_2$ is the diameter of the product discharged from the mill, and $E$ is the energy input.

The efficiency of the milling process is influenced by the nature of the force as well as by its magnitude. The rate of application of force affects comminution, as there is a time lag between the attainment of maximum force and fracture. Often, materials respond as brittle materials to fast impact and as plastic materials to a slow force. The greater the rate at which the force is applied, the less effectively the energy is utilized, and the higher is the proportion of fine material produced. As the rate of milling is increased, more energy is expended. To produce a new surface in milliseconds may require three to four times as much energy as the production of the same new surface area in seconds.

## Energy for Comminution

The energy required to reduce the size of particles is inversely proportional to the size raised to some power. This general differential equation may be expressed mathematically as:

$$\frac{dE}{dD} = -\frac{C}{D^n} \qquad \text{... (3)}$$

where, $dE$ is the amount of energy required to produce a change in size, $dD$, of unit mass of material, and where $C$ and $n$ are constants.

### *Kick's Law*

In 1885, Kick suggested that the *energy requirement, E, for size reduction is directly related to the reduction ratio* ($D_1/D_2$). Kick's theory may be expressed as:

$$E = C \ln \frac{D_1}{D_2} \qquad \text{... (4)}$$

where, $D_1$ and $D_2$ are the diameters of the feed material and discharged product, respectively.

The constant $C$ may be regarded as the reciprocal efficiency coefficient. In the engineering literature, $C = K_k f_c$, where $f_c$ is the crushing strength of the material and $K_k$ is known as Kick's constant.

- If $n = 1$, the general differential equation reduces to Kick's equation.
- Based on Kick's law, if a certain horsepower is required to mill a given weight of material from 1000 to 500 μm, the same energy would be required to reduce the size from 500 to 250 μm.
- Kick's proposal was developed on a stress-strain diagram for cubes under compression and represents the energy required to effect elastic deformation before fracture occurs.
- Kick's equation assumes that the material has flaws distributed throughout its internal structure that are independent of the particle volume.
- Experimental and theoretical values apply best to coarse milling.

### *Rittinger's Law*

In 1867, von Rittinger proposed that the *energy required for size reduction is directly proportional to the increase in specific area surface* as expressed by the following relationship:

$$E = k_1(S_2\text{-}S_1) \ldots (5)$$

where, $k_1$ denotes the relationship between the particle surface and diameter, and $S_1$ and $S_2$ are the specific surface before and after milling, respectively. In terms of particle diameters:

$$E = C'\left(\frac{1}{D_2} - \frac{1}{D_1}\right) \qquad \ldots (6)$$

In the engineering literature, $C' = K_r f_c$, where $K_r$ is known as Rittinger's constant.

- If $n = 2$ (because the surface is proportional to the square of the diameter), the solution of the general differential equation yields Rittinger's equation.
- Equation (6) applies precisely only under the conditions in which all the energy is transferred into surface energy, and the energy of comminution required per unit of surface is independent of particle size.
- Rittinger's equation is less applicable if appreciable deformation occurs.
- It is most applicable to brittle materials undergoing fine milling, in which there is minimal deformation and a rapid production of new surfaces with concomitant surface-energy absorption.
- Rittinger's theory ignores particle deformation before fracture although work is the product of force and distance.

## *Bond's Law*

In 1952, Bond suggested that the *energy required for size reduction is inversely proportional to the square root of the diameter of the product.* This may be expressed mathematically as:

$$W_t \alpha 1/\sqrt{D_2} \qquad \ldots (7)$$

where, $W_t$ is the total work of comminution in kilowatt hours per short ton of milled material, and $D_2$ is the size in micrometers through which 80% by

weight of the milled product will pass. The total work, $W_t$, if defined as the kilowatt hours per ton required to subdivide from an infinitely large particle size to a certain product of size $D_2$, is proportional to $1/\sqrt{D_2}$, since $1/\sqrt{D_1}$ is infinitely small when $D$ is infinitely large. However, $W$ is proportional to $(1/\sqrt{D_2})-(1/\sqrt{D_1})$ when $W$ is the work in kilowatt hours per ton of material to mill from $D_1$ μm to $D_2$ μm. Thus:

$$\frac{W_t}{1/\sqrt{D_2}} = \frac{W}{(1/\sqrt{D_2})-(1/\sqrt{D_1})} \quad \ldots (8)$$

If $W_t$ is called the work index, and $W_i$ is the work input required to subdivide materials from an infinitely large size to a product size of 100 μm. Then, by substitution in equation (8), we get:

$$W_t = W\left(\frac{\sqrt{D_1}}{\sqrt{D_1}-\sqrt{D_2}}\right)\sqrt{\frac{D_2}{100}} \quad \ldots (9)$$

If the work index is known, the total energy input necessary to mill with the same efficiency from any feed size to any product size measured in micrometers may be found from the following equation:

$$W = W_t\left(\frac{\sqrt{D_1}-\sqrt{D_2}}{\sqrt{D_1}}\right)\sqrt{\frac{100}{D_2}} \quad \ldots (10)$$

If $n$ = 1.5, the solution of the general differential equation yields Bond's equation:

$$E = 2C'\left(\frac{1}{\sqrt{D_2}}-\frac{1}{\sqrt{D_1}}\right)$$
$$= 2C'\sqrt{\frac{1}{D_2}\left(1-\frac{1}{\sqrt{D_1}\sqrt{D_2}}\right)} \quad \ldots (11)$$

## *Holme's Equation*

Bond's equation (10) can be written in the form of Holme's equation (12) as

$$W_t = 10W_t\left(\frac{1}{\sqrt{D_2}}-\frac{1}{\sqrt{D_1}}\right) \quad \ldots (12)$$

where, 10 $W_i$ equals *2C*.

If fracture characteristics of a material were constant over all size ranges,

and if the efficiencies of all mills were equal, then the work index would be a true constant, and the energy required to mill from any feed size to any product size could be readily calculated from one test establishing the work index. In fact, the work index is not a true constant, but a parameter that changes with shifts in the particle size distribution. Best use of the work index occurs when conditions under which the work index is determined approximate those of the final application.

## Distribution and Limit of Comminution

As discussed, the variation in size is commonly expressed as a size-frequency distribution curve. As milling progresses, the particle size-frequency distribution has a narrower range and a finer mean size. As shown in Fig. 2.2, a material with initially a monomodal size distribution develops a bimodal size distribution as milling occurs.

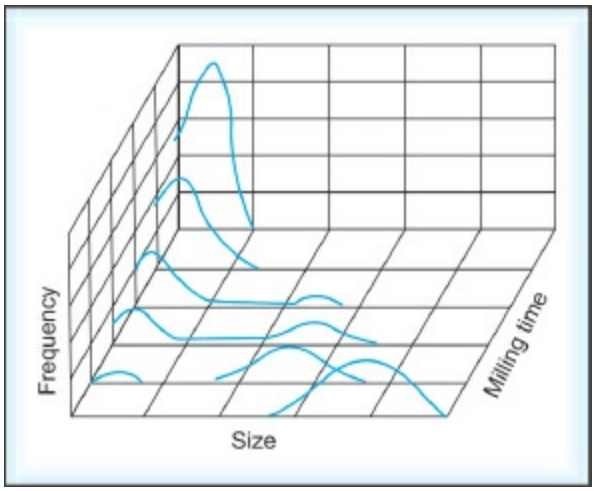



**Fig. 2.2:** Diagrammatic representation of the effect of progressive milling on particle size-frequency distribution

The primary component gradually decreases in weight, and the secondary component increases in weight. This reduction of weight is accompanied by a decrease in modal size of the primary component and is caused by preferential fracture of larger particles. The modal size of the secondary component remains essentially constant. Continued milling tends to eliminate the primary component. The process is repeated if the material is then transferred to a second mill for finer size reduction. As the particle size distribution changes, milling characteristics (work index) change because the abundance or shortage of flaws vart at different sizes.

A minimum of two specifications are necessary to characterize a specific size distribution (85% through a 60-mesh screen and 5% through a 325-mesh screen). In the simplest case, one number establishes the limits of particle sizes involved, and another number determines the weight relationships in the various size ranges. *Schuhmann* experimentally verified the empiric equation:

$$y = 100\ (D/k)^a \ldots (13)$$

where, $y$ is cumulative weight percentage smaller than size D, $k$ is the size modulus for a given distribution, and $a$ is distribution modulus. The size relationship of many fractured homogeneous materials is described by the Schuhmann equation. When the cumulative weight percentage less than a stated size is plotted on logarithm-logarithm paper against size, a straight line is obtained with a slope of $a$, which intersects the 100% ordinate line at the theoretic maximum-sized particle equal to k. For impact milling, $a \to$ and for abrasion milling, $a \to$.

In pharmacy, relatively small amounts of materials are milled, and the extent of size reduction is determined by the enhancement of clinical efficacy and product characteristics, and by the facilitation of production, rather than by energy expenditure. The proposed theories of comminution are suitable for specific applications and are to be used in a qualitative manner. The only reliable means for determining the size reduction provided by a given mill is experimental testing with the actual material.

According to the general differential equation of size reduction and its special cases, a material given sufficient time may be milled to unlimited fineness, however, in these equations, the size and energy inputs are inadequately defined. If the ultimate size reduction by mechanical means were attributed to the unit of the crystal lattice, the limit of comminution would be approximately $10^{-3}$ μm (or a specific surface of roughly $6 \times 10^7$ $cm^2/cm^3$. Milling limit refers to the size distribution to which a milling operation tends as a consequence of the mill characteristics, material properties, and operating conditions, when given sufficient time.

The milling process is affected by time. When the resident time in a mill is brief, the material is subjected to a relatively constant fracture-producing environment. Changes in milling conditions that are insignificant for short milling periods may be controlling factors in prolonged milling. In prolonged milling, the milling environment may not be constant.

As the particle becomes smaller with prolonged milling, the probability that an individual particle will be involved in a fracture diminishes. As size reduction proceeds, the mean stress required to cause fracture increases through the depletion of cracks, while the magnitude of available local stress decreases. Because of diminishing local stress and increasing aggregation,

increase in energy expenditure is useless, and size reduction reaches some practical milling limit. An empirical equation as suggested by *Harris* to express a limiting specific surface, $S_m$ is as follows:

$$S = S_m\,[1 - exp(-KE^n)] \quad \ldots 14$$

where, $E$ is energy input, and $K$ is a constant that depends on milling conditions.

As shown in Fig. 2.3, after 5 hours of ballmilling, the size reduction of sulfadimethoxine reaches a limiting value. The data fits the equation:

$$\frac{ds}{dt} = k_1 \exp(-k_2 S) \qquad \ldots (15)$$

where, $k_1$ and $k_2$ are parameters dependent on the physical properties of material and coherency, respectively.

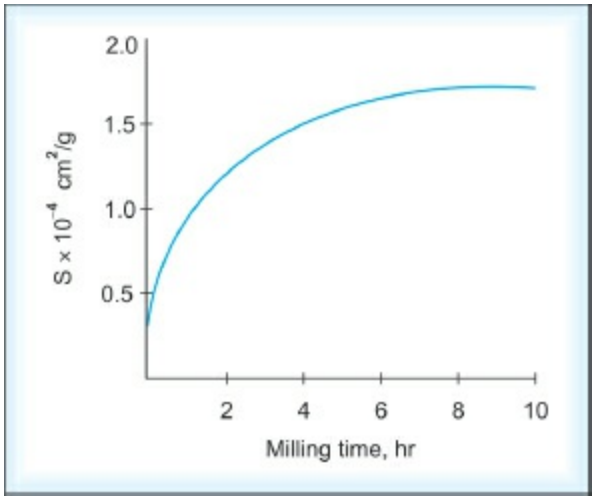



**Fig. 2.3:** Increase in surface area of sulfadimethoxine with passage of ball-milling time

The type of mill and its operation influence the milling limit. Excessive clearance between the impacting surfaces limits size reduction. In wet milling, the fineness decreases with increased viscosity, which depends on the dispersion medium, the size and concentration of particles, and the shear rate. In tumbling mills, as the particles become smaller and more numerous, friction diminishes, and the material behaves as a semisolid. Larger particles can arch and protect smaller ones from impact. Fine particles may coat the grinding medium and cushion larger particles from impact.

## Milling Rate

The mass and size of particles and the time in the mill affect the milling rate. It has been reported that batch milling of brittle materials in small mills follows the first-order law. The original particles are fractured to produce first-generation particles, which are then fractured to produce second-generation particles, which are also fractured, and so on. As this is analogous to the process of radioactive decay, milling rate is expressed in terms of a decay constant, A, which is a function of particle size and varies with the size of the material being introduced into the mill, and the size of the grinding medium in a ball mill.

If impact milling follows the first-order rate law, the number of particles, $N$ that survive fracture is the product of the initial number, $N_0$ and the probable fraction surviving fracture at time t:

$$N = N_0 \exp(-\lambda t) \ldots (16)$$

When the average mass of a particle in a given size range is constant, then $N/N_0$ is equivalent to $M/M_0$, where $M$ is the mass of particles yet unfractured and $M_0$ is the initial mass. The milling rate is given by:

$$\frac{dM}{dt} = -\lambda M \qquad \ldots (17)$$

For brief milling time, the survival probability is approximately (1-$\lambda t$), and the most probable fraction of the initial material fractured is 1-(1-$\lambda t$), or $\lambda t$. The weight of material that has been comminuted is given by:

$$M_0 - M = M_0 \lambda t \ldots (18)$$

The total mass of material in a size smaller than the largest size may increase, although the amount of original material of that size must always decrease. As particles formed from fracture of larger particles may enter the size of interest faster when the original material is fractured to a smaller size, the total mass of the size of interest is increased.

## Mechanisms of Comminution

Mills are equipments designed to impart energy to the material and cause its size reduction. There are four main methods of effecting size reduction, involving different mechanisms:

1. **Cutting:** It involves application of force over a very narrow area of material using a sharp edge of a cutting device.
2. **Compression:** In compression, the material is gripped between the two surfaces and crushed by application of pressure.
3. **Impact:** It involves the contact of material with a fast moving part which imparts some of its kinetic energy to the material. This causes creation of internal stresses in the particle, there by breaking it.
4. **Attrition:** In attrition, the material is subjected to pressure as in compression, but the surfaces are moving relative to each other, resulting in shear forces which break the particles.

## EQUIPMENTS

A mill consists of three basic parts: (1) feed chute, which delivers the material, (2) grinding mechanism, usually consisting of a rotor and stator, and (3) a discharge chute. The principle of operation depends on cutting, compression, impact from a sharp blow, and attrition. In most mills, the grinding effect is a combination of these actions. If the milling operation is carried out so that the material is reduced to the desired size by passing it once through the mill, the process is known as *open-circuit milling*. A *closed-circuit mill* is the one in which the discharge from the milling chamber is passed through a size-separation device or classifier, and the oversize particles are returned to the grinding chamber for further reduction of size. Closed-circuit operation is most valuable in reduction to fine and ultrafine size. The classification of most commonly used mills in pharmaceutical manufacturing is given in Fig. 2.4.

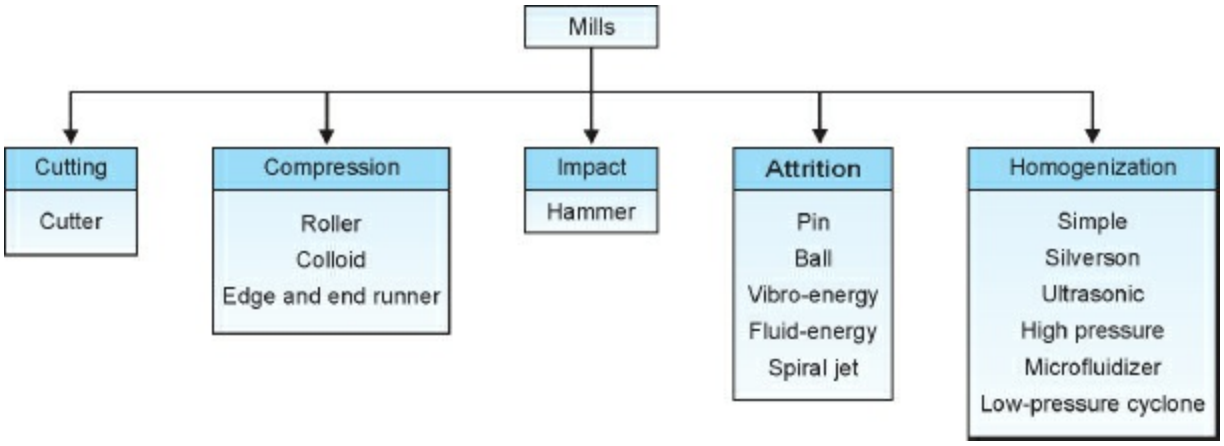



**Fig. 2.4:** Classification tree of mills

## Cutter Mill

The rotary knife cutter has a horizontal rotor with 2 to 12 knives spaced uniformly on its periphery turning from 200 to 900 rpm and a cylindrical casing having several stationary knives (Fig. 2.5). The bottom of the casing holds a screen that controls the size of the material discharged from the milling zone. The feed size should be less than 1 inch in thickness and should not exceed the length of the cutting knife. For sizes less than 20-mesh, a pneumatic product-collecting system is required. Under the best operating conditions, the size limit of a rotary cutter is 80-mesh. Cutting mills are used for tough, fibrous materials and provide a successive cutting or shearing action rather than attrition or impact. A *disc mill* consists of two vertical discs, each may be rotating in the opposite directions (double-runner disc mill), or only one may be rotating (single-runner disc mill), with an adjustable clearance. The disc may be provided with cutting faces, teeth, or convolutions. The material is pre-milled to approximately 40-mesh size and is usually suspended in a stream of air or liquid when fed to the mill.

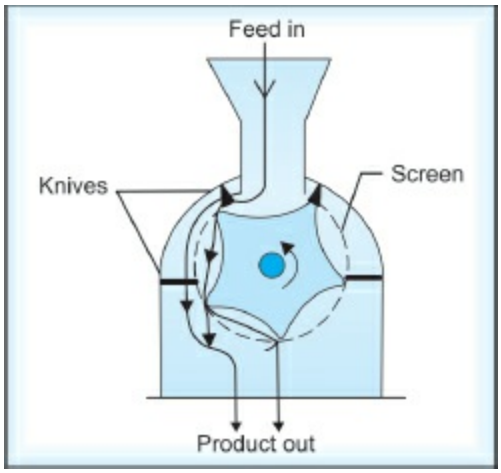



**Fig. 2.5:** Diagrammatic representation of cutter mill

## Roller Mill

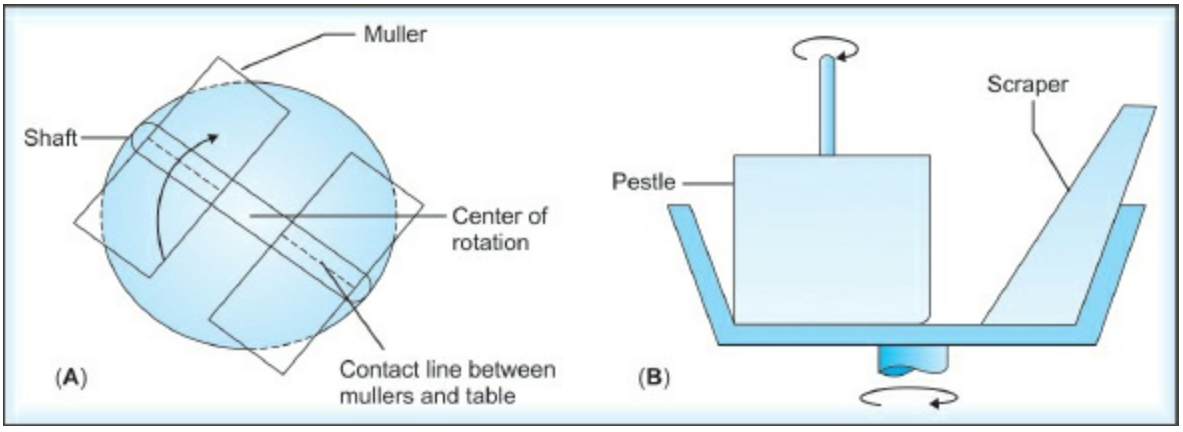
Roller mills consist of two to five smooth rollers operating at different speeds. Thus, size reduction is effected by a combination of compression and shearing action. The construction and working of a triple roller mill has been discussed in semi-solid mixing (Chapter 1).

## Colloid Mill

Colloid mills are a group of machines used for wet grinding and dispersion. They operate by shearing relatively thin layers of material between two surfaces, one of which is moving at a high angular velocity relative to the other. Although very fine dispersions can be produced, they are not, as the name implies, of colloidal dimensions. A typical colloid mill has been discussed in semi-solid mixing (Chapter 1).

## Edge-and End-runner Mill

The edge-runner mill consists of one or two heavy granite or cast iron wheels or mullers mounted on a horizontal shaft and standing in a heavy pan. The material is fed into the center of the pan and is worked outward by the mulling action. Milling occurs by compression, due to the weight of the muller, and by shearing (Fig. 2.6A). The end-runner mill is similar in principle and consists of a rotating pan or mortar made of cast iron or porcelain. A heavy pestle is mounted vertically within the pan in an off-center position (Fig. 2.6B). The mechanism of size reduction is compression due to the weight of the pestle, and shear. Both mills operate at slow speeds on a packed bed. Both produce moderately fine powders and operate successfully with fibrous materials. Wet grinding with very viscous materials, such as ointments and pastes, is also possible.




**Figs 2.6A and B:** Diagrammatic representation of runner mill: (A) Edge runner mill, (B) End runner mill

## Hammer Mill

*Construction and working*. The hammer mill is an impact mill using a high-speed rotor (up to 10,000 rpm) to which a number of swinging hammers are fixed (Fig. 2.7). The material is fed at the top or center, thrown out centrifugally, and ground by impact of the hammers or against the plates around the periphery of the casing. The clearance between the housing and the hammers contributes to size reduction.

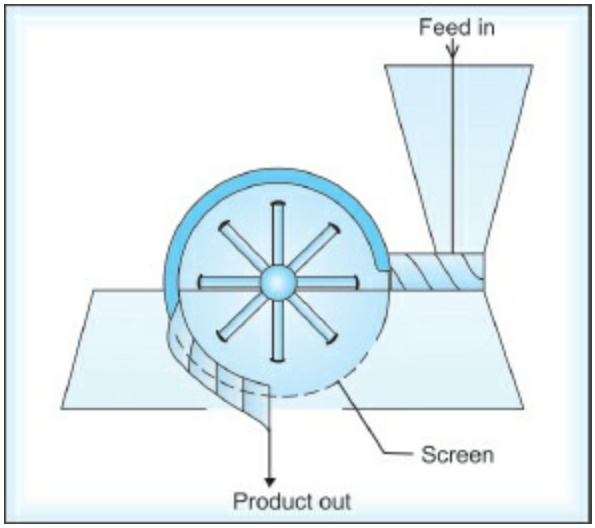



**Fig. 2.7:** Diagrammatic representation of hammer mill

The material is retained until it is small enough to fall through the screen that forms the lower portion of the casing. Particles fine enough to pass through the screen are discharged almost as fast as they are formed. A universal mill employs a variety of rotating milling elements such as a pin disk, wing or blade beater, turbine rotor, or hammer-type rotor, in combination with either a matched pin disk (that may or may not rotate), or perforated screen or stator. Some internal classification can be achieved by appropriate selection of milling tools. The particle size that can be achieved will depend on the type of milling tool selected, rotor speed (calculated as tip speed at the outermost rotating part), and solid density in the mill or solid feed rate. *Applications*. The hammer mill can be used for almost any type of size reduction. Its versatility makes it popular in the pharmaceutical industry, where it is used to mill dry materials, wet filter-press cakes, ointments, and slurries. Comminution is effected by impact at peripheral hammer speeds of

up to 7,600 meters per minute, at which speed most materials behave as if they were brittle. Brittle material is best fractured by impact from a blunt hammer and fibrous material is best reduced in size by cutting edges. Some models of hammer mills have a rotor that may be turned 180 degrees to allow use of either the blunt edge for fine grinding or the knife edge for cutting or granulating.

In the preparation of wet granules for compressed tablets, a hammer mill is operated at 2,450 rpm with knife edges, using circular or square holes of a size determined by what will pass without clogging (1.9 to 2.54 cm). In milling the dried granulation, the mill is operated at 1,000 or 2,450 rpm with knife edges and circular holes in the screen (0.23 to 0.27 cm). Hammer mills range in size from 5 to 500 horsepower units, the smaller mills being especially useful for developmental and small-batch milling.

A hammer mill can be used for granulation and close control of the particle size of powders. The size of the product is controlled by selecting the speed of the hammers, and the size and type of the screen. Speed is crucial. Below a critical impact speed, the rotor turns so slowly that a blending action rather than comminution is obtained. This results in overloading and a rise in temperature. Microscopic examination of the particles formed when the mill is operating below the critical speed shows them to be spheroidal, indicating not an impact action, but an attrition action, which produces irregularly-shaped particles. At very high speeds, there is possibly insufficient time between hammers for the material to fall from the grinding zone. In wet milling of dispersed systems with higher speeds, the swing hammers may lay back with an increased clearance. For such systems, fixed hammers would be more effective. Figure 2.8 shows the influence of speed on the particle size-frequency curves for boric acid milled at 1,000, 2,450, and 4,600 rpm in a hammer mill fitted with a screen having 6.35 mm circular holes.

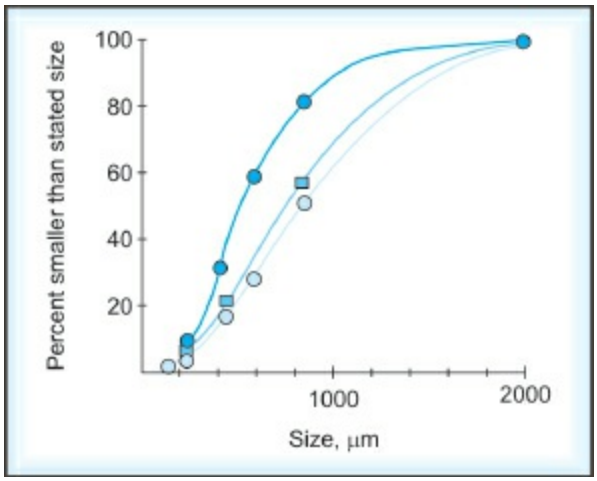
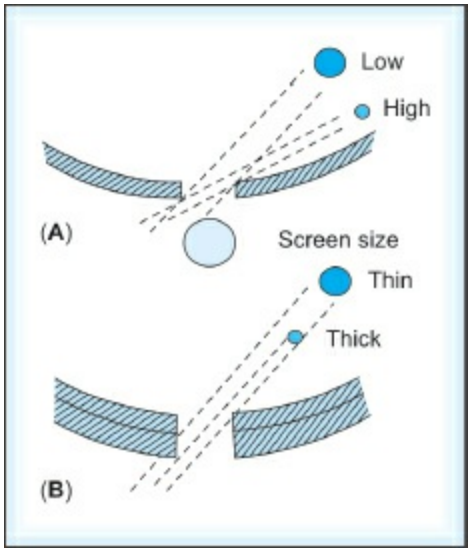



**Fig. 2.8:** Influence of speed on the size-frequency distribution of boric acid flakes milled by a hammer mill operating with impact edge forward and fitted with a round hole No. 4 screen (hole diameter: 6.35 mm). Key: ■, 1,000 rpm; ○, 2450 rpm; and ●, 4600 rpm

*Criticality*. The screens that retain the material in the milling chamber are not woven but perforated. The particle size of the discharged material is smaller than the screen hole or slot, as the particles exit through the perforations on a path approximately tangential to the rotor. For a given screen, a smaller particle size is obtained at a higher speed, as is shown in Fig. 2.9. Efforts to strengthen a screen by increasing its thickness influence particle size. For a given rotor speed and screen opening, a thicker screen produces a smaller particle, which is also illustrated in Fig. 2.9.



**Figs 2.9A and B:** In a hammer mill: (A) Particle size is influenced by speed, (B) Thickness of screen

Figure 2.9 shows the influence of screen size on the size-frequency distribution of a tablet granulation that was passed through a 4-mesh screen after wet granulation with acacia, a blend of *Terra alba*, and two active ingredients, together constituting 4.2% of the formulation. The dried granulation was milled at 2,450 rpm through a Type A plate having 1.65 mm openings, and a Type B screen having 0.84 mm openings. The granulation was also milled in a hammer mill with a vertical rotor operating at a medium speed and fitted with a number 10 screen. A comparison of the particle size-frequency distributions is shown in Fig. 2.10.

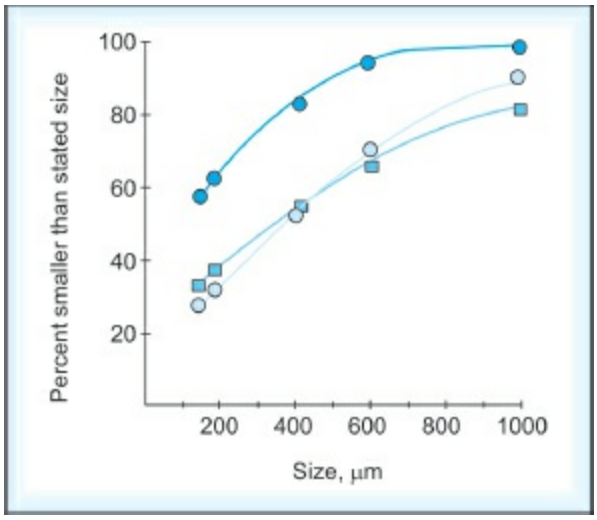



**Fig. 2.10:** Influence of screen size on the size-frequency distribution of a *Terra alba* granulation milled with a hammer mill operating at 2450 rpm, and a comparison with the granulation milled by a vertical hammer mill fitted with a no. 10 screen. Key: ○, hammer mill, 0.84 mm; ● hammer mill, 1.65 mm; and ■, vertical hammer mill

A circular hole design is the strongest screen and the most difficult to keep from clogging. It is recommended for the grinding of fibers. The herringbone design consists of a series of slotted holes repeated across the surface of the screen at an angle of 45 degrees to the length of the screen. The herringbone design is preferred for grinding crystalline materials and for continuous operation. The herringbone design with the width of the slot equal to the diameter of a round hole grinds more coarsely than the round hole. It

should not be used for fibrous materials, as it is possible for the fibers to align themselves along the slots and pass through with inadequate size reduction.

A cross slot at right angles to the path traveled by the hammer is not used in fine grinding because it clogs readily and thus a cross slot is recommended for milling slurries. The jump-gap screen is a series of bars so arranged that the particle approaches a ramp, which deflects the particle into the chamber away from the opening of the screen. The jump-gap screen is for abrasive and clogging materials.

*Advantages and disadvantages*. Hammer mills are compact with a high capacity. Size reduction of 20 to 40 μm may be achieved, however, a hammer mill must be operated with internal or external classification to produce ultrafine particles. Because the inertial forces vary with mass as the inverse cube of the diameter, small particles with a constant velocity impact with much less kinetic energy than larger ones, and the probability that particles less than a certain size will fracture decreases rapidly. In addition, small particles pass through the screen almost as fast as they are formed. Thus, a hammer mill tends to yield a relatively narrow size distribution. Hammer mills are simple to install and operate. The speed and screen can be rapidly changed. They are easy to clean and may be operated as a closed system to reduce dust and explosion hazards.

## Pin Mill

*Construction and working*. Pin mills consist of two horizontal steel plates with vertical projections arranged in concentric circles on opposing faces, and becoming more closely spaced towards the periphery (Fig. 2.11). The projections of the two faces intermesh. The feed is introduced at a controlled rate to the milling chamber through the center of the stator and is propelled through intermeshing rings of rotor and stator pins by centrifugal motion. The passage between the pins leads to size reduction by impact and attrition. The material is collected in the annular space surrounding the disks and passes to a separator. The large volumes of air drawn through the mill are discharged through the separator. The final particle size achieved in a pin mill is governed by the rotor speed, solids feed rate, and air flow rate through the mill.

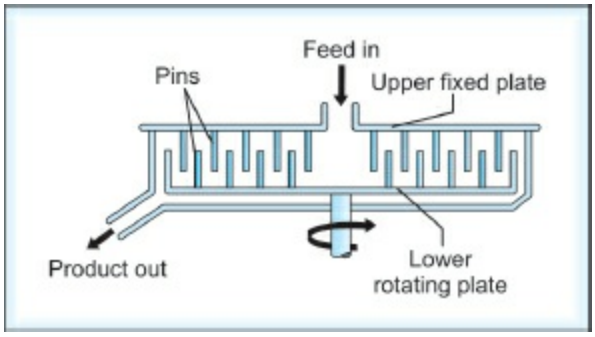


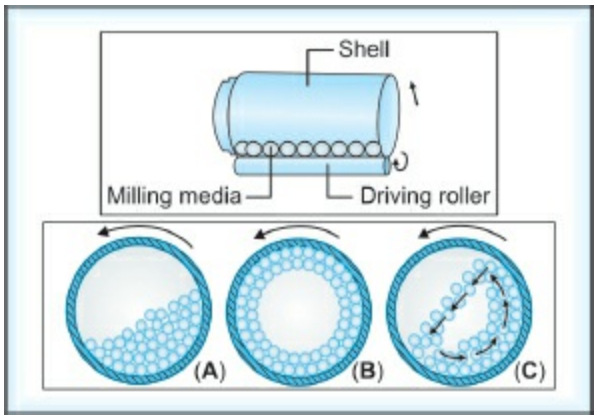

**Fig. 2.11:** Diagrammatic representation of pin mill

Smaller particles can be generated by maximizing the rotor tip speed and minimizing both product feed and air flow rate.

*Advantages and disadvantages*. Absence of screens and gratings provides a clog-free action. This type of milling is typically able to achieve smaller average particle size than wet rotor-stator milling. The machine is suitable for grinding soft, non-abrasive powders, and low milling temperatures permit heat-sensitive materials to be processed. The fineness of the grind may be varied by the use of disks with different dispositions of pins. As equipment scale is increased, maintaining rotor tip speed is one reliable way to achieve milled particle sizes comparable to small-scale results.

## Ball Mill

*Construction and working*. As depicted in Fig. 2.12, the ball mill consists of a horizontally rotating hollow vessel of cylindrical shape with the length slightly greater than its diameter. The mill is partially filled with balls of steel or pebbles, which act as the grinding medium. If pebbles are used, it is known as a pebble mill, and if rods or bars are used, it is known as a rod mill. The rod mill is particularly useful with sticky material that would hold the balls together, because the greater weight of the rods causes them to pull apart. The tube mill is a modified ball mill in which the length is about four times that of the diameter and in which the balls are somewhat smaller than in a ball mill. Because the material remains in the longer tube mill for a greater length of time, the tube mill grinds more finely than the ball mill. The ball mill may be modified to a conical shape and tapered at the discharge end. If balls of different size are used in a conical ball mill, they segregate according to size and provide progressively finer grinding as the material flows axially through the mill.




**Figs 2.12A to C:** Diagrammatic representation of ball mill and its operation: (A) Sliding at low speed, (B) Centrifuging at high speed, (C) Cascading at optimum speed

Most ball mills utilized in pharmacy are batch-operated, however, continuous ball mills are available, which are fed through a hollow trunnion at one end, with the product discharged through a similar trunnion at the opposite end. The outlet is covered with a coarse screen to prevent the loss of the balls.

*Criticality*. In a ball mill rotating at a slow speed, the balls roll and cascade over one another, providing an attrition action. As the speed is increased, the balls are carried up the sides of the mill and fall freely onto the material with an impact action, which is responsible for most size reduction (Fig. 2.12). Ball milling is a combination of impact and attrition. If the speed is increased sufficiently, the balls are held against the mill casing by centrifugal force and revolve with the mill. The critical speed of a ball mill is the speed at which the balls just begin to centrifuge with the mill. Thus, at the critical speed, the centrifugal force is equal to the weight of the ball, and the critical angular velocity, $\omega_c$, may be expressed as:

$$\omega_c = \sqrt{\frac{g}{r}} \qquad \text{... (19)}$$

where, $r$ is the radius of the ball mill. For example, a ball mill 1.2 m in diameter is run at 48 rpm and is found to be milling unsatisfactorily. The critical angular velocity of the mill is:

$$\omega_c = \sqrt{\frac{9.8}{0.6}}$$

$$\omega_c = 4.04 \text{ radians per second}$$

The actual angular velocity of the mill is $2p$ (48/60) or 5.02 radians per second. Therefore, the speed of rotation is too high, and the balls are being carried around in contact with the walls with little relative movement. If 0.6 $\omega_c$ is selected, 0.6×4.04 or 2.42 radians per second would be the angular velocity, which is equivalent to 60×2.4/2π or 23 rpm. This is half of the speed of the unsatisfactory operation. At and above the critical speed, no significant size reduction occurs. The critical speed, *nc* is given by the equation:

$$n_c = \frac{76.6}{\sqrt{D}} \qquad \text{... (20)}$$

where, $D$ is the diameter of the mill in feet. A larger mill reaches its critical speed at a slower revolution rate than a smaller mill (a 228.6 cm ball mill and a 11.4 cm jar mill may have critical speeds of 28 and 125 rpm, respectively).

Ball mills are operated from 60 to 85% of the critical speed. Over this range, the output increases with the speed. However, the lower speeds are for finer grinding. An empirical rule for the optimum speed of a ball mill is as follows:

$$n = 57–40 \log D \quad \text{... (21)}$$

where, $n$ is the speed in revolutions per minute and $D$ is the inside diameter of the mill in feet. In actual practice, the calculated speed should be used initially in the process and modified as experience is acquired.

For a given feed, smaller balls give a slower but finer grinding. The smaller balls provide smaller voids than the larger balls, consequently, the voids through which material can flow without being struck by a ball are less, and the number of impacts per unit weight of material is greater. It has been suggested that the optimum diameter of a ball is approximately proportional to the square root of the size of the feed:

$$D_{ball}^2 = kD \quad \text{... (22)}$$

where, $D_{ball}$ and $D$ are the diameters of the ball and the feed particles, respectively. If the diameters are expressed in inches, $k$ may be considered to be a grindability constant varying from 55 for hard to 35 for soft materials. Small balls facilitate the production of fine materials, but they are ineffective in reducing large-sized feed.

The charge of balls can be expressed in terms of percentage of volume of the mill (a bulk volume of balls filling one-half of a mill is a 50% ball charge). To operate effectively, a ball charge from 30 to 50% of the volume of the mill is required.

The amount of material to be milled in a ball mill may be expressed as a material-to-void ratio (ratio of the volume of material to that of the void in the ball charge). The efficiency of a ball mill is increased as the amount of material is increased until the void space in the bulk volume of ball charge is filled and then, the efficiency of milling is decreased by further addition of material.

Increasing the total weight of balls of a given size increases the fineness of the powder. The weight of the ball charge can be increased by increasing the number of balls, or by using a ball composed of a material with a higher density. Since optimum milling conditions are usually obtained when the bulk volume of the balls is equal to 50% of the volume of the mill, variation in weight of the balls is normally effected by the use of materials of different densities. Thus, steel balls grind faster than porcelain balls, as they are three times denser. Stainless steel balls are also preferred in the production of

ophthalmic and parenteral products, as there is less attrition and less subsequent contamination with particulate matter.

In dry milling, the moisture should be less than 2%. With batch processing, dry ball milling produces a very fine particle size. With wet milling, a ball mill produces 200-mesh particles from slurries containing 30 to 60% solids. From the viewpoint of power consumption, wet grinding is more efficient than dry grinding. A slower speed is used in wet milling than in dry milling to prevent the mass from being carried around with the mill. A high viscosity restricts the motion of the grinding medium, and the impact is reduced. With 1.27 cm steel balls, a viscosity from 1,000 to 2,400 centipoises (cp) is satisfactory for wet milling.

Wetting agents may increase the efficiency of milling and physical stability of the product by nullifying electrostatic forces produced during comminution. For those products containing wetting agents, the addition of the wetting agent at the milling stage may aid size reduction and reduce aggregation.

*Advantages and disadvantages*. In addition to being used for either wet or dry milling, the ball mill has the advantage of being used for batch or continuous operation. In a batch operation, unstable or explosive materials may be sealed within an inert atmosphere and satisfactorily ground. Ball mills may be sterilized and sealed for sterile milling in the production of ophthalmic and parenteral products. The installation, operation, and labor costs involved in ball milling are low. Finally, the ball mill is unsurpassed for fine grinding of hard and abrasive materials.

## Vibro-energy Mill

To overcome the limitations (low milling rates and long milling times) of a ball mill, mills that work on the principle of vibration are designed (Fig. 2.13). A simple form of a vibratory mill consists of an annular grinding chamber accommodating cylinders, which align coaxially in a three-dimensional vibratory field to give close packing and line contact between the moving surfaces. The mill body is supported on springs which permit an oscillatory movement. This vibration is usually, but not necessarily, in a vertical plane. The resultant chattering of the mill gives comminution by attrition. In additional attrition is induced by porcelain or stainless steel balls used as grinding media. The more efficient use of the energy applied and the shorter grinding times usually result in lower milling temperatures and products with narrow particle size distributions than are found in a ball mill. Furthermore, the vibratory mill is usually more flexible than the ball mill. Construction, however, is more complex and the mill is not suitable for grinding resilient materials which cannot be ground by impact since the shear forces developed are less than those found in a ball mill.

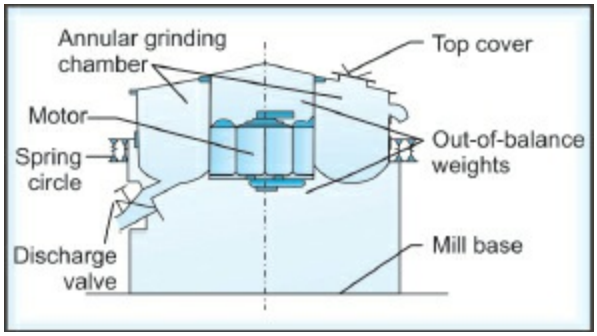



**Fig. 2.13:** Diagrammatic representation of vibro-energy mill

## Fluid-energy Mill

*Construction and working*. In the fluid-energy mill or micronizer, the material is suspended and conveyed at high velocity by air or steam, which is passed through nozzles at pressures of 100 to 150 pounds per square inch (psi). The violent turbulence of the air and steam reduces the particle size chiefly by interparticular attrition. Air is usually used because most pharmaceuticals have a low melting point or are thermolabile. As the compressed air expands at the orifice, the cooling effect counteracts the heat generated by milling.

As shown in Fig. 2.14, the material is fed near the bottom of the mill through a venturi injector (A). As the compressed air passes through the nozzles (B), the material is thrown outward against the wall of the grinding chamber (*impact*) (C) and other particles (*attrition*). The air moves at high speed in an elliptical path carrying with it the fine particles that pass out of the discharge outlet (D) into a cyclone separator and a bag collector. The large particles are carried by centrifugal force to the periphery, where they are further exposed to the attrition action. The design of the fluid-energy mill provides internal classification, which permits the finer and lighter particles to be discharged and the heavier oversized particles, under the effect of centrifugal force, to be retained until reduced to a small size.

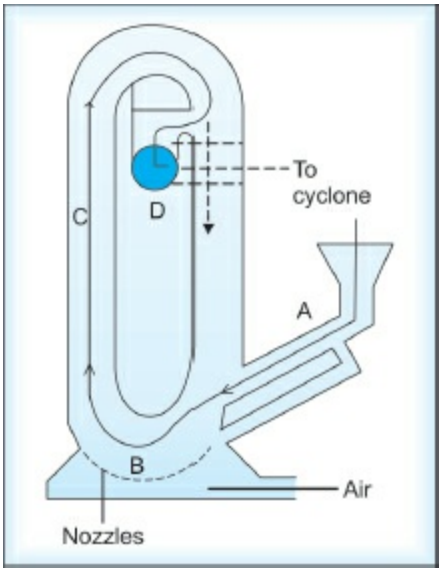



**Fig. 2.14:** Diagrammatic representation of fluid-energy mill

*Criticality*. Fluid-energy mills reduce the particle size to 1 to 20 μm. The feed should be premilled to approximately a 20-to 100-mesh size to facilitate milling. A 2 inch laboratory model using 20 to 25 cubic feet per minute of air at 100 psi mills 5 to 10 grams of feed per minute. For a given machine, size reduction depends on the size of the feed, its rate of introduction to the grinding chamber, and the pressure of the grinding fluid. The most important machine-related factors are the grinding chamber geometry and the number and angle of the nozzles. In selecting fluid-energy mills for production, the cost of a fluid-energy source and dust collection equipment must be considered in addition to the cost of the mill.

*Advantages and disadvantages*. Powders with all particles below a few micrometers may be quickly produced by this method. The disadvantage of high capital and running costs may not be so serious in the pharmaceutical industry because of the high value of the materials which are often processed. One drawback of this type of mill is the potential for build-up of compressed product in the mill or on the classifier. This can affect milled particle size by changing the open volume in the mill or open area in the classifier, especially if classifier vanes or gas nozzles become plugged or blocked.

## Spiral Jet Mill

The laboratory spiral jet mill (Fig. 2.15) comprises a flat, cylindrical milling chamber. The milling gas (air, vapor or nitrogen) with a pressure between 1–16 bars expands through 8–10 milling nozzles, tangential in the milling chamber, thus creating a spiral vortex. The feed material enters the milling chamber accelerated by an injector gas stream. The milling takes place by autogeneous impact grinding between the particles. Those particles milled fine enough leave the mill with the gas stream through the vortex finder. Spiral jet mills are used for dry fine comminution and deagglomeration of materials to particle sizes below 10 μm.

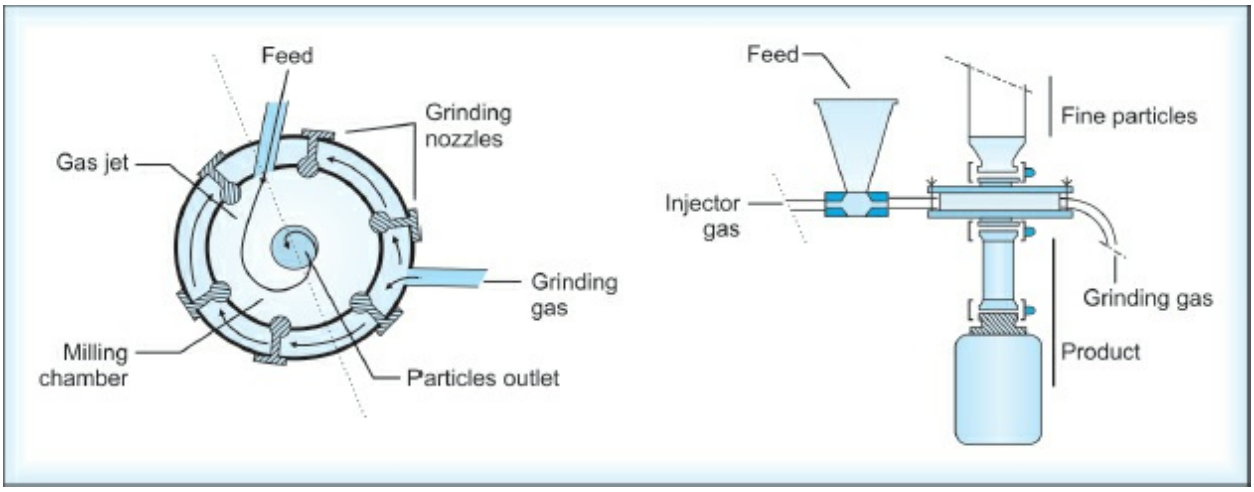



**Fig. 2.15:** Diagrammatic representation of spiral jet mill

Compared with grinding media mills, the spiral jet mill offers the advantage of (i) high fineness with narrow particle size distribution, (ii) low milling chamber temperatures, (iii) being insusceptible to dust explosions and contamination due to absence of any moving part and (iv) high mass and heat transfer rates.

## FACTORS INFLUENCING MILLING

The properties of a solid determine its ability to resist size reduction and influence the choice of equipment used for milling. The specifications of the product also influence the choice of a mill. The grindability of coal is expressed in terms of the number of revolutions of a standardized ball mill required to yield a product of which 80% passes through a 200-mesh screen. Although a similar expression could be applied to pharmaceutical materials, no quantitative scale has been adopted to express hardness.

## Nature of Material

It is perhaps implicit to speak of hard, intermediate, and soft materials. Hard materials (iodine, pumice) are those that are abrasive and cause rapid wear of mill parts immediately involved in size reduction. The physical nature of the material determines the process of comminution. Fibrous materials (*Glycyrrhiza, Rauwolfia*) cannot be crushed by pressure or impact and must be cut. Friable materials (dried filter cake, sucrose) tend to fracture along well-defined planes and may be milled by attrition, impact, or compression.

## Moisture Content

The presence of more than 5% water hinders comminution and often produces a sticky mass upon milling. This effect is more pronounced with fine materials than with larger particles. At concentrations of water greater than 50%, the mass becomes a slurry, or fluid suspension. The process is then a wet milling process, which often aids in size reduction. An increase in moisture can decrease the rate of milling to a specified product size. Glauber's salt and other drugs possessing water of crystallization liberate the water at low temperatures, causing clogging of the mill. Hygroscopic materials (calcium chloride) rapidly absorb moisture to the extent that the wet mass sticks and clogs the mill.

## Temperature

The heat during milling softens and melts materials with a low melting point. Synthetic gums, waxes, and resins become soft and plastic. Heat-sensitive drugs may be degraded or even charred. Pigments (ocher and sienna) may change their shade of color if the milling temperature is excessive. Unstable compounds and almost any finely-powdered material may ignite and explode if the temperature is high.

## Particle Shape

An impact mill produces sharp, irregular particles, which may not flow readily. When specifications demand a milled product that will flow freely, it would be better to use an attrition mill, which produces free-flowing spheroidal particles.

## Polymorphism

Milling may alter the crystalline structure and cause chemical changes in some materials. Wet milling may be useful in producing a suspension that contains a metastable form of material causing crystal growth and caking. For example, when cortisone acetate crystals are allowed to equilibrate with an aqueous vehicle, subsequent wet milling provides a satisfactory suspension.

Starch, amylose, and amylopectin may be broken down by a vibratory mill to a wide molecular weight range. Powdered povidone breaks down into lower molecular weight polymers during ball milling. Pure $C_{12}$-and $C_{16}$-fatty acids may be decarboxylated and converted to the hydrocarbon containing one less carbon atom by ball milling with wet sand. Milling well-dried microcrystalline cellulose for 1 to 25 hours decreases its crystallinity. Excessive shear of a colloid mill may damage polymeric suspending agents so that there is a loss of viscosity. A decrease in particle size of crystals in a hammer mill was reported to increase the rate of crystal growth during storage, owing to alterations in crystal lattice and formation of active sites. Specifically, crystals of phenobarbital (initial size of 310 μm) milled to 22.7 μm grew to 38.9 μm after 4 weeks at 60°C. However, crystals milled to 31.5 μm showed little growth on storage.

## Feeding Rate

The manner in which an operator feeds a mill markedly affects the product. If the rate of feeding is relatively slow, the product is discharged readily, and the amount of undersize particles or fines is minimized. If the mill is choke fed at a fast rate, the material is in the milling chamber for a longer time, as its discharge is impeded by the mass of material. This provides a greater reduction of particle size, but the capacity of the mill is reduced, and power consumption is increased. Choke feed is used when a small amount of material is to be milled in one operation.

The rate of discharge should be equal to the rate of feed, which is such that the milling parts can operate most effectively. Most mills used in pharmaceutical operations are designed so that the force of gravity is sufficient to give free discharge generally from the bottom of the mill. For ultrafine grinding, the force of gravity is replaced by a fluid carrier. A current of steam, air, or inert gas removes the product from the attrition, fluid-energy, or high-speed hammer mill. The powder is removed from the fluid by cyclone separators or bag filters.

## TECHNIQUES OF MILLING

In addition to the standard adjustments of the milling process (i.e., speed, screen size, design of rotor and load), special techniques of milling may be useful.

## Special Atmosphere

Hygroscopic materials can be milled in a closed system supplied with dehumidified air. Thermolabile, easily oxidizable, and combustible materials should be milled in a closed system with an inert atmosphere of carbon dioxide or nitrogen. Almost any fine dust (dextrin, starch, sulfur) is a potential explosive mixture under certain conditions and especially if static electrical charges result from the processing. All electrical switches should be explosion-proof, and the mill should be grounded.

## Temperature Control

As only a small percentage of the energy of milling is used to form new surface, the bulk of the energy is converted to heat. This heat may raise the temperature of the material by many degrees, and unless the heat is removed, the solid will melt, decompose, or explode. To prevent these changes in the material and avoid stalling of the mill, the milling chamber should be cooled by means of a cooling jacket or a heat exchanger. Stainless steel equipment (type 304 with number 4 finish) is routinely used in preparing pharmaceuticals because it minimizes contamination and reaction with the drugs. With the use of refrigerants, the mill must be constructed of stainless steel since cast iron becomes brittle and may shatter at low temperatures. Waxy and low-melting-point materials are chilled before milling. If this is not sufficient to embrittle the material, it may be fed to the mill simultaneously with dry ice. Stearic acid and beeswax may be reduced in a hammer mill to 100-mesh size with the use of dry ice.

## Pretreatment

For a mill to operate satisfactorily, the feed should be of the proper size and enter at a fairly uniform rate. If granules or intermediatesized particles are desired with a minimum of fines, pre-sizing is vital. Pretreatment of fibrous materials with high-pressure rolls or cutters facilitates comminution.

## Subsequent Treatment

If extreme control of size is required, it may be necessary to recycle the larger particles, either by simply screening the discharge and returning the oversized particles for a second milling, or by using air-separation equipment in a closed circuit to return the oversized particles automatically to the milling chamber. With materials to be reduced to micron size, an integrated air-separation, conveyor, and collection element are usually required.

## Dual Process

The milling process may serve simultaneously as a mixing process if the feed materials are heterogeneous. If hot gas is circulated through a mill, the mill can be used to comminute and dry moist solids simultaneously. The fluid-energy mill has been suggested as a means of simultaneous size reduction and dispersion. It has been suggested that the particles in a fluid-energy mill can be coated with almost a monomolecular film by pre-mixing with as little as 0.25% of the coating agent.

## Wet and Dry Milling

The choice of wet or dry milling depends on the use of the product and its subsequent processing. If the product undergoes physical or chemical change in water, dry milling is recommended. In dry milling, the limit of fineness is reached in the region of 100 μm when the material cakes on the milling chamber. The addition of a small amount of grinding aid may facilitate size reduction. The use of grinding aids in pharmacy is limited by the physiologic and toxicologic restrictions on medicinal products. In certain cases, the addition of ammonium salts, aluminum stearate, arylalkyl sulfonic acid, calcium stearate, oleic acid, and triethanolamine salts have been useful. These dispersing agents are especially useful in the revolving mill if coating of the balls occurs. The addition of less than 0.1% of a surface-active agent may increase the production rate of a ball mill from 20 to 40%. Wet grinding is beneficial in further reducing the size, but flocculation restricts the lower limit to approximately 10 μm. Wet grinding eliminates dust hazards, and is usually done in low-speed mills, which consume less power. Some useful dispersing agents in wet grinding are the silicates and phosphates. Comparison of dry and wet milling is summarized in Table 2.1.

**Table 2.1:** Comparison of dry and wet milling

| Characteristics | Dry milling | Wet milling |
|---|---|---|
| Medium | Air | Liquid |
| Medium property | Compressible | Incompressible |
| Force/energy transfer | Particle–particle collision, particle–equipment collision | Particle–particle collision, particle–equipment collision, liquid shear forces |
| Inertial dampening of forces | Weak to none | Strong |

## SELECTION OF A MILL

In general, the materials used in pharmaceuticals may be reduced to a particle size less than 40-mesh by means of ball, roller, hammer, and fluid-energy mills. General characteristics of the types of mills are given in Table 2.2 and the types of mills used to reduce some typical pharmaceutical materials are shown in Table 2.3. The choice of a mill is based on (1) product specifications (size range, particle size distribution, shape, moisture content and physical and chemical properties of the material), (2) capacity of the mill and production rate requirements, (3) versatility of operation (wet and dry milling, rapid change of speed and screen, safety features), (4) dust control (loss of costly drugs, health hazards and contamination of plant), (5) sanitation (ease of cleaning and sterilization), (6) auxiliary equipment (cooling system, dust collectors, forced feeding and stage reduction), (7) batch or continuous operation and (8) economical factors (cost, power consumption, space occupied and labor cost).

**Table 2.2:** Comparison of dry and wet milling

| Type of mill | Action | Product size | Used for | Not used for |
|---|---|---|---|---|
| Cutter | Cutting | 20- to 80-mesh | Fibrous, crude animal and vegetable drugs | Friable materials |
| Revolving | Attrition and impact | 20- to 200-mesh | Fine grinding of abrasive material | Soft material |
| Hammer | Impact | 4- to 325-mesh | Almost all drugs | Abrasive material |
| Roller | Compression | 20- to 200-mesh | Soft material | Abrasive material |
| Attrition | Attrition | 20- to 200-mesh | Soft and fibrous material | Abrasive material |
| Fluid-energy | Attrition and impact | 1 to 30 μm | Moderately hard and friable material | Soft and sticky material |

**Table 2.3:** Index for milling to less than 40-mesh size

| Materials | Mill |
|---|---|
| Acetanilide | Ball, roller, hammer, fluid-energy |
| Alum | Ball, roller, hammer, fluid-energy |
| Antibiotics | Ball, hammer, colloid, fluid-energy |
| Ascorbic acid | Ball, roller, hammer, fluid-energy |

| | |
|---|---|
| Barium sulfate | Hammer, colloid, fluid-energy |
| Benzoic acid | Hammer, fluid-energy |
| Boric acid | Ball, roller, hammer, fluid-energy |
| Caffeine | Roller, hammer |
| Calcium stearate | Hammer, colloid, fluid-energy |
| Carboxymethylcellulose | Ball, hammer, colloid, fluid-energy |
| Citric acid | Hammer, fluid-energy |
| Color, dry | Ball, hammer, fluid-energy |
| Color, wet | Hammer, colloid, fluid-energy |
| Filter-cake fluid | Ball, roller, hammer, colloid energy |
| Gelatin | Hammer |
| Iodine | Hammer, fluid-energy |
| Methylcellulose | Ball, hammer |
| Sodium acid phosphate | Hammer, fluid-energy |
| Sodium benzoate | Hammer, fluid-energy |
| Sodium metaphosphate | Hammer, fluid-energy |
| Sodium salicylate | Hammer, fluid-energy |
| Stearates | Hammer, fluid-energy |
| Sugar | Hammer, fluid-energy |
| Urea | Ball, hammer, fluid-energy |
| Vitamins | Ball, hammer, fluid-energy |
| Wax | Hammer, colloid, fluid-energy |

After consideration of these factors (listed in Table 2.4) for a specific milling problem, it is suggested that the equipment manufacturer be consulted and its pilot laboratory be utilized, as there exists a wide variety of mills differing in details of design and modifications. The industrial pharmacist should evaluate the pilot study personally to observe the temperature of the inlet and outlet air, the temperature of the milled material, and the size reduction performance at different mill speeds. A size-frequency analysis should be made on samples from each condition of operation. The pilot

evaluation is important because laboratory procedures of size reduction do not duplicate milling conditions in production mills.

| **Table 2.4:** Factors in selection of a mill |
|---|
| **Material** |
| Physical properties: hard, soft, fibrous, elastic, hygroscopic, solvated |
| Size |
| Moisture content |
| Melting point |
| Flammability |
| Thermolability |
| Subsequent processing |
| **Operation** |
| Size specification of milled material |
| Ease of sanitization |
| Ease of sterilization |
| Ease of adjustments during operation |
| Contamination of milled material |
| Versatility |
| Capacity |
| Batch or continuous |
| Wet or dry |
| Rate of introduction of material |
| Space occupied Labor cost |
| **Auxiliary equipment** |
| Dust collector |
| Mechanical introduction of material |
| Temperature control: jacket, refrigerated air, liquid nitrogen, dry ice |
| Inert atmosphere: carbon dioxide, nitrogen air sweep |
| **Safety** |
| Explosivity |
| Irritativity |

Toxicity
Safety features incorporated into mill

# HOMOGENIZATION

Homogenization is a technique used to mill the particles by processing a suspension of solid particles. Traditionally, homogenizers have been used in the pharmaceutical industry for emulsification. However, they are finding increasing applications in the manufacture of liposomes, nano-suspensions, solid-lipid nanoparticles, tablet coating dispersions, microencapsulation and cell disruption. Homogenization is also used for particle size reduction in pharmaceutical suspensions. Important factors controlling the formation of pharmaceutical emulsions and dispersions are mechanical and/or formulation related. Mechanical forces during homogenization cause droplet or particle size reduction by shear, turbulence, impact, and cavitation. The homogenizers available operate using a combination of these forces.

## Equipments

### *Simple Homogenizer*

In a simple homogenizer, the dispersion of two liquids is achieved by forcing their mixture through a small inlet orifice at high pressures. A homogenizer generally consists of a pump that raises the pressure of the dispersion to a range of 500 to 5,000 psi and an orifice through which this fluid impinges upon the homogenizing valve held in place on the valve seat by a strong spring (Fig. 2.16). As the pressure builds up, the spring is compressed, and some of the dispersion escapes between the valve and the valve seat. At this point, the energy that has been stored in the liquid as pressure is released instantaneously and subjects the product to intense turbulence and hydraulic shear.

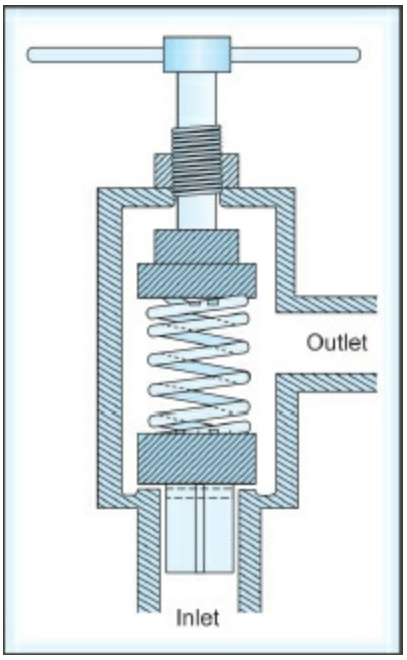



**Fig. 2.16:** Diagrammatic representation of a simple homogenizer

## *Silverson Homogenizer*

In silverson homogenizer the droplets and particles are subjected to a wide variety of high shear rates. Silverson homogenizers are designed with fine tolerance rotor/stator gaps that promote the high shear rates and high amounts of shear per pass through. The maximum shear rates occur in the gap between the high-speed rotating impeller and the stationary housing. Many ingenious designs are available that use various configurations with the purpose of increasing the probability that the solid particles or liquid droplets travel through the rotor/stator zones where maximum shearing occurs. Almost all designs have some form of teeth or blades, which are meshed into an accompanying stationary housing as illustrated in Fig. 2.17.

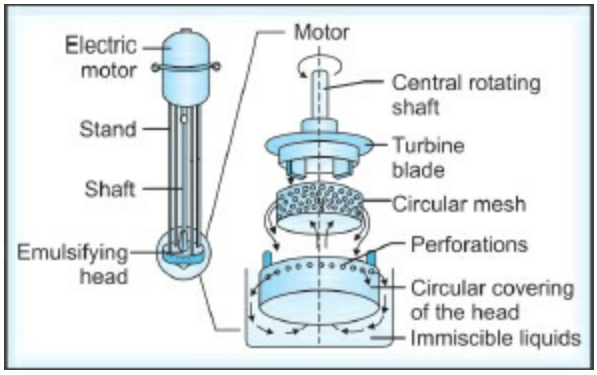


**Fig. 2.17:** Diagrammatic representation of a silverson homogenizer

Of particular importance is the fact that, when a mixture containing solid particles agglomerations or emulsion droplets is pumped through a fixed-gap rotor/stator mixer, not all of the droplets or particles pass through the highest shearing zones. Some particles passing through the machine may escape the rotor/stator gaps and are exposed only to lower shear zones. The more open the design of the rotor/stator, the greater the probability that the particles will miss the high shear zones on a single pass through the device. For this reason the discharge of many rotor/stator devices is restricted with a series of holes or bars, or with a grid, which performs two separate tasks. First, a discharge grid absolutely limits the maximum particle size that can exit the mixer. If the holes on the discharge grid are 2 mm, it is virtually impossible for particles larger than that diameter to pass through the machine. Silverson homogenizers are made with restricting grids as small as 0.5 mm, and a wide range of larger sizes up to 50 mm or even unrestricted outlets. But even the smallest usable discharge grid does not necessarily provide absolute control over the required size distributions in the range of fine dispersion.

## *Ultrasonic/Rapisonic Homogenizer*

The use of ultrasonic energy to produce pharmaceutical emulsions has been demonstrated, and many laboratory-size models are available (Fig. 2.18). The ultrasonic homo-genizer uses a positive displacement pump to force the premixed liquid through an elliptical opening at a speed of 100 m/s or more. This high-speed flow impinges on to the edge of a blade-shaped obstacle, called a vibrating knife. Based on the principle of the Pohlman liquid whistle, in some designs the blade is caused to vibrate at the ultrasonic frequency by the action of the fluid, while in others this vibration is caused by an electrically powered piezoelectric crystal. When a liquid is subjected to ultrasonic vibrations, alternate regions of compression and rarefaction are produced in the liquid. Cavities are formed in the region of rarefaction, which subsequently collapse in the region of compression resulting in high shear. The vibrating element is an extra maintenance item, especially in heavy or abrasive service. These units have seen application in both a two-step premix-homogenization procedure and true multi-component continuous processing. Capacities and pressures of systems ranging from laboratory units to full-scale production units are available.

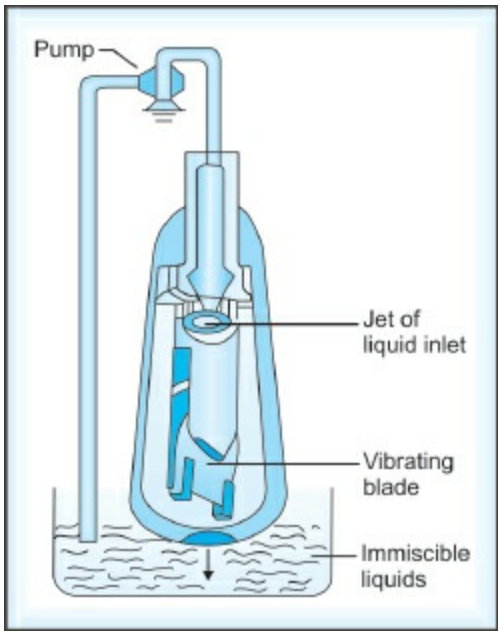



**Fig. 2.18:** Diagrammatic representation of rapisonic homogenizer

## *High-pressure Homogenizer*

The high-pressure homogenizer consists of a positive displacement pump that forces liquid into the valve area at a high pressure (Fig. 2.19). As the product is forced through the adjustable gap, its velocity increases tremendously with a corresponding decrease in pressure. The emerging product then impinges on the impact ring. This sudden change in energy causes increased shear, turbulence, and/or cavitation, resulting in droplet size reduction and uniform dispersion of particles. High pressure homogenizers are used in emulsification, preparation of microparticles and nanodispersions, liposomes and in cell disruption. Another commonly used approach is the multiple-pass homogenizer, if a very narrow particle size distribution is needed. This can be achieved by using a series of homogenizers or processing several discrete passes through the same machine.

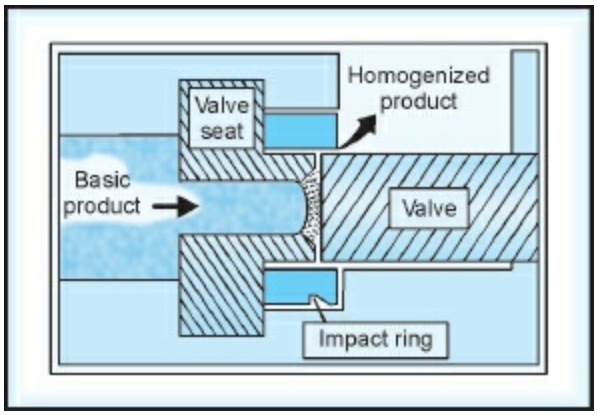



**Fig. 2.19:** Valve assembly of a high-pressure homogenizer

## *Microfluidizer*

The microfluidizer is a high-pressure homo-genizer that works on a different principle. The pre-homogenized liquid is forced through an interaction chamber using a high-pressure pump. The interaction chamber consists of ceramic micro-channels, which cause the liquid feed to split into two streams. These streams are then recombined at very high velocities producing forces of shear, impact, and cavitation (the formation, growth, and implosive collapse of vapor bubbles in liquid), which cause droplet or particle-size reduction in emulsions and suspensions. Reduction of particles to a size less than 1 μm, i.e. nanosize has been reported for cavitation milling. Because of their efficient droplet size reduction and ease of scale-up, microfluidizers are frequently used to prepare parenteral emulsions. A schematic diagram of a microfluidizer process is shown in Fig. 2.20.

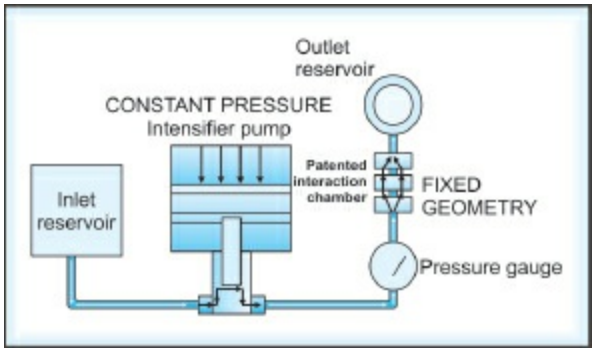



**Fig. 2.20:** Diagrammatic representation of flow path through the microfluidizer

## *Low-pressure Cyclone Emulsifiers*

Another unique apparatus used for the formation of emulsions and suspensions is the cyclone emulsifier. This device uses a positive-displacement pump to feed a special chamber with a tangential entry port. The product is forced to circulate in concentric layers towards the center and the ends of the chamber, where it is expelled. The shear arises from the difference in the velocity of the fluid as the fluid travels in a spiral towards the center as illustrated in Fig. 2.21. These devices require substantially less pressure than ultrasonic homogenizers and high-pressure homo-genizers. They operate in the 200 psig range and have operating capacities of 7.5–225 L/min. They are capable of producing emulsions in the 2–10 μm range, which is very similar to that obtained from fixed-gap rotor/stator machines and some settings of colloid mills. There are no moving or requiring adjustment parts in the cyclone-type emulsifier. The recommended viscosity limit is 1–2,000 centipoise, but some acceptable results have been obtained at higher viscosities.

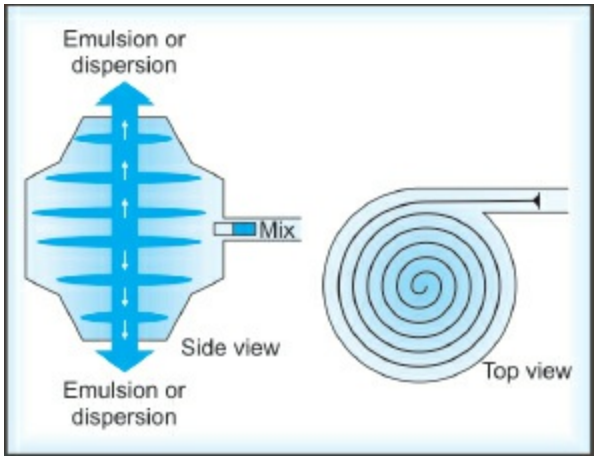



**Fig. 2.21:** Mixing chamber of cyclone homogenizer

## PARTICLE SIZE AND SIZE DISTRIBUTION

In naturally occurring particulate solids and milled solids, the shape of particles is irregular, and the size of the particles varies within the range of the largest and smallest particle. There is no known method of defining an irregular particle in geometric terms, however, statistical methods have been developed to express the size of an irregular particle in terms of a single dimension referred to as its diameter. If this diameter is measured by a standardized procedure for a large number of particles, the values may be expressed by several diameters. It is only required that the surface area is proportional to the square of the diameter and the volume is proportional to the cube of the diameter.

For an irregular particle, an equivalent particle with the same surface or volume may be substituted. For convenience of mathematical treatment, an irregular particle is considered in terms of an equivalent sphere. The size of the particle can then be expressed by a single parameter, $d$ (the diameter). The volume of a particle may be determined by displacement in a liquid and equated to the volume of a hypothetical sphere possessing an equivalent diameter. As the volume of a sphere is $\pi d^3/6$, the equivalent diameter of an irregular particle with a volume $V$ is given by:

$$d = \sqrt[3]{\frac{6V}{\pi}} \qquad \text{... (23)}$$

The effective diameter of particles based on their rate of sedimentation is commonly used in pharmacy. The time required for the particle to settle between two fixed points in a suitable liquid is experimentally determined and allows the evaluation of the rate of sedimentation. By use of Stokes' equation (*see* under "Sedimentation" in this chapter), the effective diameter is calculated. This effective, or Stokes' diameter is the diameter of a sphere that requires the same time to settle between two fixed points in the liquid as does the irregular particle.

In addition to the two effective diameters described, several other diameters are defined and their values are calculated in Table 2.5 for 261 particles measured by microscopy. The arithmetic average diameter is the sum of the diameters of the separate particles divided by the number of particles. If $n_1$, $n_2$, and $n_n$ are the number of particles having diameters $d_1$, $d_2$,

and $d_n$, respectively, the average diameter is:

$$d_{avg} = \frac{n_1 d_1 + n_2 d_2 + \ldots n_n d_n}{n_1 + n_2 + \ldots n_n} = \frac{\sum(nd)}{\sum n} \quad \ldots (24)$$

The average diameter of a group of 261 particles can be calculated from the data in Table 2.5. The average diameter is:

$$d_{avg} = \frac{5366}{261} = 20.6\ \mu m$$

The geometric mean diameter is the *nth* root of the product of the *n* particles measured:

$$d_{geo} = \sqrt[n]{d_1 d_2 \ldots d_n} \quad \ldots (25)$$

Using the logarithmic form of this equation, the geometric mean diameter of the 261 particles is calculated by use of the following:

$$\log d_{geo} = \frac{\sum(n \log d)}{\sum n}$$

$$= \frac{336.0874}{261} = 1.2876 \quad \ldots (26)$$

and $d_{geo}$ = antilog 1.2876 = 19.4 μm

The median diameter is the diameter for which 50% of the particles measured are less than the stated size. An inspection of Table 2.6 shows that 134 particles of the 261 are less than 18 μm. Therefore, the median diameter is approximately 18 μm. Cumulative plots are those in which the percentage of particles less than (or greater than) a given particle size are plotted against size. As shown in Fig. 2.22 for the data in Table 2.6, the cumulative percentage less than the stated size is plotted against size, and the median diameter is read from the 50% value of the curve.

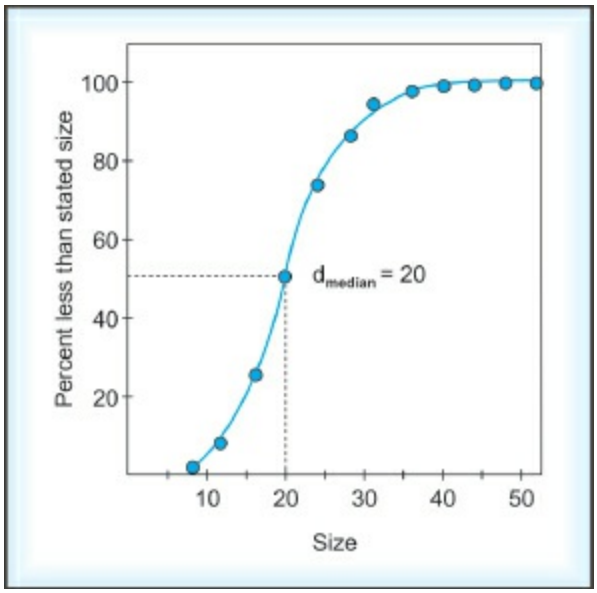



**Fig. 2.22:** Cumulative distribution plot used to determine the median size

The arithmetic or geometric means and the median have no physical significance. The meaningful choice of diameter depends on its relevance to some significant physical property. The packing and flow of a powder or granulation depends on its volume. Thus, if packing is a prime consideration, the size should be expressed as a mean volume diameter. Dissolution and adsorption processes are a function of the surface area of the particles, and with these processes, the particle size should be expressed as a surface mean diameter. As sedimentation is an important property of suspensions, the size of the suspended solids should be expressed as Stokes' diameter.

## Representation of Data

When a material is milled, the particles have a variety of sizes as determined by flaws in the structure. The purpose of particle size measurement is to determine the percentage frequency of distribution of particle sizes. The most precise method of data presentation is tabular form, as in Table 2.5. The data may be presented as a bar graph or histogram of the frequency as a function of particle size. Size-distribution data are commonly presented graphically because a graph is more concise and permits easy visualization of the mean and skewness of distribution. A size-frequency curve is a plot of the percentage frequencies of various particles against the mean of size-groups. The size-frequency curve in Fig. 2.23 is drawn from the data in Table 2.5. The arithmetic and geometric mean diameters are indicated. The mode is the maxima in the size-frequency curve.

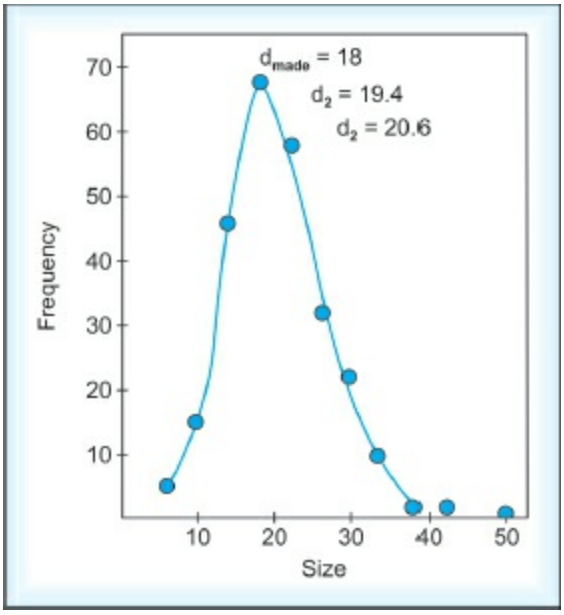



**Fig. 2.23:** Size-frequency distribution of 261 particles measured by microscopy

**Table 2.5:** Definitions of various diameters and their values for 261 particles measured by means of an optical micrometer

| Size | Mean Size $d$ | Frequency $n$ | $nd$ | log $d$ | $n$ log $d$ | $nd^2$ | $nd^3$ | $nd^4$ |
|---|---|---|---|---|---|---|---|---|
| 4 to 7.9 µm | 6 | 5 | 30 | 0.7782 | 3.9910 | 180 | 1080 | 6480 |
| 8 to 11.9 | 10 | 15 | 150 | 1.0000 | 15.0000 | 1500 | 15,000 | 150,000 |
| 12 to 15.9 | 14 | 46 | 644 | 1.1461 | 52.7206 | 9016 | 126,224 | 1,767,136 |
| 16 to 19.9 | 18 | 68 | 1224 | 1.2553 | 85.3604 | 22,032 | 396,476 | 7,138,368 |
| 20 to 23.9 | 22 | 58 | 1276 | 1.3424 | 77.8592 | 28,072 | 617,584 | 13,586,848 |
| 24 to 27.9 | 26 | 32 | 832 | 1.4150 | 45.2800 | 21,632 | 562,432 | 14,623,232 |
| 28 to 31.9 | 30 | 22 | 660 | 1.4771 | 32.4962 | 19,800 | 594,000 | 17,820,000 |
| 32 to 35.9 | 34 | 10 | 340 | 1.5315 | 15.3150 | 11,560 | 393,040 | 1,336,336 |
| 36 to 39.9 | 38 | 2 | 76 | 1.5798 | 3.1596 | 2888 | 109,744 | 2,085,136 |
| 40 to 43.9 | 42 | 2 | 84 | 1.6232 | 3.2464 | 3528 | 148,176 | 6,222,392 |
| 44 to 47.9 | 46 | 0 | 0 | 1.6628 | 0 | 0 | 0 | 0 |
| 48 to 51.9 | 50 | 1 | 50 | 1.6990 | 1.6990 | 2500 | 125,000 | 6,250,000 |
| | | 261 | 5366 | | 336.0874 | 122,708 | 3,088,756 | 70,985,928 |

| Diameter | Definition | Diameter for 261 particles |
|---|---|---|
| Mean-surface | $d_s = \sqrt{\frac{\sum nd^2}{\sum n}}$ | $d_s = \sqrt{\frac{122,708}{261}} = 21.7\ \mu m$ |
| Volume mean | $d_v = \sqrt[3]{\frac{\sum nd^3}{\sum n}}$ | $d_v = \sqrt[3]{\frac{3,088,756}{261}} = 22.8\ \mu m$ |
| Volume-surface mean | $d_{vs} = \frac{\sum nd^3}{\sum nd^2}$ | $d_{vs} = \frac{3,088,756}{122,708} = 25.2\ \mu m$ |
| Weight mean | $d_w = \frac{\sum nd^4}{\sum nd^3}$ | $d_w = \frac{70,985,928}{3,088,756} = 22.9\ \mu m$ |

An infinite number of particle size distributions may have the same average diameter or median. For this reason, parameters other than a median or average diameter are required to define the size of a powder. A powder should be characterized with a size-frequency curve.

Size distributions that follow the probability law are referred to as normal or Gaussian distributions, as shown in Fig. 2.23. This normal-probability distribution is symmetric about a vertical axis. The size-frequency distribution of ground material is usually skewed with the number of particles increasing with decreasing size. It is believed that the size distributions of

milled material follow an exponential law. If the distribution is asymmetric or skewed, it can be frequently made symmetric and will follow the normal-probability law if the sizes are replaced by the logarithms of the sizes.

Size-frequency data are conveniently plotted on an arithmetic-probability or logarithm-probability grid. For a normal distribution, a plot of the cumulative percentage less (or greater) than the stated size against size produces a straight line. In Fig. 2.24, using the data from Table 2.6 (the size is plotted against the cumulative percentage less than the stated size using an arithmetic-probability grid. For a skewed distribution, a plot of the cumulative percentage less (or greater) than the stated size against the logarithm of size generally produces a straight line, as shown in Fig. 2.25.

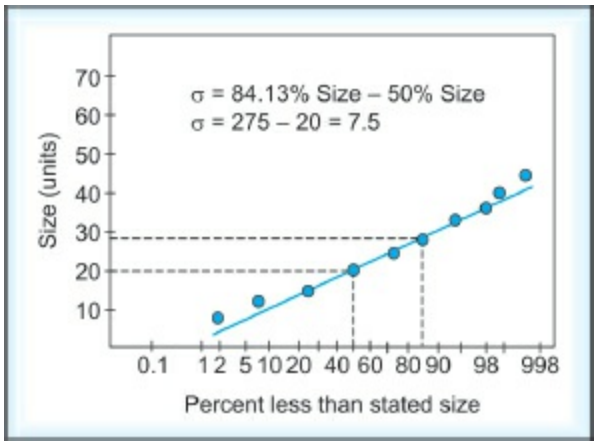



**Fig. 2.24:** Arithmetic-probability plot of data in Table 2.6

**Table 2.6:** Summation for the determination of the median diameter of 261 particles measured by an optical micrometer

| Size-group | Number in each size-group, *n* | Number less than maximum of size-group | Percentage of particles in each size-group | Percentage of particles-less than maximum size of group |
|---|---|---|---|---|
| 4 to 7.9 μm | 5 | 5 | 1.9 | 1.9 |
| 8 to 11.9 | 15 | 20 | 5.8 | 7.7 |
| 12 to 15.9 | 46 | 66 | 17.7 | 25.4 |
| 16 to 19.9 | 68 | 134 | 26.0 | 51.4 |
| 20 to 23.9 | 58 | 192 | 22.2 | 73.6 |
| 24 to 27.9 | 32 | 224 | 12.4 | 85.8 |
| 28 to 31.9 | 22 | 246 | 8.4 | 94.2 |
| 32 to 35.9 | 10 | 256 | 3.8 | 98.0 |
| 36 to 39.9 | 2 | 258 | 0.8 | 98.8 |
| 40 to 43.9 | 2 | 260 | 0.8 | 99.6 |
| 44 to 47.9 | 0 | 260 | 0 | 99.6 |
| 48 to 51.9 | 1 | 261 | 0.4 | 100 |

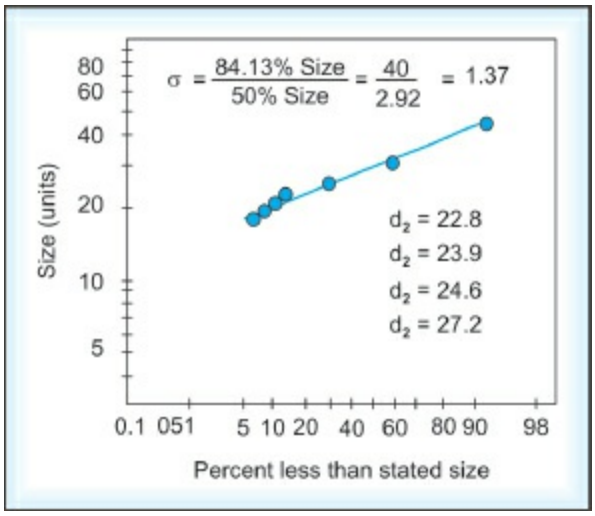



**Fig. 2.25:** Logarithm-probability plot of data in Table 2.6

When the plots are made on either probability grid, the distributions must be asymptotic on both extremes. In practice, there may be a largest and smallest particle in the material measured, therefore, the distribution is not asymptotic, and the plots on the probability grids often depart from linearity at the extremes. This does not detract from the usefulness of such plots, as the areas extending from the extremes to infinity are negligible as compared to the areas contained under the distribution curve between the largest and smallest particles measured.

The calculations involved in computing the mean diameter and the

standard deviation are reduced by the use of probability grids. The median diameter for both grids is obtained by reading from the curve the size corresponding to 50% value on the probability scale. In the arithmetic-probability plot, the mean is the arithmetic average and in the logarithm-probability plot, the 50% size is the geometric mean. The standard deviations can be obtained from the arithmetic-probability plot from the relation:

$$\sigma = 84.13\%\ \text{size} - 50\%\ \text{size} \quad \ldots (27)$$

$$\sigma = 50\%\ \text{size} - 15.87\%\ \text{size} \quad \ldots (28)$$

and from the logarithm-probability plot from the relations:

$$\sigma_{geo} = \frac{84.13\%\ \text{size}}{50\%\ \text{size}} \quad \ldots (29)$$

$$\sigma_{geo} = \frac{50\%\ \text{size}}{15.87\%\ \text{size}} \quad \ldots (30)$$

Using these probability functions, Hatch derived equations relating various types of diameters by use of the standard deviation and mean. These statistical parameters are a function of the size and numeric frequency of the particles for a given size. To calculate their values, the size-distribution data must be expressed in terms of a number frequency. In microscopy, this requirement is met directly. However, in sieving and sedimentation methods, the data obtained provide a weight distribution. Fortunately, as shown in Table 2.7, equations have been derived relating the weight distribution data to statistical diameters. The prime on the $d'_{geo}$ and $\sigma'_{geo}$ signify a weight distribution rather than a number distribution. The geometric standard deviations for a weight and number distribution are practically identical.

**Table 2.7:** Definitions of diameters of non-uniform particulate systems in terms of the parameter of size distribution curves by number and by weight

| Diameter | Numbers distribution | Weight distribution |
|---|---|---|
| Geometric mean, $d_{geo} = \frac{\sum(n \log d)}{\sum n}$ | | $\log d_{geo} = \log d'_{geo} - 6.9078 \log^2 \sigma'_{geo}$ |
| Arithmetic mean, $d_{ave} = \frac{\sum nd}{\sum n}$ | $\log d_{ave} = \log d_{geo} + 1.151 \log^2 \sigma_{geo}$ | $\log d_{ave} = \log d'_{geo} - 5.756 \log^2 \sigma'_{geo}$ |
| Mean surface, $d_s = \sqrt{\frac{\sum nd^2}{\sum n}}$ | $\log d_s = \log d_{geo} + 2.3026 \log^2 \sigma_{geo}$ | $\log d_s = \log d'_{geo} - 4.6052 \log^2 \sigma'_{geo}$ |
| Mean volume, $d_v = \sqrt{\frac{\sum nd^3}{\sum n}}$ | $\log d_v = \log d_{geo} + 3.4539 \log^2 \sigma_{geo}$ | $\log d_v = \log d'_{geo} - 3.4539 \log^2 \sigma'_{geo}$ |
| Mean volume-surface $d_{vs} = \frac{\sum nd^3}{\sum nd^2}$ | $\log d_{vs} = \log d_{geo} + 5.7565 \log^2 \sigma_{geo}$ | $\log d_{vs} = \log d'_{geo} - 1.1513 \log^2 \sigma'_{geo}$ |

To illustrate the use of these equations, the data from a sample of magnesium hydroxide given in Table 2.6 are plotted in Fig. 2.25 with the cumulative percentage less than stated size on the probability grid and the size on the logarithmic grid. The geometric mean diameter corresponding to the 50% value on the cumulative percentage axis is 29.2 μm. The geometric standard deviation is:

$$\sigma'_{geo} = \frac{84.13\% \text{ size}}{50\% \text{ size}} = \frac{40}{29.2} = 1.37 \text{ μm}$$

Knowing the value of the geometric mean diameter and the geometric standard deviation, the Hatch-Choate equations may be used to calculate statistical diameters given in Fig. 2.4. For example, the mean surface diameter of the sample of magnesium hydroxide is:

$$\log d_s = \log d_{geo} - 4.606 \log^2 \sigma'_{geo}$$
$$\log d_s = \log 29.2 - (4.6060 \times 0.0187) = 1.3793$$
$$d_s = 23.9 \text{ μm}$$

## Particle Size Determination

The particle size distribution in a powder may be quantified by:

1. Determining the number of particles-*microscopy*
2. Determining the weight of particles-*sieving*, *sedimentation*, *centrifugation*, *elutriation*
3. Determining the volume of particles-*coulter counter*
4. Determining light scattering by particles-*dynamic and laser light scattering*.

### *Microscopy*

Microscopy is the most direct method for size distribution measurement. Its lower limit of application is determined by the resolving power of a lens. A particle cannot be resolved if its size is close to the wavelength of the light source. For white light, an ordinary microscope is used to measure particles from 0.4 to 150 μm.

With special lenses and ultraviolet light, the lower limit may be extended to 0.1 micron. In the ultramicroscope, the resolution is improved by the use of a dark field illumination. The size range of the ultramicroscope is from 0.01 to 0.2 μm.

The diameters of the particles on the slide are measured by means of a calibrated filar micrometer eyepiece. The hairline of the eyepiece is moved by the micrometer to one edge of a particle, and the reading on the micrometer is recorded. The hairline is then moved to the opposite edge of the particle being measured, and the micrometer is read. The difference between the two readings is the diameter of the particle. All of the particles are measured along an arbitrary fixed line.

Graticules or eyepieces with grids of circles and squares are used to compare the crosssectional area of each particle in the microscopic field with one of the numbered patterns. The number of particles that best fits one of the numbered circles is recorded. The field is changed, and the procedure is repeated with another numbered circle. This procedure is repeated until the entire size range is covered.

In both techniques, the magnification is determined by the use of a

calibrated stage micrometer, as the magnification is not equal to the product of the nominal magnification of the objective and the eyepiece. The particulate field to be counted should be random. The total number of fields to be counted depends on the number of particles per field. In principle, the number of particles measured should be enough so that the results do not change on measuring a larger number of particles. The British Standard on microscopic counting recommends at least 625 particles. If the particle size distribution is wide, it may be necessary to count more particles. If the particle size distribution is narrow, as few as 200 particles may be sufficient.

There is a considerable variation among operators using the microscopic technique. Photomicrographs, projections, and automatic scanners have been used to lessen the operator fatigue. The diameters measured by a microscope and denned in Table 2.5 are number parameters, since microscopy involves a counting procedure.

## *Sieving*

Sieving is the most widely used method for measuring particle size distribution because it is inexpensive, simple, and rapid with little variation between operators. Although the lower limit of application is generally considered to be 50 μm, micromesh sieves are available for extending the lower limit to 10 μm. A sieve consists of a pan with the bottom of wire cloth with square openings. In the United States, two standards of sieves are used. The Tyler Standard Scale is based on the size of opening (0.0029”) in a wire cloth having 200 openings per linear inch, i.e. 200-mesh. In the Tyler Standard Scale, the ratio of the width of openings in successive sieves is $\sqrt{2}$. The United States Standard Scale proposed by the National Bureau of Standards in general uses the ratio $\sqrt{2}$, but it is based on an opening of 1 mm (18-mesh). The two standard sieves are compared in Table 2.8.

**Table 2.8:** Designations and dimensions of US Standard and Tyler Standard sieves

| US Standard | | Tyler Standard | |
|---|---|---|---|
| Micrometer | Mesh | Micrometer | Mesh |
| 5660 | 3½ | 5613 | 3½ |
| 4760 | 4 | 4699 | 4 |
| 4000 | 5 | 3965 | 5 |
| 3360 | 6 | s3327 | 6 |
| 2830 | 7 | 2794 | 7 |
| 2380 | 8 | 2362 | 8 |
| 2000 | 10 | 1651 | 10 |
| 1680 | 12 | 1397 | 12 |
| 1410 | 14 | 1168 | 14 |
| 1190 | 16 | 991 | 16 |
| 1000 | 18 | 883 | 20 |
| 840 | 20 | 701 | 24 |
| 710 | 25 | 589 | 28 |
| 590 | 30 | 495 | 32 |
| 500 | 35 | 417 | 35 |
| 420 | 40 | 351 | 42 |
| 350 | 45 | 295 | 48 |
| 297 | 50 | 246 | 60 |
| 250 | 60 | 208 | 65 |
| 210 | 70 | 175 | 80 |
| 177 | 80 | 147 | 100 |
| 149 | 100 | 124 | 115 |
| 125 | 120 | 104 | 150 |
| 105 | 140 | 88 | 170 |
| 88 | 170 | 74 | 200 |
| 74 | 200 | | |
| 62 | 230 | | |
| 53 | 270 | | |
| 44 | 325 | | |
| 37 | 400 | | |

The procedure involves the mechanical shaking of a sample through a series of successively smaller sieves (Fig. 2.26), and the weighing of the portion of the sample retained on each sieve. The type of motion influences sieving: vibratory motion is most efficient, followed successively by side-tap motion, bottom-tap motion, rotary motion with tap, and rotary motion. Time is an important factor in sieving. The load or thickness of powder per unit area of sieve influences the time of sieving. For a given set of sieves, the time required to sieve a given material is roughly proportional to the load placed on the sieve. Therefore, in size analysis by means of sieves, the type of motion, time of sieving, and load should be standardized.

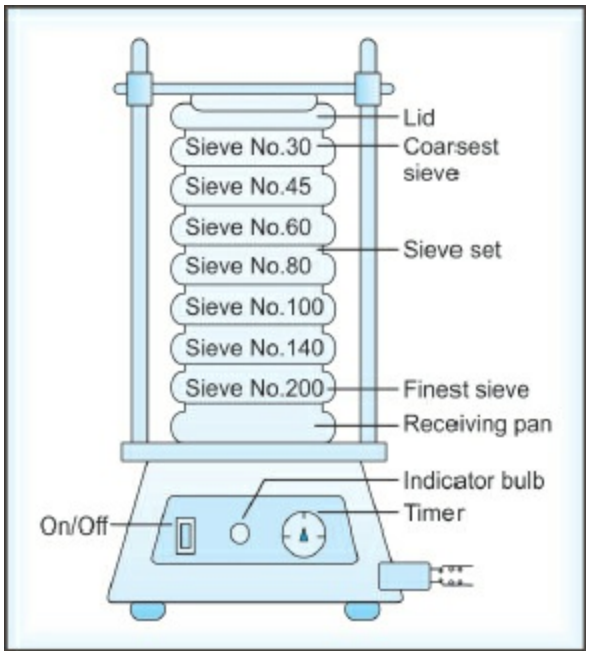



**Fig. 2.26:** Arrangements of sieves in sifter

A typical size-weight distribution obtained by sieving is shown in Table 2.9. The size assigned to the sample retained is arbitrary, but by convention, the size of the particles retained is taken as the arithmetic or geometric mean of the two sieves (a powder passing a 30-mesh and retained on a 45-mesh sieve is assigned an arithmetic mean diameter of (590 + 350)/2 or 470 μm).

If the weight distribution obtained by sieving follows a logarithm-probability distribution, the Hatch-Choate equations given in Table 2.7 permit conversion of the weight distribution to a number distribution.

**Table 2.9:** Weight-size distribution of granular sodium bromide as measured by US Standard sieves

| Sieve number (passed/retained) (1) | Arithmetic mean size of openings (2) | Weight retained on smaller sieve (3) | % Retained on smaller sieve (4) | Weight size (2) × (4) |
|---|---|---|---|---|
| 30/45 | 470 μm | 57.3 g | 13.0 | 6100 μm |
| 45/60 | 300 | 181.0 | 41.2 | 12,380 |
| 60/80 | 213 | 110.0 | 25.0 | 5320 |
| 80/100 | 163 | 49.7 | 11.3 | 1840 |
| 100/140 | 127 | 20.0 | 4.5 | 572 |
| 140/200 | 90 | 22.0 | 5.0 | 450 |
| | | 400.0 | 100.0 | 26,662 |

## *Sedimentation*

Sedimentation method may be used over a size range of 1 to 200 μm to obtain a size-weight distribution curve and to permit the calculation of particle size. The sedimentation method is based on the dependence of the rate of sedimentation of the particles on their size as expressed by Stokes' equation:

$$d_{stokes} = \sqrt{\frac{18\eta}{(\rho - \rho_0)} \frac{x}{t}} \qquad \ldots (31)$$

where, $d_{stokes}$ is the effective or Stokes' diameter, pis the viscosity of the dispersion fluid, $x/t$ is the rate of sedimentation or distance of fall $x$ in time $t$, $g$ is the gravitational constant, and $\rho$ and $\rho_0$ are the densities of the particle and the medium, respectively. Stokes' equation is applicable to free spheres that are falling at a constant rate. If the concentration of the suspension does not exceed 2%, there is no significant interaction between the particles, and they settle independent of one another.

The ***Andreasen pipet method*** is the simplest means of incremental particle size analysis. A 1% suspension of the powder in a suitable liquid medium is placed in the pipet (Fig. 2.27). At given intervals of time, samples are withdrawn from a specified depth without disturbing the suspension and are dried so that the residue may be weighed. By means of Stokes' equation, the particle diameter corresponding to each interval of time is calculated, with $x$ being the height of the liquid above the lower end of the pipet at time $t$ when each sample is withdrawn. As the sizes of the particles are not uniform, the particles settle at different rates. The size-distribution and concentration of the particles vary along the length of the suspension as sedimentation occurs. The larger particles settle at a faster rate and fall below the pipet tip sooner than the smaller particles, thus, each sample drawn has a lower concentration and contains particles of smaller diameter than the previous sample.

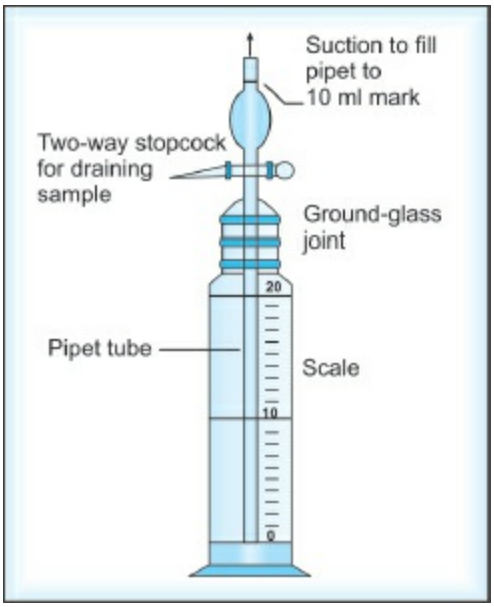



**Fig. 2.27:** At measured time intervals, a 10 ml sample is withdrawn by aspiration at a depth that can be read from the scale etched on the Andreasen pipet

From the weight of the dried sample, the percentage by weight of the initial suspension is calculated for particles having sizes smaller than the size calculated by Stokes' equation for that time. The weight of each sample residue is called the *weight undersize*, and the sum of the successive weights is known as the *cumulative weight undersize*. Typical data obtained by use of the Andreasen pipet are given in Table 2.10.

**Table 2.10:** Weight-size distribution of magnesium hydroxide as determined by Andreasen pipet using water as medium

| Time | Height (cm) | Weight of residue (units) | Percentage of initial suspension | Particle diameter calculated by means of Stokes' equation (μm) |
|---|---|---|---|---|
| 120 | 20.0 | 0.0912 | 92.4 | 44.5 |
| 240 | 19.6 | 0.0591 | 59.9 | 30.5 |
| 360 | 19.2 | 0.0321 | 32.5 | 25.1 |
| 420 | 18.8 | 0.0134 | 13.9 | 22.9 |
| 480 | 18.4 | 0.0107 | 10.9 | 20.7 |
| 600 | 18.0 | 0.0089 | 9.0 | 18.5 |
| 720 | 17.2 | 0.0069 | 7.0 | 16.8 |

In Fig. 2.25, the plot of logarithm of size against the percentage less than the stated size produces a straight line and allows the evaluation of the ******ebook converter DEMO Watermarks*******

geometric mean diameter and standard deviation. The geometric mean diameter corresponding to the graphic 50% size is 29.2 μm. The standard deviation, which is evaluated as the ratio of the 84.13% size to the 50% size (40/29.2), is 1.37. With these two values, the weight distribution obtained by sedimentation may be converted into number distribution by use of the Hatch-Choate equations. For example, if one wishes to calculate the mean surface diameter, the appropriate equation selected from Table 2.7 is:

$$\log d_s = \log d'_{geo} - 4.6052 \log^2 \sigma'_{geo}$$
$$= \log 29.2 - (4.6052 \log^2 1.37)$$
$$d_s = \text{antilog } 1.3793 = 23.9 \text{ μm}$$

## *Centrifugation*

When particles are small, normal sedimentation methods are very slow and factors such as Brownian movement interfere with the results. This can be overcome by applying the same basic principles, but utilizing centrifugal force instead of gravitational force, where settling velocities can be increased greatly. The rate of sedimentation may be observed by methods similar to those used for gravitational sedimentation. Calculation of results is by Stroke's law, with an appropriate factor to indicate the number of times the centrifugal force is greater than the gravitational force. It will be realized that the method is most useful for dealing with very small particles.

## *Elutriation*

Eluriation is a procedure in which the fluid moves in a direction opposite to the sedimentation movement, so that in the gravitational process. For example, the particles will move vertically upwards. Then the velocity of the fluid is less than the settling velocity of the particle, the latter will move downwards against the stream of the fluid. If however, the reverse applies, the settling velocity of the particle will be insufficient to overcome the velocity of the fluid, and the particle will be carried upwards. In case of sedimentation, the separation is dependent upon the time for settling of particles while in elutriation the separation depends on the velocity of the fluid and is independent of time. Separation into several fractions may be effected by using a number of vessels of increasing diameter, with the suspension entering the bottom of the narrowest column, overflowing from the top to the bottom of the next widest column, and so on. Since the mass

flow remains the same, the greater diameter will cause the fluid to flow at a lower velocity. As the tubes become wider, therefore, particles of decreasing size will be separated.

## *Coulter Counter*

Counter Coulter, which is recently developed method for the determination of particle size, is based on conductivity measurement. The particles are suspended in an electrically conductive fluid and the suspension is allowed to flow through a suitable aperture with an immersed electrode on either side and the particle concentration is arranged so that only one particle travels through the aperture. As the particle passes through the aperture, some electrolyte is displaced and changes the resistance between the electrodes which causes a pulse in the voltage. The magnitude of pulse will be proportional to the volume of the particle. This is a unique feature of the coulter counter, since it is the only method that measures a property having a direct relationship with the volume of the particle (Fig. 2.28). The changes in voltage are amplified and impulses above a predetermined threshold value are counted, so that the recorder provides a count of the number of particles over a certain size. The process is repeated with different threshold values, the coulter counter carrying out counts of particles which are oversize to a series of pre-determined values.

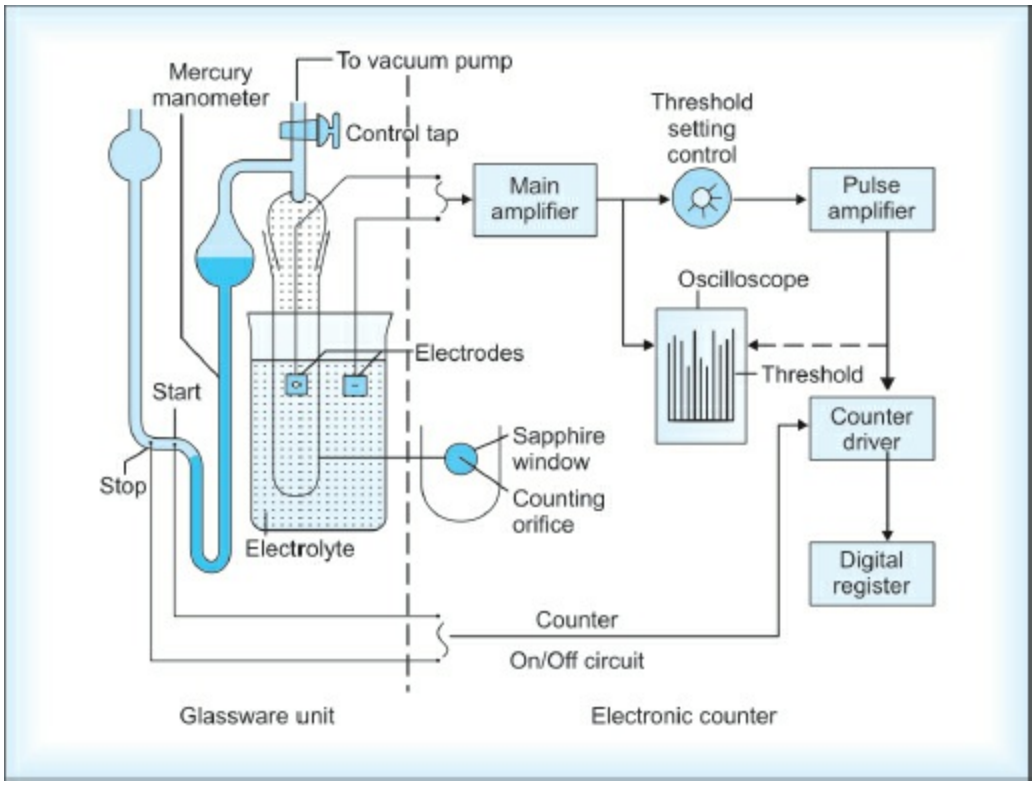



**Fig. 2.28:** Schematic representation of Coulter counter

The various advantages associated with this technique are: (i) results are expressed in terms of particle volume from which it is a simple matter to calculate the diameter of the sphere of equivalent volume, (ii) instrument can operate with particles between 0.5 and 1000 μm, (iii) operation is very rapid, (iv) since a large number of particles are counted, the results are more reliable, and (v) operational simplicity of the method reduces operator variables, enabling reproducible results to be obtained.

## *Dynamic Light Scattering (DLS)*

Light scattering techniques have been the mainstay for the determination of particle size and particle size distributions. When light is directed at a particle, it can either be deflected or absorbed by the particle, which is dependent on the size of the particle relative to the wavelength of the light. Particles that are large compared to the wavelength of the light tend to diffract the light. Smaller particles, close to the size of the wavelength or smaller tend to scatter light. Most measurements are conducted on dilute particle dispersions, rather than single particles. DLS is a relatively simple and inexpensive technique to conduct measurements, and computer-

controlled acquisition and analysis are available. DLS is well suited for measuring nanosize particles because it takes the advantage of Brownian motion, which is unique to colloidal particles. For this reason, DLS is limited to particles that are small and not affected by settling forces. DLS measures the velocity of a particle in motion and correlates this velocity to size. A laser light is utilized as a radiation source and the frequency variation of scattered light owing to particle motion are directly related to particle velocity and used to calculate particle size based on the theory of Brownian motion. The theory assumes that the particles are spherical, and the refractive index of the dispersion medium and particles must be known. This technique can also be used to determine droplet size of emulsions and microemulsions. The equipment based on the principle of DLS is described in Fig. 2.29. The summary of various particle size determination techniques and range of particle size analyzed by them are shown in Table 2.11.

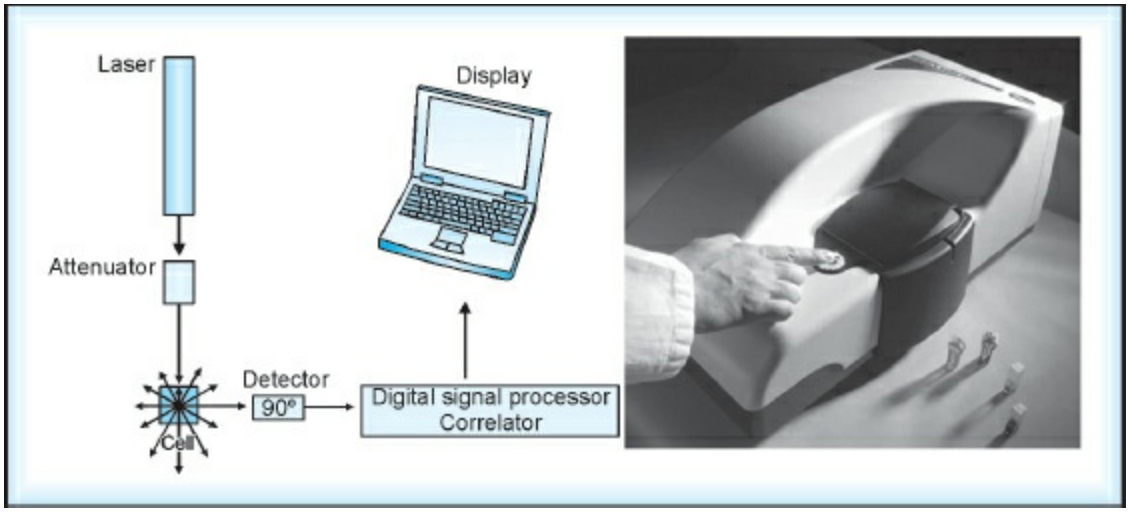



**Fig. 2.29:** Equipment for particle size analysis based on DLS principle

**Table 2.11:** Summary of particle size analysis methods

| Method | Particle size range analyzed |
|---|---|
| Sieving | > 50 µm |
| Andreasen pipet | 1–200 µm |
| Centrifugation | 0.05–5 µm |
| Elutriation | 2–100 µm |
| Optical microscopy | > 0.2 µm |

| | |
|---|---|
| Electron microscopy | > 0.01 μm |
| Coulter counter | > 0.5 μm |
| Light scattering | 0.01–1 μm |

# 3:
# Evaporation and Distillation

Practically, all the operations that are carried out involve either production or absorption of energy in the form of heat. Unfortunately, this process is taken so much for granted that principles behind it are generally overlooked. This chapter introduces the pharmacists to the general principles behind heat transfer and two of its most important applications: evaporation and distillation.

Heat is a form of energy associated with random movement of molecules. For our purposes, it can be well understood with the help of an example in which a metal vessel containing liquid is heated, the molecules of the vessel will start vibrating with a higher energy and with the quest of coming to a lower energy state, will release their energy to the water molecules in contact. In this way, heat energy is transferred from more energetic molecules to those with lesser energy.

The temperature of a material is an indication of its internal energy, the greater the molecular motion, the greater the internal energy, and higher the temperature. Therefore, at a particular temperature all the molecules will have some energy. It is only at absolute zero (0 K), that the substance will have no heat content. Temperature gradient gives rise to spontaneous transfer of heat and the rate of heat transfer indicates how quickly heat is exchanged. It is expressed in units $Js^{-1}$ or watts (W).

**Pharmaceutical applications:** Many pharmaceutical processes involve heat transfer. Some of the direct purposes of heat transfer

include sterilization of the products using autoclaves and hot air ovens, evaporation of liquids for yielding concentrated products, drying of granules for tablet production, separation of mixture components using distillation, melting of substances, creating an elevated temperature during production of creams, suppositories or ointments, heating of the solvents to hasten dissolution processes, etc.

Evaporation, distillation and drying are the three direct applications 

involving heat transfer. The processes of evaporation, distillation and drying can be distinguished from each other based upon the desired product. Evaporation is carried out when a concentrated liquid residue is needed, distillation is done when condensed vapour is required and drying is done when a dried solid residue is required as the product.

The basic knowledge of principles and mechanisms of heat transfer is of prime importance before getting into the details of each of these processes.

## Heat Transfer Mechanisms

Heat can flow by one or more of the three basic mechanisms viz, *conduction, convection and radiation*. In *conduction*, heat transfer occurs by transmission of momentum of individual molecules without the involvement of actual mixing. So, it is limited to solids and 'static' fluids that are bound in some way to prevent their free motion. In solids, heat is mainly transferred by virtue of electron movement. This is the reason why metals are good conductors (as they contain a free electron) and non-metals are not. In fluids, heat transfer by actual molecular collisions is of greater effect. Therefore, in case of gases, conductivity increases remarkably due to the increase in the movement of molecules at higher temperatures, whereas most of the liquids except water show poor conductivity at higher temperatures.

The basic law of heat transfer by conduction can be written as:

$$\text{Rate} = \frac{\text{Driving force}}{\text{Resistance}} \qquad \ldots (1)$$

Here, driving force is the temperature gradient across the solid, heat flows from the region of higher energy to the region of lower energy and resistance can be defined in terms of *Fourier's law*, which states that the rate of heat flow through a uniform material is proportional to the cross-sectional area and the temperature drop, and inversely proportional to the length of the path of the flow.

$$\frac{\partial Q}{\partial t} = \frac{KA\partial t}{\partial L} \qquad \ldots (2)$$

where, ∂Q/∂t is the rate of heat transfer, *A* is the area of cross section of heat flow path, *∂t/∂L* is the temperature gradient per unit path length, and *k* is the coefficient of thermal conductivity of the medium. It is the heat passing in unit time from one face of a cube of unit side to the opposite face, at a unit temperature difference. The numerical value of *k* depends upon the material of which the body is made and upon its temperature. The thermal conductivities vary through a wide range, being highest for metals and lower for non-metals, liquids and gases.

In *convection*, heat flow results from mixing or turbulence within the fluids wherein the molecules are free to move around. Natural convection occurs when there is heat transfer due to difference in density within the fluids. Hotter regions have lower density owing to greater expansion, hence

currents are set up as the warm less dense fluid rises and mixes with colder fluid. If quick heat transfer is desired, then the fluids are forced to move by means of mixer blades or by making the use of baffles. Heat transfer by this means is termed as forced convection.

The third basic method of heat transfer is through space by electromagnetic waves, which is known as *radiation*. To understand the principles of radiation, let us consider a hot body that acts as an emitter. Emission of energy occurs when an electron from the higher energy level moves to a lower energy level. Energy being released has all the properties of electromagnetic waves. These waves have photons which strike the electron or nucleus, whichever is susceptible to the energy level of the photon of the *receiver*. Such a collision results in increase in the energy of the receiver. The waves falling on the receiver are either reflected, transmitted or absorbed. In solids, due to denser surface most of the radiations are absorbed, whereas in liquids, a higher proportion is transmitted and still higher penetration is there in gases. The absorbed radiations are transformed into heat. The extent to which the radiations are absorbed depends upon the properties of the receiver and wavelength of the radiation. A *black body* is the one which converts all the incident radiation into heat and emits all thermal energy as radiation. It can be stated that:

$$\alpha + \tau + r = 1 \quad \ldots (3)$$

where, $\alpha$ is the absorptivity, which can be defined as the fraction of the incident radiation that is absorbed, $r$ is the reflectivity, fraction of incident radiation that is reflected and $\tau$ is the transmissivity, that is the fraction of incident radiation that is transmitted. For most of the solids transmissivity is zero and therefore, the above equation becomes:

$$\alpha + r = 1 \quad \ldots (4)$$

The heat transfer through radiation is based on two important laws viz Kirchhoff's law and Stefan-Boltzmann law.

*Kirchhoff's law* gives the relationship between emissive power of the surface and its absorptivity. The law can be understood with the help of an example of a small body being placed inside a large evacuated enclosure with a wall temperature $T$. Heat exchange occurs until equilibrium is attained. According to the law, if G is the rate at which energy falls from the hot wall of the evacuated enclosure on the body, $\alpha$, is its absorptivity and $E$ is the

emissive power of the body then at equilibrium, the energy balance can be given by:

$$G\alpha = E \ldots (5)$$
$$G = E/\alpha \ldots (6)$$

where, G is the function of the temperature and geometrical arrangement of the surface. If the body is small as compared to the enclosure and its effect upon the irradiation is also negligible then G will remain constant. And hence, the Kirchhoff's law states that the ratio of emissive power to the absoptivity is same for all the bodies in thermal equilibrium and is given by the equation:

$$E_1/\alpha_1 = E_2/\alpha_2 \ldots (7)$$

As black body is perfect radiator, it is used for comparison of emissive powers. The ratio of emissive power of a surface (E) to the emissive power of a perfectly black body ($E_b$) at the same temperature is known as emissivity ($\varepsilon$) of the surface. This can be expressed as:

$$\varepsilon - E/E_b \ldots (8)$$

Therefore, at thermal equilibrium the emissivity of a body is equal to its absorptivity. Although emissive power of a surface varies with the wavelength, for certain material it is constant fraction of the emissive power of a perfectly black body ($E_b$), i.e. $E/E_b$ is constant. Such materials which have constant emissivity are known as *grey bodies*. Thus for grey bodies to apply Kirchhoff's law, it is not necessary that the two bodies should be at the same thermal equilibrium.

*Stefan-Boltzmann law* states that emissive power of a black body is proportional to the fourth power of the absolute temperature, which can be expressed as:

$$E_b = \sigma T^4 \ldots (9)$$

where, $\sigma$ is Stefan-Boltzmann constant, its numerical value being $5.67 \times 10^{-8}$ $W/m^2 \times K^4$.

## EVAPORATION

Evaporation refers to the change in the phase of a component from liquid to gas. In the pharmaceutical industry, evaporation is chiefly associated with removing water and other solvents in batch operations. The suspended components do not appear in the vapour phase, if they do appear then the same operation is referred to as *distillation*. However, the same principles apply to both operations and are derived from studying heat transfer to the boiling liquid, the pertinent physical properties of the liquid, and the thermal constancy of its components.

## Principle

The heat required to boil a liquid in any vessel is usually transferred from a heating fluid, such as steam or hot water, across the wall of a jacket or tube around or inside which the liquid boils. Rate of heat flow is governed by the following equation:

$$Q = IA \Delta T \ldots (10)$$

where, Q is the rate of heat flow, U is the overall heat transfer coefficient, *A* is the area over which heat is transferred, and ΔT is the difference in temperature between the fluids. The overall heat transfer coefficient is derived from a series of individual coefficients that characterize the thermal barriers. It includes the resistance given by metal wall and liquid films (both condensed steam film and liquid side film) present on the either side of the wall. If the solid barrier consists of a thin metal wall, the resistance to heat flow is small, whereas a glass wall may provide the largest thermal resistance of the system. As we know that, for the heating fluid, the film coefficient for a condensing vapour, such as steam is high, therefore, the condensed liquid on the steam side should be removed immediately, as soon as it is formed using a suitable steam trap. With liquid heating media, the velocity of fluid inside should be as high as possible as this will maintain a low boundary film thickness and thus resistance on the liquid side can be minimized. High boiling temperature and rapid circulation of the liquid promote high film coefficients.

Other factors affecting the heat transfer includes area across which the heat transfer has to take place and the difference in temperature of the boiling liquid and the heating surface. Surface area can be kept high by making use of long and coiled tubes. It is important to keep the boiling point of the liquid low, which rises due to material in the solution and hydrostatic head. For dilute solutions, the rise in the boiling point with solute concentration can be calculated from *Raoult's law*. However, this procedure is not applicable to concentrated solutions or to solutions of uncertain composition. For aqueous concentrated solutions, *Duhring's rule* may be used to obtain the boiling point rise of a solution at any pressure. This rule states that the boiling point of a given solution is a linear function of the boiling point of water at the same pressure. Therefore, a plot of the temperature of the constant

concentration solution versus the temperature of a reference substance where the reference substance and solution exert the same pressure, results in a straight line. Reference substance is generally pure water. Duhring line can be drawn by knowing two vapour pressures and temperature points required for the solution to boil. A family of lines is required to cover a range of concentration as shown in Fig. 3.1.

The liquid which is at the bottom is subjected to the pressure of the liquid column above it. Therefore, hydrostatic boiling occurs at the bottom at higher temperature than the surface. The difference between boiling point of the solution under pressure and solution at evaporating surface is boiling point rise due to hydrostatic head. Thus, hydrostatic head should be kept minimum, so as to keep high evaporation rates.

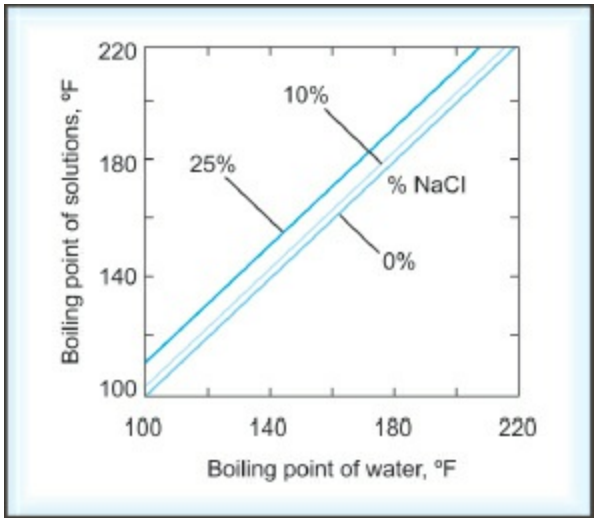



**Fig. 3.1:** Duhrings lines of sodium chloride

## Evaporators

It is convenient to classify evaporators on the basis of heating surface.

### *Evaporating Pan/Steam Jacketed Kettle*

As the name implies, heat is transferred to the aqueous extract by *conduction* and *convection* through steam which is supplied to the outer jacket. Rise in temperature increases the escaping tendency of the solvent molecules into the vapour phase. Stirring further abet the vapourization of solvent molecules. Pans are hemispherical or shallower constructed from a suitable material such as tinned copper and stainless steel. Such a shape gives the best surface/volume ratio for heating and the largest area for rescue of vapours. Recovery of the product from such a shallower form is through the outlet provided at the bottom as shown in the Fig. 3.2. The outer jacket which is generally made up of tinned iron is provided with an inlet for the steam and vent at the top for non-condensed gases. It is constructed in a way to withstand high steam pressure of about 40lb/inch$^2$. The rate of heat transfer in such kettles may vary from 50 to 300 Btu (sq ft) (°F) (h) depending upon the viscosity of the liquid being evaporated and the amount of agitation. Such evaporating pans are simple, easy to operate, clean and maintain. The vexation about it is its lower heat economy and thus higher running costs.

**Fig. 3.2:** Evaporating pan

### *Horizontal Tube Evaporator*

Norbert Rilliieux in Louisiana in 1843 for the first time built evaporators which received recognition. It was designed using horizontal tubes. The same was invigorated in 1879 into *Wellner-Jelinek* type from which evolved the most modern horizontal tube evaporator. The standard type horizontal tube evaporator is as shown in the Fig. 3.3.

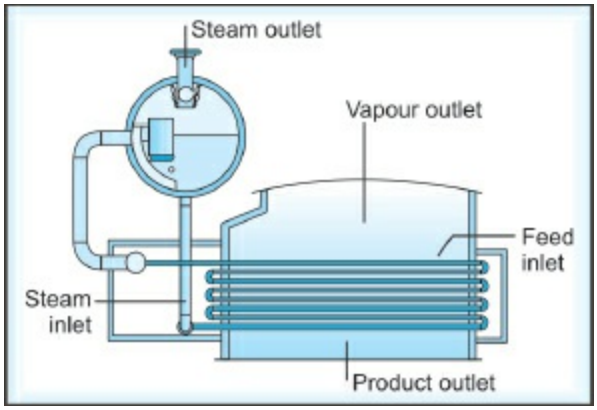



**Fig. 3.3:** Horizontal tube evaporator

Heat is transferred through the horizontal tubes which are ducked in pool of liquid to be evaporated. The liquid just above the tubes gets heated and solvent molecules evaporate leaving behind the concentrated liquid which is collected from the bottom. It consists of a large cylindrical body with conical or dome-shaped top and bottom. It is made up of cast iron or plate steel. The diameter of horizontal tube evaporator may vary widely, ranging from 36 inch to 11 ft, and the tubes are usually from 7/8 to 1M inches in outside diameter. There is inlet for steam, vent for non condensed gases and an outlet for the assortment of concentrated product from the bottom. Apart from this, there is feed inlet, an outlet for vapour placed at the top of the dome. The feed is introduced until all the horizontal tubes are adequately immersed. Steam when introduced will transfer the heat to the feed due to temperature gradient. Solvent gets evaporated and the vapour then escapes through the outlet placed at the top. This type of evaporator is best suited for non viscous solutions that do not deposit scale or crystals on evaporation. Its cost per square foot of heating surface is usually less as compared to other types of evaporators.

## *Short Vertical Tube Evaporator*

Though not first to be built, it was the first to gain popularity. It was built by

Robert, director of sugar factory at Seelowitz, Austria in 1850. It was known as Robert type. It became so popular in Europe that it was then later on called *standard evaporator*. Principle being, liquid which has to be evaporated is filled in the vertical tubes and steam is passed outside these tubes. Heat is transferred through the tubes and liquid inside the tubes gets heated. The rescue of the vapours and collection of the concentrated liquid is from the top and bottom, respectively. General construction and working are as shown in the Fig. 3.4, where it is seen that lower portion consists of a group of about, 1000 tubes which are about 1–2 m in length and from 40 to 80 mm in diameter with length to diameter ratio of 15:1. This portion of the evaporator is known as *calandria*. Four important reins are feed valve, to control the feed level which is always kept slightly above the top of the tube, the space above this being left for the rescue of vapour from the boiling liquid. Liquid levels are critical in this type of evaporator, the one present at the bottom is subjected to hydrostatic head of the liquid above it, and hence boiling point may not be reached. Liquid may leave the chamber without complete evaporation, i.e. it may 'short circuit'. Slight increase in the levels will increase the boiling point of the liquid at the bottom, and the capacity of evaporator may decrease. Second control being vent valve to remove any residual air in the steam space. Third being steam valve to provide control to the steam input to the steam space and lastly, the condensate valve to remove the condensate as soon as it is formed so as to maintain high heat transfer rates.

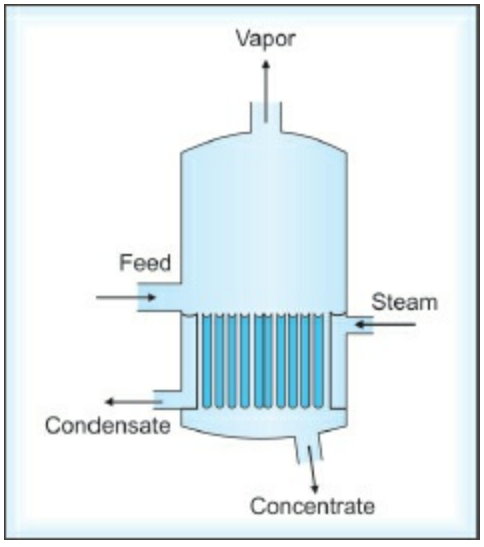



**Fig. 3.4:** Diagrammatic representation of a standard evaporator with

calandria

The first design was build without the downtake which was never satisfactory. But later, changes were proposed and different types of downtakes with different crosssection, located centrally, eccentrically or entirely external to the evaporators were included to improve the performance.

Another short tube evaporator is *basket type*. Here also, the liquid is inside the tubes and steam outside, but the downtake is annular. The whole of the heating filament can be removed as such. Steam may be introduced as shown in the Fig. 3.5. Another important feature of this type is the ease with which a deflector may be added in order to reduce entrainment through spurting. Baffles can also be introduced in these types easily in order to prevent aforementioned losses.

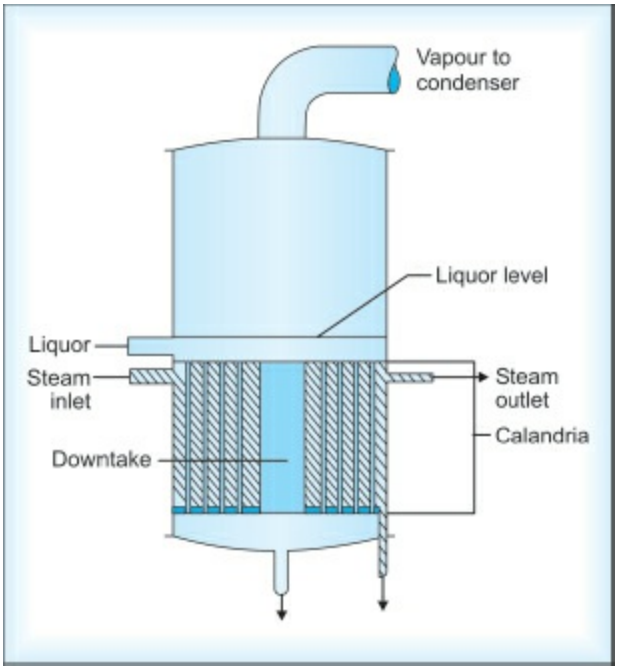



**Fig. 3.5:** Diagrammatic representation of a basket type evaporator

## *Multiple Effect Evaporators*

All of the above discussion is based on what is known as single effect evaporation. An alteration of this system is known as multiple effect evaporation for obtaining higher steam economies than is possible in single effect. In single effect evaporation described earlier, steam provides latent

heat of vapourization to the liquid; the vapours so formed are given off to the condenser, where the heat is given to the cold water. Thus, a great deal of heat goes into the vane. Principle behind multiple effect evaporation is that the number of evaporator units are connected in such a way that it is only the first calandria that is heated by steam; vapours from this effect will then heat the calandria of the next effect (Fig. 3.6). In other words, instead of using cold water, calandria of the next effect is used as a condenser and latent heat of vapourization is being utilized instead of being going into the drain. The vapours of the last effect the then taken to the condenser in the usual way.

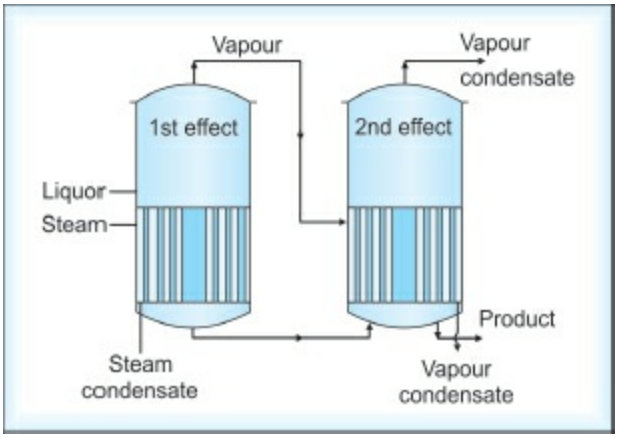



**Fig. 3.6:** Diagrammatic representation of a multiple effect evaporator

The point of consideration while constructing such multiple effect evaporators is method of supplying the effects with feed. Most common is parallel feed method in which separately the feed enters and product is collected from each effect. This method is suitable where not much concentration of the feed is required, whereas if higher concentration is desired then either forward or backward feed methods are used depending upon the requirement. Forward feed is used for scale the forming liquids. The product is obtained at lowest temperature. As feed moves from high pressure to lower pressure chambers, no pumps are used. On the other hand, backward feed method utilizes pumps to move the feed from the last to the first effect. It is suitable for cold feeds and viscous products. Higher concentrated products are at higher temperatures, hence lower viscosity, so heat transfer is maintained high.

Theoretically, any number of evaporators can be added in this way serially but practically this number is limited. The reasons being, provision

for temperature gradients and economics of the whole process. The successive effects should have decreasing temperatures to maintain temperature gradient and allow heat transfer in each effect. This can be done by operating each unit at a lower pressure in order to lower the boiling point of the liquid. Suppose the steam at the first effect is at 120°C temperature and water boils at 100°C at atmospheric pressure so, there is a gradient of 20°C. Now the vapours reaching the next effect are at 100°C, the decrease in the atmospheric pressure of the second effect will decrease the boiling point of the liquid say to 80°C. Again there is gradient of 20°C. In this way, decreasing the pressure in each will limit the number of effects that can be added consecutively. Moreover, increase in the economy of maintenance of each effect is another factor limiting the number of effects. Taking both factors into consideration, the number of effects that can be added is expressed by the following equation:

$$N = \{C/m[\mathrm{U}(T_i - T_N)/\lambda^n(1-f_N)\ W_0]^{1-n}\}^{½} \ldots (11)$$

where, $N$ is the number of effects that can be added, $C$ is the cost of steam per kilogram, $T_i$ is the temperature of input steam in effect I, $T_N$ is the temperature of vapours generated in $N^{th}$ effect, λ is latent heat of vapourization of water, $W_0$ is amount of water in feed solution, $f_N$ is fraction of water remaining in concentrate leaving $N^{th}$ effect, and $m$ and $n$ are constants.

*Capacity*, i.e. mass of vapours produced per unit time is given by the rate of heat transfer. In single effect the rate of heat transfer is given by $Q = UA\Delta T$, whereas in multiple-effect each evaporator has its own rate of heat transfer that is given by $Q_n = U_nA\Delta T$. The net temperature difference i.e. the temperature of the boiling liquid and the condensed film in the steam space is less for multiple-effect as the vapours are produced under pressure and condense in the steam space of the next effect at the lower pressure. Hence, net ΔT is less in multiple effects. Therefore, the capacity of multiple effects is less than single effect evaporator.

*Economy* of any evaporator is expressed as the ratio of total mass of vapours produced to that of the total mass of steam supplied. In single effect the steam produces vapours only once and thus, Economy of a single effect evaporator is 1, whereas in multiple-effect the vapours are produced $n$

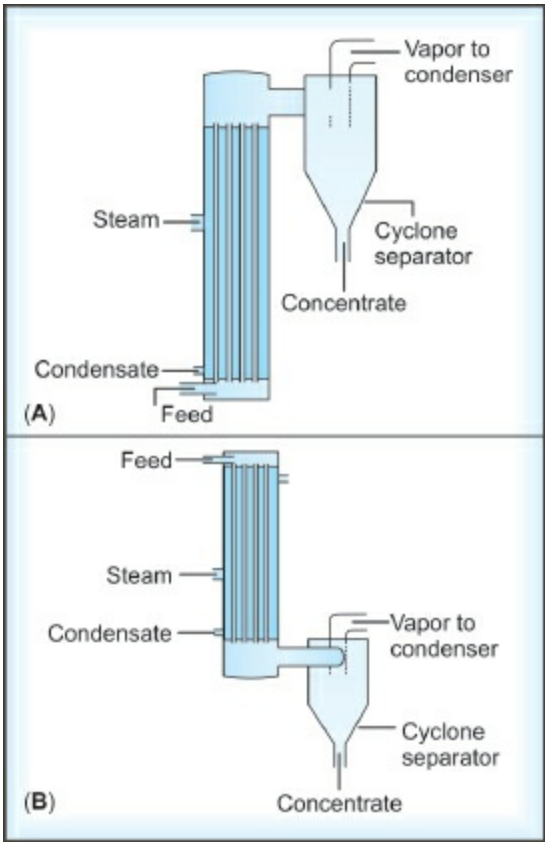
number of times, depending upon the number of effects connected. Hence, economy becomes *n*. In other words the economy of a multi effect evaporator is *N* times the economy of single effect evaporator.

## *Long Tube Evaporators*

As the name suggests these are taller and slender tubes having the L/D ratio greater than 150. The evaporator may have the tubes about 60 mm in diameter and 7 m in length. Such longer sizes allow higher velocities due to violent pumping action of bubbles inside the tubes. Higher velocities are desirable because they tend to decrease the thickness of the viscous film and also assist in the quick removal of the steam bubbles as soon as they are formed. This improves the overall coefficient which is lower for the film on the boiling liquid side and comparatively higher for the condensing steam side. The first evaporator utilizing this effect was *Kestner*, which was patented in 1899.

These longer tubes can have either *natural or forced circulation system*. The most common forms of natural circulation are climbing and falling film types in (Figs 3.7A and B). In *climbing film evaporators*, the feed enters at the bottom of the evaporator, is preheated and subjected to the pressure of column above it. Liquid flows up through the heated tubes as a thin film. The explanation to this can be owed to the vapourization of the liquid near the walls to form bubbles of smaller size which unite together to form larger ones, filling the width of the tube and trapping a slug of liquid above the bubble. The tube gets filled with the vapour, whereas liquid is spread as the film over the walls. It is the only reason of the climbing of film and not the reduced pressure. The evaporators function equally under reduced and atmospheric pressures. As the vapour rises the film is also dragged upwards. It travels up with the velocities of about 5–8 m/s. Such high velocities increase turbulence and heat transfer rates, and also make the evaporator suitable for heat-sensitive substances as the time of contact between the heating surface and liquid is very short. The layer of the liquid which is blown up from the top, strikes deflector (entrainment separator) kept above, which separates both liquid concentrate and vapours. The entrainment separator makes it suitable for use with foam-forming liquids as it can easily break the foams. Feed rate is a critical factor as higher feed rates can lead to insufficient concentration and lower feed rates will not maintain the proper

films. The equipments have the disadvantage that these are not suitable for viscous, salting and scaling liquids.




**Figs 3.7A and B:** Diagrammatic representation of long tube evaporators with natural circulation system: (A) Climbing film evaporator, (B) Falling film evaporator

To overcome this disadvantage, an alternative form of this evaporator was designed which was called as *falling film evaporator*. It is the inverted form of climbing film type so that the feed enters over the top of the tubes and the concentrate, and vapours leave at the bottom. Movement of the liquid film is supported by gravity. Moderately viscous fluids can easily be handled. Even feed distribution is accomplished using a perforated plate above the tubes or utilizing spray nozzles. Such an arrangement makes the equipment unsuitable for suspensions, scaling and salting liquids as solids can clog the perforated plates.

Another form of the long tube evaporator is *forced circulation evaporator*. In climbing and falling film, natural circulation of the liquid was

observed. Only the low viscosity fluids can be adequately treated by natural circulation but liquids having considerably higher viscosities cannot be handled proficiently. Therefore, long tube evaporators were modified to work under forced circulation conditions where liquid is circulated through tubes at high pressures by means of a pump. Principle of the most commonly used forced circulation evaporators being maintenance of pressure and hydrostatic head in such a way that liquid does not boil, and the superheated liquid flowing in the tubes *flashes* just while entering the vapour separating chamber as static head is reduced. The construction is shown in Fig. 3.8.

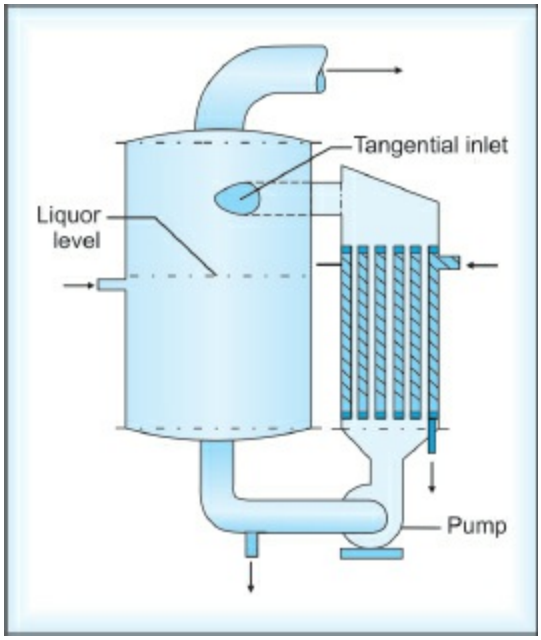



**Fig. 3.8:** Diagrammatic representation of a forced circulation evaporator

For the convenience of an industrial pharmacist these can be classified into type I, type II and type III. *Type I* consists of a bundle of long and narrow tubes (usually ¾ in inside diameter and 8 ft long) contained between two tube sheets. The part of these steam-jacketed tubes project into the vapour head which consists of a deflector/cyclone separator. This is connected to a pipe which runs downwards and enters into the inlet of the pump which delivers the liquid to the tubes with a positive velocity. As the liquid traverse the tube, it becomes heated and begins to boil so that, vapours and liquid are ejected from the top at high velocities. Finally, on striking the cyclone separator, liquid moves back to the tubes and vapours are collected separately. This design suffers from the disadvantage of long height and large

sizes. This problem is surmounted by *Type II* in which there is a separate two pass heater. The heater is not the part of the evaporator body as in case of type I. Pump takes the liquid from the evaporator body and delivers it to the heating tubes. The hydrostatic pressure is maintained here to prevent any boiling in the tubes. Vapourization takes effect only after the pressure is released after entering into the evaporator body which then happens to behave as *flash chamber*. As there is no boiling in the tubes, all heat is imparted as sensible heat resulting in the temperature rise of the circulating liquid that reduces the overall $\Delta T$ of heat transfer. After entering into the evaporator body, liquid settles down and moves through pump into the heating chamber and vapours can be collected from the top. This design is competently used for concentrating materials that deposit crystalline solids on evaporation such as salt brines as there is no evaporation and thus supersaturation, scaling and salting in the tubes. Here, around the heater and evaporator body the slurry of the solids is circulated and it poses problem in case of the solids which grow as coarse crystals. The *type III* is well adapted for such crystals as heating, flashing and crystallization are well separated. The crystallizing solids are maintained as a fluidized bed in the chamber below the vapour head and little or no solids circulate through the heater and flash chamber.

## *Wiped Film Evaporator*

Falling film evaporator suffers from the major drawback of large temperature differences and hot spots. In order to trounce it, evaporators are attached with wipers or other suitable rotary device. As with the falling film evaporator the liquid enters onto the heated wall from the top but a fast rotating wiper filament spreads it perfunctorily. The vapours produced flow in upward direction counter currently. The concentrated solution is haggard off at the bottom of the evaporator. The solution of viscosities up to 2,800–3,200 centipoises can be handled. Rotor speed is maintained between 200 and 400 rpm. Various designs of wiper blades are available. Scarpers and solid wipers are the commonest of all. *SAMBAY® evaporator* (Fig. 3.9) consists of a central rotor tube carrying movable wiper blades. This special rotor design provides efficient squeezing of residues at relatively low rotor speeds. With this rotor, it is possible to handle materials with viscosities up to about 35,000 mPas and materials having tendency to form crust or coatings on the wall. The heating jacket is also divided into a number of separately heated

segments and therefore, temperature conditions in each segment can be easily adjusted. The evaporator can be adapted optimally to a wide variety of applications by varying the wiper blade frequency and pressure applied by the blades.

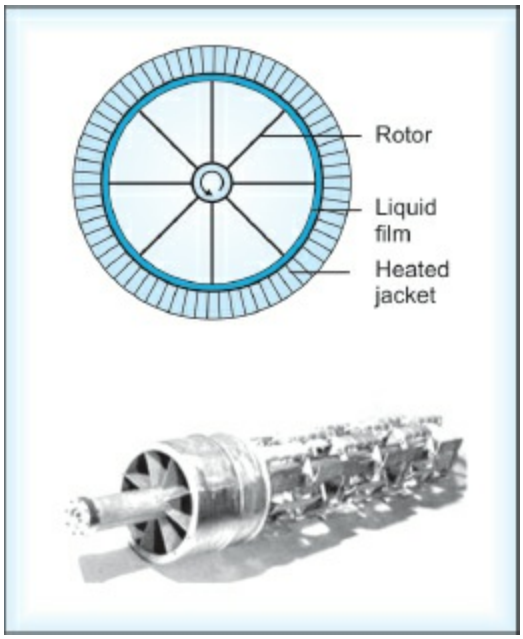



**Fig. 3.9:** Cross section of SAMBAY® wiped film evaporator and SAMBAY® rotor

## *Rotary Evaporator*

Rotary evaporator, often called as ‘rotovap’ is of great importance to the pharmaceutical industry (Fig. 3.10). Vacuum is a key component of this device, and is used to aid in the evaporation of the solvent. Vacuum will lower the air pressure above a liquid, thereby lowering the boiling point of the liquid being heated. Use of a rotavap therefore allows liquid solvents to be removed without excessive heating. Rotary evaporation is most conveniently applied to separate low boiling solvents such as ethyl acetate or n-hexane from compounds which are solid at room temperature and pressure. Evaporation of high boiling hydrogen bond-forming solvents such as water is often a last recourse, due to the fact that in such solvents, the tendency to “bump” is accentuated. The modern centrifugal rotary evaporator is particularly useful for high boiling solvents and when one has many samples to do in parallel. A centrifugal rotary evaporator comprises of a vacuum

pump connected to a centrifuge chamber in which the samples are placed. The centrifugal force creates a pressure gradient within the solvent contained in the tubes, this means that the samples boil from the top down, helping to prevent bumping. The most advanced systems apply the vacuum slowly and run the rotor at speeds of 500 × gravity.

**Fig. 3.10:** Laboratory scale rotary evaporator

## Factors Affecting Evaporation

The process of evaporation is based on several factors and their relationship may be expressed mathematically as:

$$M = K\,S\,(b\text{-}b')/p \ldots (12)$$

where, $M$ is mass of vapour formed in unit time ($m^3/s$), $K$ is a constant (m/s), $S$ is surface area of the liquid exposed ($m^2$), $b$ stands for the maximum vapour pressure at the temperature of air, $b'$ stands for pressure due to vapour of the liquid actually present in the air and $p$ stands for atmospheric pressure.

Therefore, by the formula we can observe that evaporation is directly proportional to the surface of the liquid exposed, temperature used during the evaporation, dryness of the air and indirectly proportional to the atmospheric pressure.

### *Surface Area of the Liquid*

Equation depicts that with increase in surface area of the liquid exposed the rate of evaporation increases. This justifies the efficiency of hemispherical pans providing maximum surface area.

### *Pressure*

The relation between the vapour pressure of the liquid and its evaporation rate is directly proportional. Liquids with low boiling point have high vapour pressure and such liquids evaporate quickly. If the outer atmosphere is dry the value of $b'$ will be low and will be greater. If the vapours of the liquid are removed as soon as they are formed (applying vacuum) the space above the liquid does not become saturated with the vapours and hence evaporation proceeds faster. Furthermore, applying vacuum to the system also reduces the atmospheric pressure and facilitates evaporation.

### *Temperature*

For the substances that decompose at temperatures below 100°C, the evaporation temperatures have to be carefully selected so, these operations are generally carried at reduced pressures. Residence time also plays a significant role. Therefore, evaporation carried out at relatively high temperatures but for a very short period of time is less destructive as

compared to lower temperatures for a longer period of time. Selection of evaporating temperature also depends upon the moisture because for some cases the moisture is destructive as it can cause hydrolysis whereas, for others water may be required as a reaction medium.

## *Type of the Product*

The method of evaporation used also depends upon the type of the product desired. Pans or stills can be used to obtain liquid or dry products, respectively, whereas film evaporators will yield only liquid products.

## *Concentration*

With evaporation liquid becomes concentrated, and its viscosity increases. The increasing solid content elevates the boiling point of the solution. Moreover, increased viscosity will adversely affect heat transfer and also there will be greater risk of decomposition of the thermolabile substances.

## *Economic Considerations*

The method of evaporation should provide economy of labour, fuel and materials. The recovery of solvents and heat also contributes significantly to cost reduction. A multiple-effect evaporator is operated primarily with the aim to achieve increased steam economies.

## Selection of Evaporators

Selection of evaporator relies on the heat transfer coefficient under the desired operating conditions. When power is required to induce circulation past the heating surface, the coefficient must be even higher to counterbalance the cost of power of circulation. Another factor on which evaporator economy is rated is the use of steam. Pounds of solvent evaporated per pound of the steam used affects the evaporator economy. Heat is required to raise the temperature of the liquid from its initial point to the boiling temperature. This can be achieved by either making use of the pressure in order to reduce the boiling point temperature or by making use heat interchange mechanism between feed and residual product. Maximum steam economy is achieved by reusing the vapourized solvent in multiple-effect evaporators.

Other points that must be taken into consideration include crystallization, salting, scaling, product quality, corrosion, and foaming, etc.

## DISTILLATION

For the purpose of industrial pharmacist, distillation can be defined as the separation of the constituents of the mixture including the liquid by partial vapourization of the mixture with subsequent condensation at another place. The term can be employed only at those places where vapourization of the liquid mixture yields a vapour phase with more than one component and it is desired to recover one or more of these constituents in a nearly pure state. The process involves two steps: converting a liquid into vapour phase and transferring these vapors to another place with subsequent condensation to recover the liquid. The feed liquid is known as *distilland* and the condensed liquid is known as *distillate or condensate*.

## Fundamentals

*Vapour-liquid equilibria:* Any distillation problem requires the data of equilibria between liquid and vapour phase of the system. This data can be obtained by what are known as *boiling point diagrams*. Subsequent discussion is limited to systems containing only two substances that are volatile, as multicomponent systems are rarely met within the pharmaceutical industry. These systems can be classified on the basis of miscibility: mixtures with completely-miscible components, mixtures with partially-miscible components and mixtures with immiscible components. Before we get into the details of each type it is important to understand boiling point diagrams.

Figure 3.11 represents the typical boiling point diagram of mixtures of liquid A and liquid B at all compositions and constant pressure. In these diagrams, temperatures are plotted as ordinates and compositions as abscissas. The two curves coincide at the ends. Any point on “y” on the upper curve has its abscissa giving the composition of the vapour that will just begin to condense at the temperature given by the ordinate. Whereas, any point on the lower curve “x”, will give the composition of the liquid at its abscissa that will just begin to boil at the temperature given by its ordinate. The vapour-liquid equilibria data at a particular temperature can be obtained by drawing a horizontal line from ordinate (at that temperature) coinciding both the curves. Corresponding abscissa units will give the composition of liquid and vapour phase in equilibrium with each other. At any point above the upper curve, the mixture is entirely vapour and at any point below the bottom line, the mixture is completely liquefied. If the vapours “y” are completely condensed then they will yield liquid with the composition corresponding to the point obtained by coinciding vertical line from “y” to the lower curve to abscissa. If this liquid is heated, it will give vapours, and thus the point moves towards they left of “y”. Continued distillation of the subsequent residual liquids obtained, will increase their boiling points gradually due to increase in the concentration of the less volatile component. If distillation is still further continued, pure B will be finally left in the still.

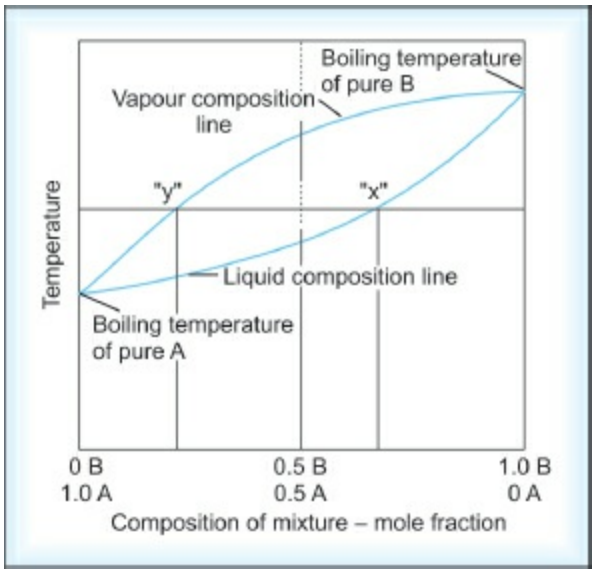



**Fig. 3.11:** Boiling point diagram of an ideal binary mixture at constant pressure

From these considerations, it can be observed that with a rough initial fractionation and repeated evaporation and condensation of the intermediate mixtures, the entire liquid mixture could eventually be separated into two fractions containing pure ‘A’ and pure ‘B’.

The data for these boiling point diagrams are determined experimentally by independently vapourizing mixtures of different compositions, collecting the condensate such that boiling point of the liquid does not appreciably change and then determining the compositions of both the liquid and condensate by a suitable method. However, for *ideal solutions* it is possible to compute the data from vapour pressure data of the pure components. Ideal solutions are those in which there is no change in the property of the components when they are mixed together. The final volume is additive of the two individual volumes. One important physicochemical property that characterizes these solutions is vapour pressure. Vapour pressure calculations are based on *Raoult’s law*, which states that at any particular temperature, the partial pressure of one component of the mixture is equal to the product of its mole fraction and vapour pressure of the pure component at that temperature. The total pressure exerted by the system is equal to the sum of the partial vapour pressures of its components (Fig. 3.12).

From the above diagram (Fig. 3.12), the partial pressure of component A is $p_A$ and is given by following equation:

$$P_A = P_A x \ldots (13)$$

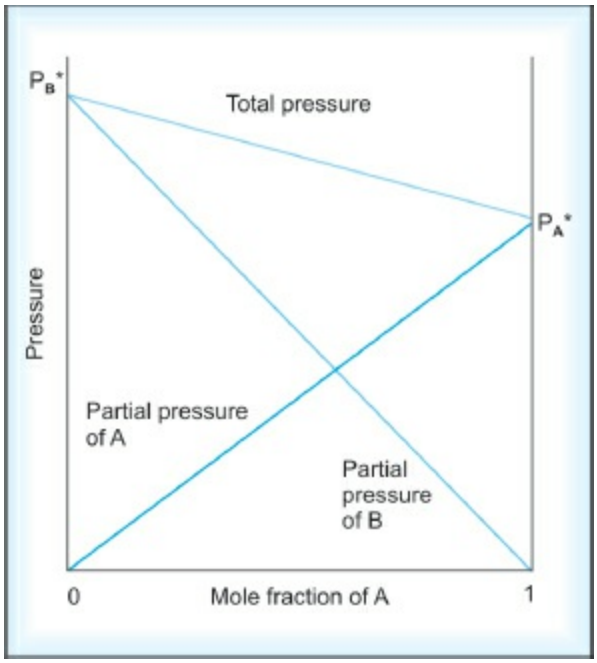



**Fig. 3.12:** Vapor composition diagram for ideal mixtures

where, $p_A$ is partial pressure of A, $P_A$ is vapour pressure of $A$ at that temperature and $x$ is mole fraction of $A$.

Total pressure $P$ is, $P = p_A + p_B \ldots (14)$

Similarly, partial pressure of B is given by,

$$p_B = P_B (1—x) \ldots (15)$$

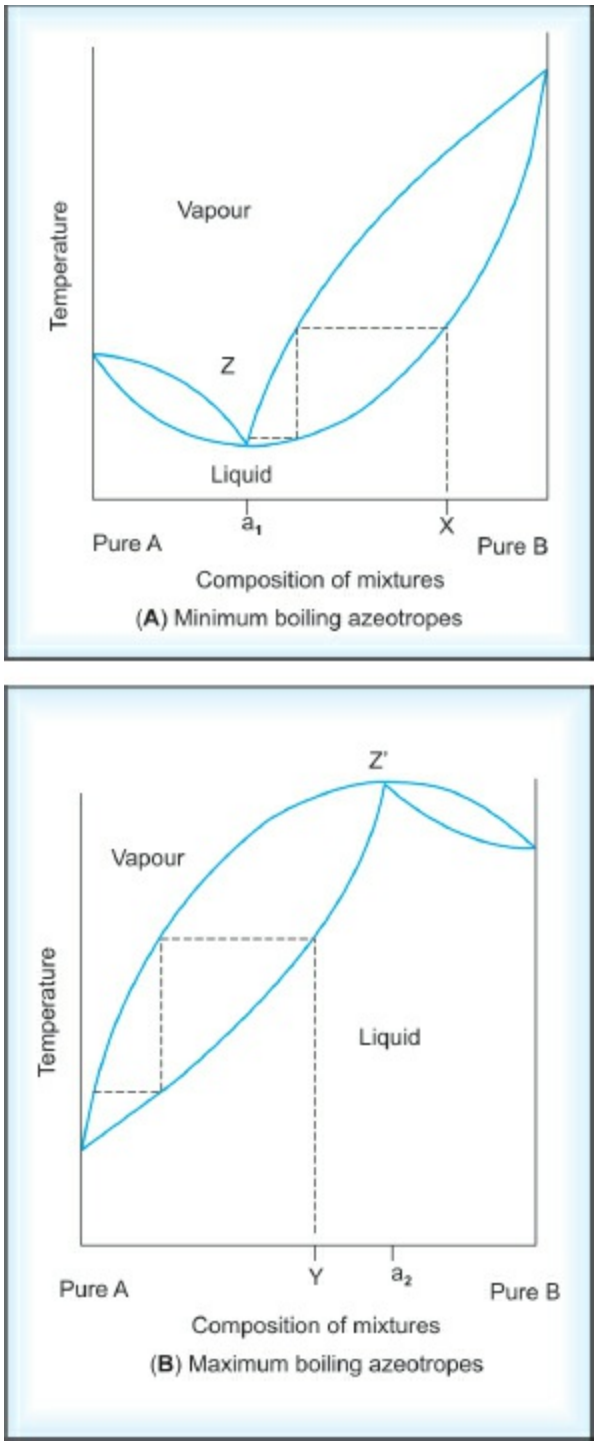
where, $y$, the mole fraction of component Ain vapour, is equal to the ratio of partial pressure of $A$ to the total pressure,i.e.

$$y = p_A/P \ldots (16)$$
$$y = P_A x/P \ldots (17)$$

By choosing temperatures intermediate boiling points of the two pure components, points can be calculated on both the vapour and liquid curves of the boiling point diagram for any given total pressure.

The boiling point diagram described above is only for ideal mixtures, however, practically we encounter many systems which do not obey Raoult's law or in other words which are *non-ideal*. This is due to the interaction

between the components of the mixture. For these systems also the boiling point varies with the composition but this is not as gradual as in case of ideal mixtures. At point A and A* as shown in the illustration (Fig. 3.13) the liquid and vapour curves are tangent showing that vapour will have the same composition as the liquid from which it is formed. Hence, the liquids with compositions A and A* cannot be separated by distillation. These are known as constant-boiling mixtures or *azeotropes*. Azeotropes can be *minimum boiling* or *maximum boiling azeotropes* as shown in Fig. 3.13. In minimum boiling azeotropes, as shown by (a), any point on the left of Z will give pure A and distillate with composition Z, and any point on right will give pure B and constant boiling mixture of composition Z. Examples of such type of mixtures include alcohol-water and benzene-ethanol mixtures. Similarly, in maximum boiling azeotropes, as shown by (b), any point on left of Z' will give pure A and distillate with composition Z' whereas any point on right of Z' will give pure B and azeotropic mixture with composition Z'. Examples of such type of mixtures include chloroform-acetone and nitric acid-water mixtures.




**Figs 3.13A and B:** Boiling point diagrams of minimum and maximum boiling azeotropes: (A) Minimum boiling azeotropes, (B) Maximum boiling azeotropes

*Equilibrium diagrams* are the simplified form of the boiling point

diagrams and give the relation between the vapour pressure and composition at constant pressure. The figures show the equilibrium diagrams for both ideal (Fig. 3.14) and non-ideal (Fig. 3.15) type of mixtures. For non-ideal type of mixtures the point of intersection across the diagonal gives the composition of the constant boiling mixtures. These diagrams can easily be constructed boiling point diagrams by reading the value of Y for different assumed values of X.

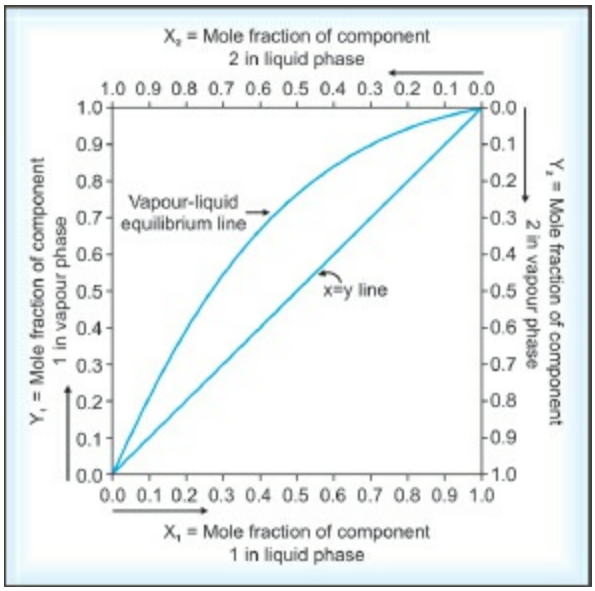
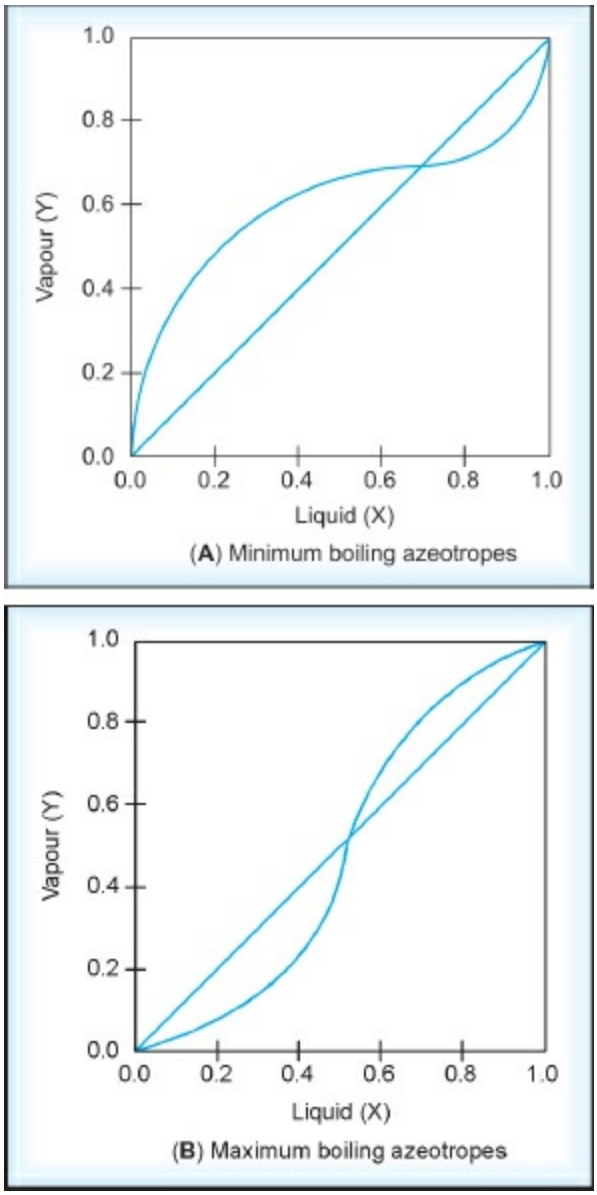



**Fig. 3.14:** Equilibrium diagram of two miscible liquids




**Figs 3.15A and B:** Equilibrium diagrams of minimum and maximum boiling azeotropes: (A) Minimum boiling azeotropes; (B) Maximum boiling azeotropes

## Distillation Methods

### *Simple Distillation*

It is based on the removal and condensation of the vapours as soon as they are formed.

Simple batch distillation is controlled by what is known as Rayleigh's equation.

$$\ln M_0/M_1 = \int_{x_2}^{x_1} dx/y-x \qquad \ldots (18)$$

where, $M_0$ and $M_1$ are the total number of moles of the liquid in the still before and after distillation, respectively. $x_0$ and $x_1$ are the mole fractions of the more volatile component in $M_0$ and $M_1$, respectively. $x$ and $y$ are the liquid vapour composition relationships given by the equilibrium diagram.

This technique is widely employed in the pharmaceutical industry when a volatile component has to be separated from a nonvolatile component. This is sometimes also known as *differential distillation* as it is based on differences in volatilities and vapour pressures of the two components in the mixture. *Volatility* is the ratio of equilibrium partial pressure of the substance in the vapour phase and its mole fraction in the solution. The simple distillation unit consists of three essential parts; distillation flask, condenser and receiver (Fig. 3.16). Mixture to be distilled is placed in the distillation flask with the height not more than two-thirds of the flask, with insulation provided using asbestos rope at its upper part to prevent the condensation taking place in the flask itself. To prevent the superheating or bumping; pumice stone, glass beads or a few pieces of porous plate are added to the distillation flask.

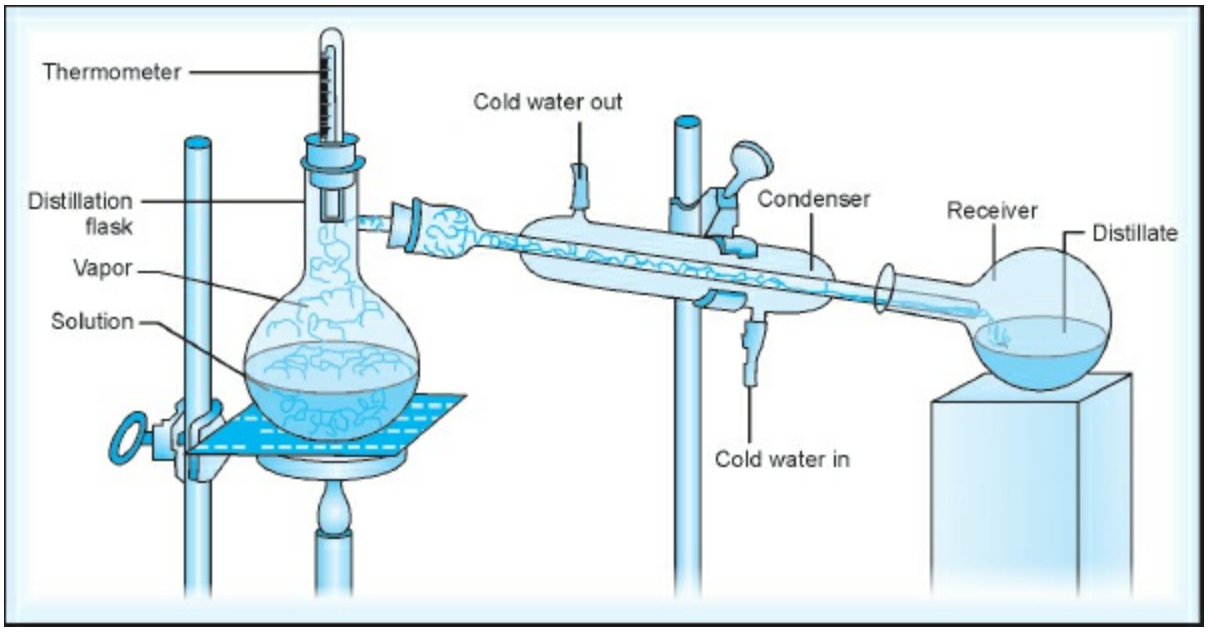



**Fig. 3.16:** Set up of a simple distillation unit

Heating of the flask is done on the basis of boiling points of the liquids in the mixture. If the boiling point is below 100°C then conveniently water or steam bath can be used. For temperatures above 220°C Woods's or Rose's metal bath is used and for still higher temperatures like more than 250°C, oil bath is most suitable. Heating mantles are also rarely used for temperatures around 400°C.

## *Vacuum Distillation/Distillation under Reduced Pressure*

For substances which are thermolabile or have high boiling points or undergo decomposition at atmospheric pressure, distillation is carried out at low pressures and hence lower temperatures. This is known as vacuum distillation. It is of great importance to the pharmacists for the separation of the menstruum which is used in the preparation of various gelanical preparations. Vacuum can also be used for recovering last traces of solvent from a semisolid residues. The distillation unit consists of a *claissen flask* with a fine capillary system to pass a carbon dioxide or nitrogen as stream of bubbles for smooth ebullition; water bath for concentration of aqueous or alcoholic extracts or for distillation of relatively high boiling liquids, and condenser and receiver are same as in differential distillation unit. The major drawback of this system is foaming. To suppress it, antifoaming agents such as capryl or octyl alcohol, alkaterage C or silicones are generally added.

## *Flash Distillation*

It is defined as the process in which the entire liquid is vapourized as a flash as it passes from a high pressure zone to a low pressure zone. The vapours are allowed to cool. The vapour phase molecules of high boiling fraction get condensed, while low boiling fractions remain as a vapour. Severance is endeavored when both liquid and vapour phases are in equilibrium with each other and therefore, the name *equilibrium distillation* is also given to it. The process is continuous and used for multicomponent systems of narrow boiling range especially in oil refineries, petroleum industries, etc. The process is not apposite for the separation of substances with similar *volatilities*. The design of unit is as illustrated in Fig. 3.17. The feed is pumped into the heater at a certain pressure. Through a pressure reducing valve liquid tends to move into vapour liquid separator where due to drop in the pressure, the hot liquid flashes which further enhance the vapourization process. Impulsive vapourization induces cooling. The vapour phase molecules of high boiling fraction get condensed, while low boiling fraction remains as a vapour. The mixture is given sufficient time so that vapour and liquid portions separate and achieve equilibrium. The vapours and condensate are collected from the top and bottom, respectively. The continuity in the process can be maintained by adjusting the operating conditions in such a way that the amount of vapours leaving the chamber is equal to the feed entering the unit so that, at any time vapour and liquid concentrations remain constant within the system.

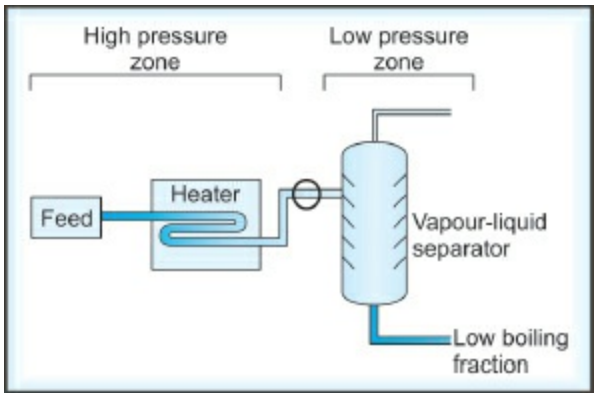



**Fig. 3.17:** Diagrammatic representation of a flash distillation unit

## *Rectification/Fractional Distillation*

If one considers the boiling point diagram of an ideal binary mixture (Fig. 3.11), it could be easily observed that, it is only with the tedious repeated simple distillations that fairly sharp separation of two components can be obtained. These repeated distillations are termed as *fractionation*. Instead of repeating the whole procedure again and again, a better degree of separation is easily achieved using a fractionating column between the still and the condenser, and the resulting method is termed as *rectification*. The method primarily differs from simple distillation in returning the part of the condensate to the still by admitting it to the top of the fractionating column to enable the recurring liquid to come into close contact with the vapour flowing in the direction of the condenser. The construction or schematic diagram of the rectification unit is shown in Fig. 3.18. The major parts of the unit are: *still, fractionating column and condenser*. Still is the boiler in which the liquid is boiled and the vapours are generated. Fractionating column is the tower to bring both the vapours and part of the liquid condensate to come into intimate contact with each other. Condenser is to condense the vapours leaving the top of the column and to send the part of the condensate back to the column. This part returning back is termed as *reflux* and the ratio of the reflux to the product is termed as *reflux ratio*.

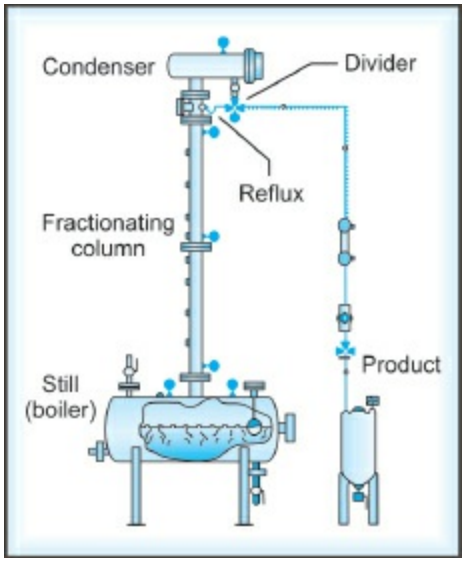



**Fig. 3.18:** Diagrammatic representation of a fractional distillation unit

## *Mechanism*

Within the rectification column, both ascending vapours from the still and

condensate sent back to the column come into intimate contact with each other. Assuming no heat loss to the surroundings, the descending liquid stream causes the condensation of the higher boiling component from the rising vapours. This is termed as *scrubbing*. Heat so liberated makes the ascending vapour stream to vapourize the lower boiling component from the descending liquid. This is termed as *stripping*. Therefore, the plates in the column which are nearer to the still will have greater fraction of the high boiling component and plates near the condenser will be richer in lower boiling component, and at the middle of the column there will be an intermediate ratio of both the components. Moreover, as the liquid descends the column it progressively becomes hotter and as vapours rises through the column they progressively become cooler. The result is the maintenance of temperature and concentration gradients throughout the tower. With ideal mixtures, a properly designed columns yields from the top of the column a vapour which constitutes practically pure lower boiling component and liquid from the still which is practically pure higher boiling component.

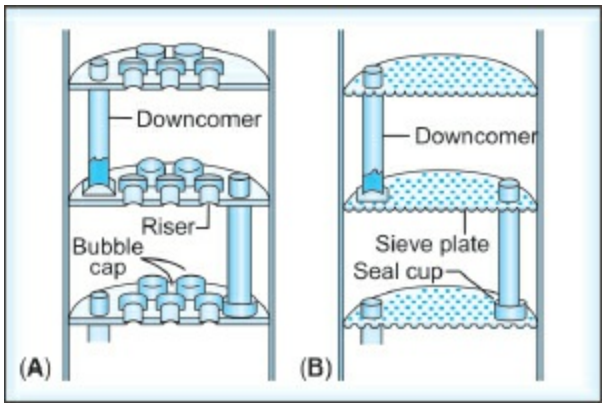
Practically, instead of using still, the feed to be separated is directly fed into a suitable plate of the column known as *feed plate*, which is somewhere within the height of the column. The part of the column which is above this feed plate is termed as *rectifying section* while the part below the feed plate is termed as *stripping section*.

## *Construction of Rectifying Columns*

In general, two types of the columns are used: plate columns and packed columns. *Plate columns* are further classified into bubble cap columns and sieve plate columns. *Bubble cap columns*, as depicted from the name, consist of hundreds of bubble caps onto the horizontal plates present throughout the column. Each bubble cap has a nipple with a bell or rectangular-shaped cap over it as shown in the diagram below (Fig. 3.19A). Each edge or wall of the cap has slots. Vapours rising above move through the nipples and are diverted by the cap and finally bubbled through the slots. A layer of liquid is maintained on the plate by means of an overflow such that the openings in the cap are submerged. Downpipe from the plate above is sealed by the liquid on the plate to ensure that vapour does not enter the pipe. A layer of liquid is maintained from the pipe above and pipe below. Both the pipes are arranged so that they are opposite to each other to ensure that liquid does not short

circuit but flows uniformly across the plate before it is delivered to the lower plate.

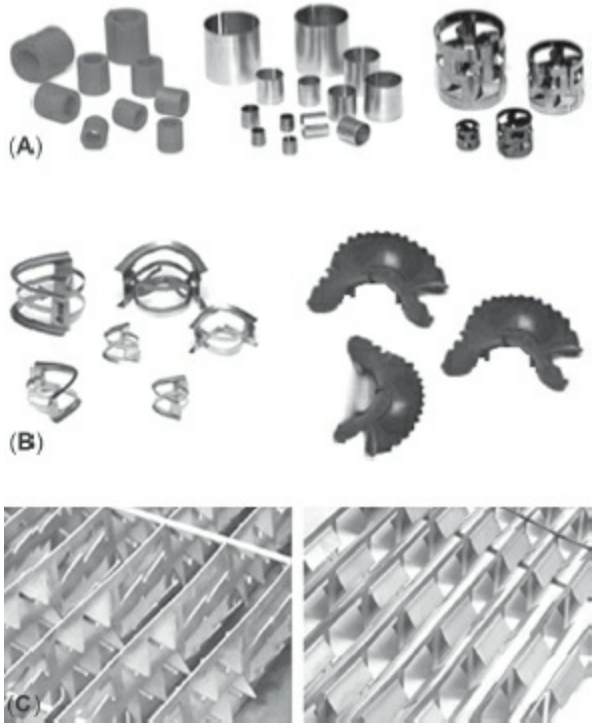
*Sieve plate columns* contain flat plates with a large number of perforations (Fig. 3.19B). The vapours move above through these perforations and pressure exerted by them prevents the downflow of the liquid through these columns. Liquid moves plate-to-plate in the same way as in the case of sieve plate column. An extension of this type is turbo grid column, which consists of a number of horizontal bars with narrow slots between them. Liquids flow through the bars and vapours rise through the slots.




**Figs 3.19A and B:** Diagrammatic representation of plate columns: (A) Bubble cap column; (B) Sieve plate column

*Packed columns* are preferred to sieve plate and bubble cap columns in processes where two process-phases are brought into intimate contact. A number of packing materials are available and some of the important designs are shown in Figs 3.20A to C. Broadly, they are classified into two types which include *general purpose and high efficiency types*. General purpose includes either *rings or saddle-type*. Simplest of the rings is *Raschig type* which is a hollow cylinder, the diameter of which is equal to the length. Modified Raschig types include lessing, spiral and cross-partition types. Among saddles, *Berl saddles* are the most popular. High efficiency types include *structured grid* packings, where packing is installed in rigid modules stacked in successive layers with a fixed orientation, which minimizes the overall pressure drop while simultaneously increasing tower capacity and/or efficiency. These packings have a number of operational advantages and are developed primarily for severe services or those susceptible to fouling,

erosion, coking and high solids content. Packed columns are preferred over plate columns for the following reasons:




**Figs 3.20A to C:** Types of column packings: (A) Ring packing; (B) Saddle packing; (C) Structured grid packing

*Low-pressure drop:* With a typical range of 0.35–2 mm Hg per theoretical plate as compared to 3.05 mm of Hg per plate, the benefit occurring from a lowered pressure drop includes; A reduced base pressure/temperature, lower rates of degradation, lower levels of heat source, and enhanced relative velocity.

*Capacity:* Depending on flow parameters, the capacity of a properly selected and packed column is greater than plate column.

*Liquid hold up:* The total hold up in packed tower is typically 2–3% higher as compared to about 10–12% of a plate column. In other words, residence of liquid phase in a plate would be 3–4 times that in a packed column. The low residence time leads to rapid responses. The time required to establish stability on specific conditions is a fraction of that normal for a plate column.

*High pressure/foaming systems:* Packed columns provide good resistance

to foaming provided that polymerization or crystallisation does not occur in vessel. However, it must be noted here that foaming is a major problem in closed-sided packings like Rasching rings. But with the advent of open packings, this problem has been overcome to a very great extent.

## *Efficiency of Fractional Distillation (McCabe-Thiele Theory)*

In the design of a fractionating column, the primary information required includes the determination of the number of plates and the column diameter. One method of determining the number of plates is to proceed with a plate-to-plate calculation starting from the known terminal conditions. Either end of the column may be selected. But this method needs some additional information other than the flow rate, composition, and enthalpy of reflux and vapour to condenser. It is therefore convenient to introduce the concept of a *theoretical plate* or *equilibrium stage*.

A theoretical plate or equilibrium stage may be defined as a plate holding a pool of liquid from which a vapour rises whose average composition is in equilibrium with a liquid whose composition is the average of that of the liquid leaving the plate (assuming that both were thoroughly mixed). Using this concept, the plate calculations will be made for theoretical plates.

The efficiency of the fractional distillation unit can be estimated by two parameters: $N_t$, the number of theoretical plates and H, the height equivalent to a theoretical plate (HETP), which is defined as $L/N_t$, where L is the length of the column.

An actual, physical plate is rarely at 100% equilibrium stage and the number of actual plates is more than the required theoretical plates.

$$N_a = \frac{N_t}{E} \qquad \text{... (19)}$$

where, $N_\alpha$ = the number of actual, physical plates or trays

$N_t$ = the number of theoretical plates or trays

$E$ = the plate or tray efficiency

To design a distillation unit, the number of theoretical plates, $N_t$, required in the process is determined, taking into consideration a likely range of mixture composition and the desired degree of separation of the components in the resultant fractions. In industrial continuous fractionating columns, $N_t$ is

determined by beginning at either the top or bottom of the column and calculating material balances, heat balances and equilibrium vapourizations for each of the successive equilibrium stages until the desired end product composition is attained.

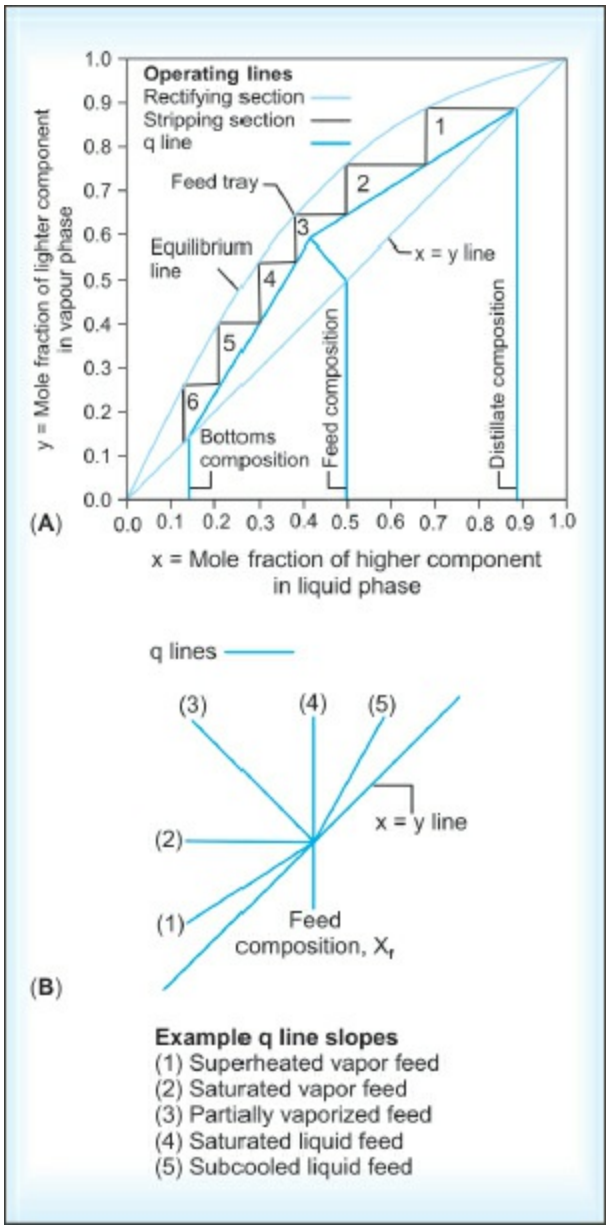
In 1925, McCabe-Thiele presented a graphical approach, the simplest and perhaps the most instructive method to determine the number of theoretical plates for binary distillation units. The McCabe-Thiele method uses the fact that the composition at each theoretical plate (or equilibrium stage) is completely determined by the mole fraction of one of the two components and is based on the assumption of constant molar overflow which requires that:

- The molal heats of vapourization of the feed components are equal.
- For every mole of liquid vapourized, a mole of vapour is condensed.
- Heat effects such as heats of solution and heat transfer to and from the distillation column are negligible.

## *The McCabe-Thiele Diagram*

The graphical solution for the determination of theoretical plates by McCabe-Thiele method requires the following steps:

1. The vapour-liquid equilibrium (VLE) data for the lower-boiling component of the binary feed are obtained.
2. Equal-sized vertical and horizontal axes of a graph are drawn. The horizontal axis (x) is for the mole fraction of the lower-boiling feed component in the liquid phase. The vertical axis (y) is for the mole fraction of the lower-boiling feed component in the vapour phase.
3. A straight line from the origin of the graph to the point where × and y both equal 1.0 is drawn, which is the × = y line in Fig. 3.21. Then, the equilibrium line using the VLE data points of the lower boiling component, representing the equilibrium vapour phase compositions for each value of liquid phase composition is drawn. Vertical lines from the horizontal axis up to the × = y line for the feed ($x_f$ to g) and for the desired compositions of the top distillate product ($x_d$ to a) and the corresponding bottoms product ($x_w$ to j) are also drawn.




**Figs 3.21A and B:** (A) Typical McCabe-Thiele diagram for fractional column problems; (B) Examples of q line slopes

4. The operating line (a to h) for the rectifying section (the section above the feed inlet) of the distillation column, is drawn. Starting at the intersection of the distillate composition line and the × = y line (a), the rectifying operating line is drawn at a downward slope ($\Delta y/\Delta x$) of L/(D + L), where, L is the molar flow rate of reflux and D is the molar flow rate of the distillate product. For example, in Fig. 3.21, assuming the molar flow rate of the reflux L is 1000 moles per hour and the molar flow rate of the

distillate D is 590 moles per hour, then the downward slope of the rectifying operating line is 1000/(590 + 1000) = 0.63 which means that the y-coordinate of any point on the line decreases 0.63 units for each unit that the x-coordinate decreases.

5. The q line from the × = y line is drawn to intersect the rectifying operating line (g to h). The parameter q is the mole fraction of liquid in the feed and the slope of the q line is q/(q-1). For example, if the feed is a saturated liquid, it has no vapour. Thus q = 1 and the slope of the q line is infinite which means the line is vertical. As another example, if the feed is all saturated vapour, q = 0 and the slope of the q line is 0 which means that the line is horizontal. Some examples of q line slopes are presented in Fig. 3.21B. As can be seen now, the typical McCabe-Thiele diagram in Fig. 3.21A uses a q line representing a partially-vapourized feed.
6. Starting from the intersection of the bottoms composition line and the x = y line, the stripping section (i.e., the section below the feed inlet) operating line is drawn up to the point where the q line intersects the rectifying section operating line.
7. The steps between operating lines and the equilibrium line are drawn and then counted. These steps represent the theoretical plates (or equilibrium stages). The required number of theoretical plates is 6 for the binary distillation depicted in Fig. 3.21.

In continuous distillation with varying reflux ratio, the mole fraction of the lighter component in the top part of the distillation column will decrease as the reflux ratio decreases. Each new reflux ratio will alter the slope of the rectifying section operating line.

In an industrial distillation column, the $N_t$ required to attain a given separation also depends upon the amount of reflux used. Using more reflux decreases the number of plates required and using less reflux increases the number of plates required. Hence, the $N_t$ is calculated at various reflux rates. $N_t$ is then divided by the plate efficiency, E, to determine the actual number of physical plates, $N_a$, needed in the separating column. Depending upon the economic balance, the final design is made which tells the number to be actually installed in the distillation unit.

## *Steam Distillation*

When a mixture of two practically immiscible, steam-volatile liquids which are not decomposed by water and whose boiling points are too high and components too heat sensitive for distillation at atmospheric pressure, are heated with agitation to expose the surfaces of both the liquids to the vapour phase, each component independently exerts its own vapour pressure as a function of temperature as if the other component were not present. Consequently, the vapour pressure of the whole system increases. When the sum of the partial pressures of the two non miscible components just exceeds the atmospheric pressure, boiling begins. In this way, many organic compounds insoluble in water can be purified at a temperature well below the point at which decomposition occurs. For example, the boiling point of bromobenzene is 156°C and the boiling point of water is 100°C, but a mixture of the two boils at 95° C. Thus, bromobenzene can be easily distilled at a temperature 61°C below its normal boiling point.

The design of the apparatus is as shown in the Fig. 3.22. Steam can be introduced either directly as *open steam*, or can be supplied in a jacketed form. Steam supplied in an open manner is directly bubbled through the still containing the preheated liquid mixture while keeping the temperature of the mixture constantly high and allowing the vapours so formed to condense in a water condenser. Sometimes, steam is bubbled keeping the mixture temperature low. This will condense the steam within the flask, forming a different aqueous layer which can be decanted and organic layer can be allowed to dry. For the synthesis of volatile oils, perfumes, etc. raw material is placed on a perforated bottom and steam is admitted through the jacket below the material. The water and the volatile vapours are condensed in the condenser. The distillate is separated into two layers and is collected by means of a *florentine reciever*. For oils having higher specific gravity, the system is reversed and for oils having specific gravity close to water entire distillate is collected and extracted with an immiscible volatile solvent. Eucalyptus oil and orange oil are obtained by this method on the industrial scale. Steam distillation is also widely used in petroleum refineries and petrochemical plants where it is commonly referred to as “steam stripping”.

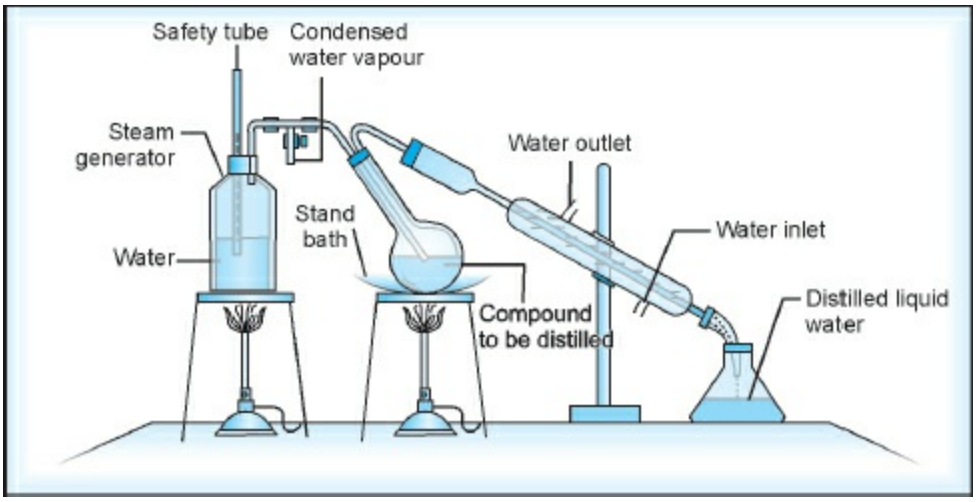



**Fig. 3.22:** Diagrammatic representation of a steam distillation unit

## *Azeotropic Distillation*

As we know that with simple distillation or rectification, the complete separation of the azeotropes or constant boiling mixtures could not be attained, so, in order to bring this to effect, another extraneous liquid is added to generate a new, lower-boiling azeotrope that is heterogeneous (e.g. producing two immiscible liquid phases). It exaggerates the differences in the volatility of two original components and thus facilitates separation. for example, addition of benzene to water and ethanol.

If the extraneous material is less volatile than the feed then the same operation is referred to as *extractive distillation*. When the solvent is passed through the column containing the mixture, the least affected volatilities are collected from the bottom along with the extraneous material whereas considerably enhanced volatilities due to this material pass overhead. As the extraneous material here is working similar to the solvent used in the solvent extraction system, it is referred to as *extractive distillation*.

If the extraneous material is more volatile than the feed then the operation is referred to as *azeotropic distillation*. Here the extraneous material is added and collected with the overhead material. At all times, volatility characteristics are such that it forms one or more azeotropes usually of minimum boiling type and hence the name *azeotropic distillation* is given to it.

A common example of distillation with an azeotrope is the distillation of

ethanol and water. Using normal distillation techniques, ethanol can only be purified to approximately 96% (hence the 96% strength of some commercially available alcohols). Once at a 96.4% ethanol/water concentration, the vapour from the boiling mixture is also at the same concentration, therefore, further distillation becomes ineffective. The 96.4% azeotrope needs to be “broken” in order to refine further. This is achieved using an extraneous material, benzene, to the ethanol/water mixture that changes the molecular interactions and eliminates the azeotrope, and thus “breaking” is achieved. The azeotrope is fed to the column and benzene is added to the feed. The vapours with ternary azeotropic composition are condensed to give a benzene phase and water phase as a top and bottom layer, respectively. Aqueous phase is used to recover benzene as ternary azeotrope in a decanter. Anhydrous alcohol is collected as a residue from the column.

## *Molecular Distillation*

Under normal conditions, the molecules leaving the evaporator surface collide with each other and with the molecules of the gas present in its vicinity. Ordinary vacuum distillations are also performed at pressure of about 1 to 8 mm Hg. So even under vacuum, the molecule could undergo a number of collisions. As a result of these collisions some of the molecules return to the surface of the evaporating liquid, and thus reduce the effective rate of distillation. Ordinary vacuum can cause the molecule to return thousands of time back to the still before reaching the exit. The frequency of collision depends upon the *mean free path* of the molecules which is defined as the average distance traveled by the single molecule between successive collisions. Its value actually depends upon existing pressure and temperature. Not all molecules travel the distance which is equal to their mean free path. Sometimes, a molecule does not travel any distance without collision. If vacuum is introduced with the residual gas not exceeding $10^{-3}$ mm Hg then the residual gas molecules offer no or minimum obstruction, and hence the mean free path is lengthened. Now, if the distance between the condenser and evaporator is reduced to become equal to the mean free path, then the molecules can reach the condenser without colliding. There would be least hindrance to the motion of the molecules and this type of arrangement is termed as *molecular distillation or short path distillation*.

This type has a number of advantages over normal distillation. As we see under normal conditions a number of collisions are affected, so, a large amount of heat dissipation occurs which is minimized in the case of molecular distillation and hence the rate is much higher for molecular distillation at given conditions of temperature and pressure when compared to normal vacuum or simple distillation set ups. As here distillation can be effected at lower temperatures it can be proficiently used for heat-sensitive materials. The process can be carried out at temperatures about 100°C less as compared to conventional distillation, therefore; it is the operation that involves minimum thermal decomposition. Further, it is the most frequently used method for the preparation of high molecular weight substances due to reduced tendency of oxidation as it involves negligible amount of residual gas molecules. Also, it does not involve any molecular changes and risk of contamination with solvents.

Design of the still can be classified as *batch or continuous type*. A simple *batch type* molecular still consists of an electrically heated tray suspended in an evacuated test tube. It has evaporative efficiency of about 0.5, which that means about 50% of all molecules are able to leave the surface and reach the condenser at first try.

As we have seen earlier that molecular distillation is mostly used for high molecular weight substances, so as in case of simple distillation, the diffusion of the liquid to the vapour phase is not continuous due to the formation of a film. In order to disrupt this film, continuously circulating stills are used. Two types of stills are available, *falling film and centrifugal types* (Fig. 3.23). In both of these, either the distilland is fed from a supply vessel and distillate and residue collected separately, or residue may be refed to the still either intermittently or continuously. Both the types keep the residence time very small and thus time of heating of the thermally sensitive materials as short as possible.

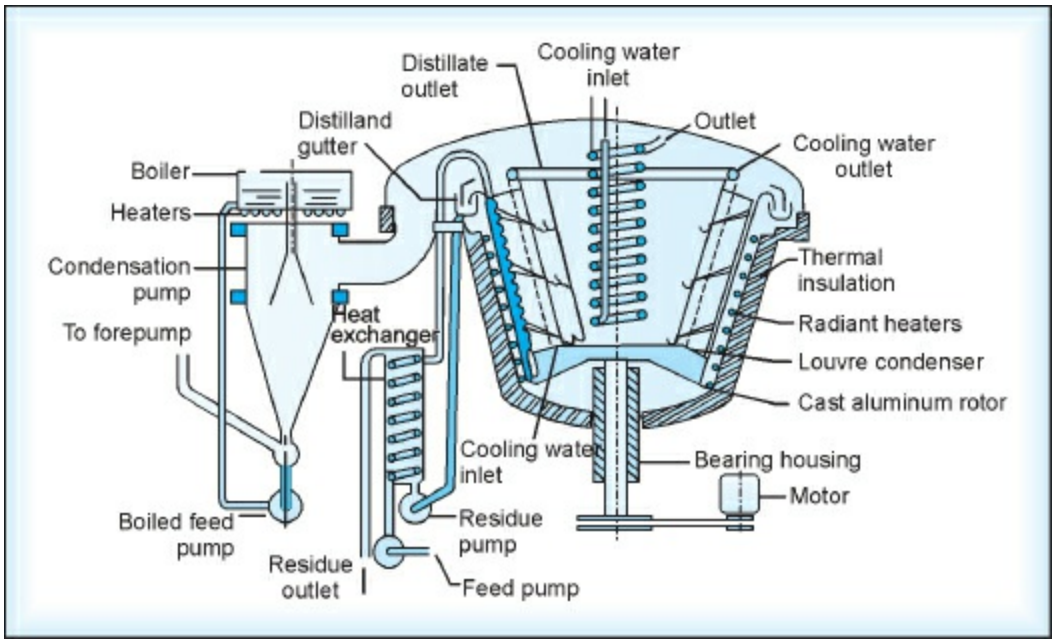



**Fig. 3.23:** Large-scale centrifugal molecular still

## Selection of Distillation Units

Selection of the distillation unit is dependent entirely upon the properties of the mixtures to be separated and the product desired. If the boiling point of the substance is less than 75°C then the mixture is subjected to simple distillation procedure, whereas if any of the components is heat labile or thermosensitive then reduced pressure, vacuum or flash distillation is of great utility. Moreover, the distillation operation to be followed also depends upon the type of the mixtures. For mixtures which are completely miscible with each other, simple, vacuum, rectification and azeotropic distillation can be conveniently used whereas for mixtures which are miscible with certain proportions or are immiscible with each other, steam distillation is most suitable. Not only miscibility but also difference in the volatility of the mixture decides the method to be used. For mixtures which have far apart volatilities can easily be separated using fractionating or rectifying columns, whereas those which have very close volatilities can be separated making use of the third entrainer substance as in the case of extractive or azeotropic distillation.

Other than these factors if the desired product has a higher molecular weight and oxidation has to be avoided during the polymerization reaction then falling film or centrifugal type of molecular stills are most suitable. So, lastly the selection of the suitable stills and units are done keeping all the essential properties and requirements of the material in mind.

# 4:
# Drying

There is hardly a pharmaceutical plant engaged in the manufacture of tablets or capsules that does not contain dryers. Unfortunately, the operation of drying is so taken for granted that efforts for achieving increased efficiency in the production of tablets do not include a study of drying. This chapter introduces the industrial pharmacist to the theory and fundamental concepts of drying.

## DEFINITIONS

For the purpose of this discussion, drying is defined as the removal of a liquid from a material by the application of heat, and is accomplished by the transfer of a liquid from a surface into an unsaturated vapour phase. This definition applies to the removal of a small amount of water from moisture-bearing table salt as well as to the recovery of salt from the sea by evaporation. Drying and evaporation are distinguishable merely by the relative quantities of liquid removed from the solid.

There are, however, many non-thermal methods of drying, for example, the *expression* of a solid to remove liquid (the squeezing of a wetted sponge), the *extraction* of liquid from a solid by the use of a solvent, *adsorption* of water from a solvent by the use of desiccants (anhydrous calcium chloride), *absorption* of moisture from gases by passage through a sulfuric acid column, and *desiccation* of moisture from a solid by placing it in a sealed container with a moisture-removing material (silica gel).

## PURPOSE

Drying is most commonly used in pharmaceutical manufacturing as a unit process in the preparation of granules, which can be dispensed in bulk or converted into tablets or capsules. Another application is found in the processing of materials, e.g. the preparation of dried aluminum hydroxide, spray drying of lactose, and preparation of powdered extracts.

Drying also can be used to reduce bulk and weight, thereby lowering the cost of transportation and storage. Other uses include aiding in the preservation of animal and vegetable drugs by minimizing mold and bacterial growth in moisture-laden material and facilitating comminution by making the dried substance far more friable than the original, water-containing drug.

Dried products often are more stable than moist ones. Drying reduces the chemical reactivity of the remaining water, which is expressed as a reduction in the water activity of the product. Various processes for the removal of moisture are used in the production of these materials. After the moisture is removed, the product is maintained at low water levels by the use of desiccants and/or low moisture transmission packaging materials. The proper application of drying techniques and moisture-protective packaging requires a knowledge of the theory of drying, with particular reference to the concept of equilibrium moisture content.

## THEORY OF DRYING

Drying involves both heat and mass transfer operations. Heat must be transferred to the material to be dried in order to supply the latent heat required for vaporization of the moisture. The drying process can be understood more easily if attention is focused on the film of liquid at the surface of the material being dried. The rate of evaporation of this film is related to the rate of heat transfer by the equation:

$$dW/d\theta = q/\lambda \ldots (1)$$

where, $dW/d\theta$ is the rate of evaporation expressed as pounds of water per hour, $q$ is the overall rate of heat transfer (BTU per hour), and $\lambda$ is the latent heat of vaporization of water (BTU per pound).

Secondly, mass transfer is involved in the diffusion of water through the material to the evaporating surface, in the subsequent evaporation of the water from the surface, and in diffusion of the resultant vapour into the passing air stream. The rate of diffusion of moisture into the air stream is expressed by rate equations similar to those for heat transfer. The driving force is a humidity differential, whereas for heat transfer, it is a temperature differential. The rate equation for mass transfer is as follows:

$$dW/d\theta = k' A(H_s - H_g) \ldots (2)$$

where, $dW/d\theta$ is the rate of diffusion expressed as pounds of water per hour; $k'$ is the coefficient of mass transfer [pounds of water/(hour) (square foot) (absolute humidity difference)]; $A$ is the area of the evaporating surface in square feet; $H_s$ is the absolute humidity at the evaporating surface (pounds of water per pound of dry air); and $H_g$ is the absolute humidity of the passing air stream (pounds of water per pound of dry air).

After an initial period of adjustment, the rate of evaporation is equal to the rate of diffusion of vapor, and the rate of heat transfer [equation (1)] can be equated with the rate of mass transfer [equation (2)], or:

$$dW/d\theta = q/dW/d\lambda = k' A(H_s - H_g) \ldots (3)$$

If the overall rate of heat transfer, q, is expressed as the sum of the rates of heat transfer by convection, radiation, and conduction, equation (4) is expanded to the form:

$$dW/d\theta = (q_c + q_r + q_k)/\lambda = k' A(H_s - H_g) \ldots (4)$$

where, $q_c$, $q_r$, and $q_k$ are the rates of heat transfer by convection, radiation, and conduction, respectively.

The rate of drying may be accelerated by increasing any of the individual terms in equation (4).

The rate of convective heat transfer, *qc*, can be increased by:

1. Increasing the air flow rate
2. Raising the inlet air temperature.

The rate of radiation heat transfer, $q_r$, can be stepped up by:

1. Introducing a high-temperature radiating heat source into the drying chamber.

The rate of conduction heat transfer, *qk*, can be stepped up by:

1. Reducing the thickness of the material being dried
2. Allowing it to come in contact with raised-temperature surfaces
3. Increasing the air velocity increasing the coefficient of mass transfer, *k*' by
4. Dehumidifying the inlet air, thus increasing the humidity differential, ($H_s$-$H_g$), is still another means of speeding up the rate of drying.

Rapid drying may also be accomplished through the application of a microwave or dielectric field. In this case, heat is generated internally by the interaction of applied electromagnetic field with the solvent. Mass transfer results from an internal pressure gradient established by the internal heat generation, while the mass concentration remains relatively uniform. The drying rate, then, primarily depends on the strength of the field applied to the material.

The foregoing discussion holds true as long as there is a film of moisture on the surface of the material being dried. When the surface becomes partially or completely dry, the heat and mass transfer equations become more complex. In this case, the rate of drying is controlled by the rate of diffusion of moisture from the interior of the material. This diffusion is greatly influenced by the molecular and capillary structure of the solid. The

process becomes further complicated when the drying surface causes a shrinkage of the solid. This phenomenon can cause blocking and distortion of the capillary structure and thus interfere with the transfer of internal water to the surface of the material. A striking example of this is the so-called “case hardening” phenomenon, in which the surface of the solid becomes harder than the interior and less permeable to the transmission of interior moisture.

## Behavior of Solids during Drying/Rate of Drying

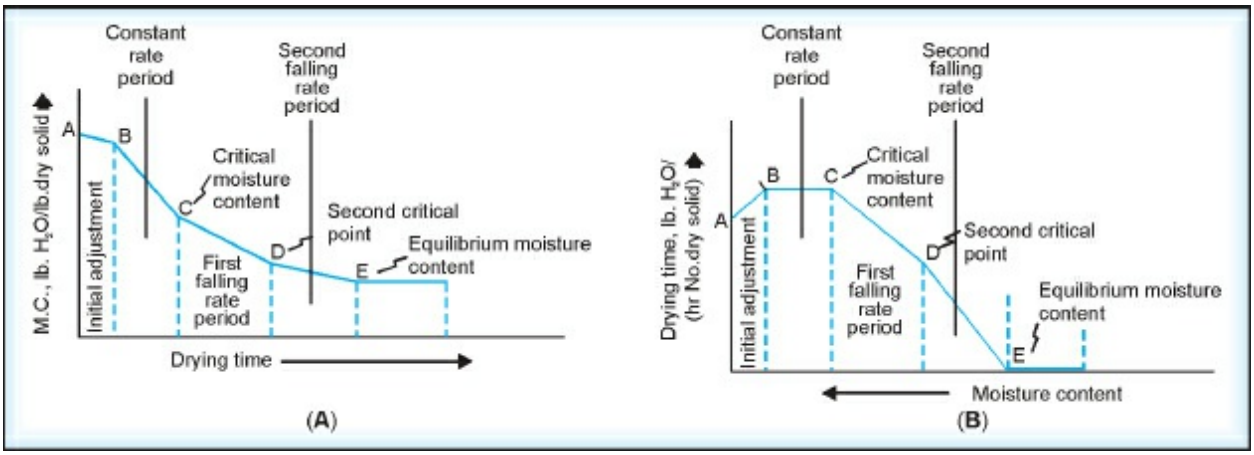
How would one know if 8 or 12 hours are required to dry a batch weight of material in a certain dryer? How can one determine the size of a particular type of dryer required for drying a substance from one moisture level to the desired moisture content?

The rate of drying of a sample can be determined by suspending the wet material on a scale or balance in a drying cabinet and measuring the weight of the sample as it dries as a function of time. In determining an accurate drying rate curve for a material in a particular oven, it is important that the drying conditions approximate the conditions in a full-sized dryer as closely as possible.

The information obtained from the drying rate determination may be plotted as moisture content versus time. The resultant curve is of the type shown in Fig. 4.1A. The changes taking place may be seen more easily if the rate of drying is calculated* and plotted against the moisture content as shown in Fig. 4.1B. Comparison of the rate of drying curve with the drying time curve is clarified when the moisture content is plotted in reverse order, i.e. with the high values to the left.

### *Segment AB-Initial Adjustment*

When a wet solid is first placed in a drying oven, it begins to absorb heat and increase in temperature. At the same time, the moisture begins to evaporate and this tends to cool the drying solid. After a period of initial adjustment, the rates of heating and cooling become equal and the temperature of the drying material stabilizes. As long as the amount of heat transfer by radiation is relatively small, the temperature reached equals the wet-bulb temperature of the drying air. This period of initial adjustment is shown as segment AB in Figs 4.1A and B. If the wet solid is initially at a higher temperature than the wet-bulb temperature, it cools down following segment AB.




**Figs 4.1A and B:** The stages of drying

## *Segment BC-Constant Rate Period*

At point B, the temperature is stabilized and remains constant as long as there is a film of moisture remaining at the surface of the drying solid. Between points B and C, the moisture evaporating from the surface is replaced by water diffusing from the interior of the solid at a rate equal to the rate of evaporation. The rate of drying is constant, and the time BC is the *constant rate period*. The drying rate during this period depends on the air temperature, humidity, and speed, which in turn determine the temperature of the saturated surface.

## *Point C-Critical Moisture Content*

At point C, the surface water is no longer replaced at a rate fast enough to maintain a continuous film. Dry spots begin to appear, and the rate of drying begins to fall off. The moisture content at which this occurs is referred to as the *critical moisture content.*

## *Segment CD-First Falling Rate Period*

At and below the critical moisture content the movement of the moisture from the interior is no longer sufficient to saturate the surface. Between points C and D, the number and area of the dry spots continue to grow, and the rate of drying falls steadily. The time CD is referred to as the *first falling rate period* or the period of *unsaturated surface drying*. Since the surface is no longer saturated, it tends to rise above the wet-bulb temperature.

## *Segment DE-Second Falling Rate Period*

At point D, the film of surface water is completely evaporated, and the rate of drying depends on the rate of diffusion of moisture to the surface of the solid. Point D is referred to as the *second critical point*. Between points D and E the rate of drying falls even more rapidly than the first falling rate, and time DE is called the *second falling rate period*. This section is controlled by vapour diffusion, a factor largely dependent on particle size due to its influence on the dimensions of the pores and channels. During this period, the surface temperature approaches the temperature of the drying air.

## *Point E-Equilibrium Moisture Content*

When the drying rate is equal to zero, starting at point E, the equilibrium moisture period begins, and the solid is in equilibrium with its surroundings, i.e. its temperature and moisture content remain constant. Continued drying after this point is a waste of time and energy.

The condition in which a material is in equilibrium with its surroundings, neither gaining nor losing moisture, may be expressed in terms of its *equilibrium moisture content, equilibrium relative humidity*, or *water activity*. These values may differ greatly for various materials, and in addition to affecting drying, they affect physical and chemical stability, susceptibility to microbial growth, and packaging requirements.

## *Equilibrium Moisture Content*

The moisture content of a material that is in equilibrium with an atmosphere of a given relative humidity is called the *equilibrium moisture content* (EMC) of the material at this humidity. At EMC, there is no driving force for mass transfer since the material exerts a water vapour pressure equal to the vapour pressure of the atmosphere surrounding it. EMC values of various materials may differ greatly under the same conditions, despite the fact that they are in equilibrium with their environment. These differences are due to the manner in which the water is held by the material. The water may be held in fine capillary pores that have no easy access to the surface, dissolved solids may reduce the vapour pressure, or the water may be molecularly bound.

At a given temperature, EMC varies with the partial pressure of the water vapours in the surrounding atmosphere. A hypothetical plot of EMC against

the relative humidity for a hygroscopic solid is shown in Fig. 4.2.

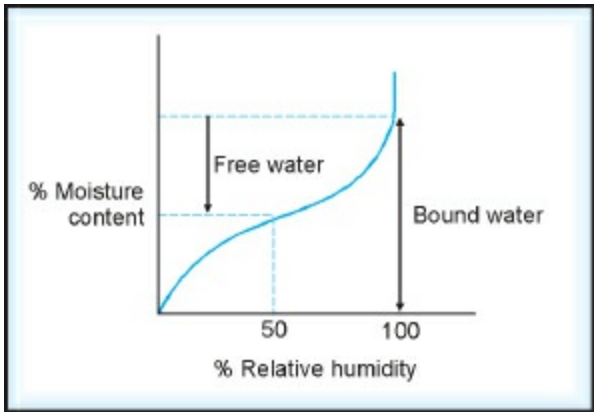



**Fig. 4.2:** Relationship between equilibrium moisture content and relative humidity for a hygroscopic solid

The EMC at 100% relative humidity represents the minimum amount of water associated with the solid that still exerts a vapour pressure equal to a separate water surface. If the humidity is reduced, only a part of the water evaporates before a new equilibrium is established. The water retained at less than 100% relative humidity must, therefore, exert a vapour pressure below that of a dissociated water surface. Such water is called *bound water*. Unlike the equilibrium moisture content, bound water is a function of the solid only and not of the surroundings. Such water is usually held in small pores bound with highly curved menisci, present as a solution, or adsorbed on the surface of the solid. Any moisture present in excess of the EMC is called *free water*.

## *Determination of EMC*

The EMC of a material can be determined by exposing samples in a series of closed chambers, such as desiccators, which are partially filled with solutions that can maintain fixed relative humidities in the enclosed air spaces. The exposure is continued until the material attains a constant weight. This process, which can take more than a month for some materials, can be accelerated by placing a revolving fan in the chamber or by passing air currents with proper humidity and temperature over the material.

Curves of the EMC versus relative humidity have been determined for many pharmaceutical substances. Typical curves are illustrated in Fig. 4.3.

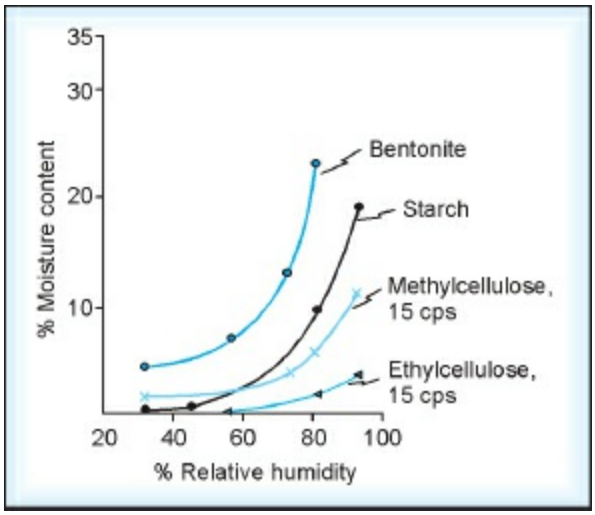



**Fig. 4.3:** Equilibrium moisture content curves for tabletting materials

## *Equilibrium Relative Humidity*

The relative humidity surrounding a material at which the material neither gains nor loses moisture is called the *equilibrium relative humidity* (ERH). At a given temperature, the ERH for a material is determined by its moisture content, just as the EMC is determined by the surrounding relative humidity.

## *Water Activity*

The *water activity* ($a_\omega$) of a material is the ratio of the water vapour pressure exerted by the material to the vapour pressure of pure water at the same temperature. Pure water is assigned an $a_\omega$ of unity, equivalent to an ERH of 100%. Thus, the water activity value for a material is the decimal fraction corresponding to the ERH divided by 100. For example, an ERH of 50% corresponds to an $a_\omega$ of 0.5.

The water activity value has special significance because it is a measure of the relative chemical activity of water in the material. It is related to the thermodynamic chemical potential by the equation:

$$U = U' + RT \ln(a_\omega) \quad \ldots (5)$$

where, U is the chemical potential of water in the material; U' is the chemical potential of pure water; $R$ is the gas law constant; $T$ is the absolute temperature; and ln is the natural logarithm.

The smaller the water activity is, the smaller are the chemical potential of

the water and the driving force for chemical reactions involving water. The most important effects of lowered water activity are increased chemical stability and reduced potential for microorganism growth. Water activity can be reduced by the addition of solutes such as sucrose, glycerin, polyols, and surfactants, as well as by reduced moisture content.

## *Determination of ERH and Water Activity*

ERH and water activity of a material can be measured by allowing the material to equilibrate in a small vapour-tight chamber, such as a glass jar, with hygrometer humidity sensor (mechanical or electric) mounted on the lid. This measurement procedure is much more rapid than the EMC technique, yields practical near-equilibrium values in several hours and true end points in a day or two.

Knowledge of the EMC versus relative humidity curve or ERH versus moisture content for a product allows more intelligent selection of the drying conditions to be used. The relative humidity of the air in the dryer must be lower than the ERH, corresponding to the desired moisture content of the product being dried. In general, the product should be dried to a moisture content corresponding to the EMC at the ambient conditions of processing and storage. If the moisture content differs markedly from this EMC, the product will pick up or lose moisture unless precautions are taken either to maintain the product under controlled humidity conditions or to use packaging materials with low water vapour transmission rates.

# Drying of Solids

## *Loss on Drying*

The moisture in a solid can be expressed on a wet-weight or dry-weight basis. On a wet-weight basis, the water content of a material is calculated as a percentage of the weight of the *wet* solid, whereas on the dry-weight basis, the water is expressed as a percentage of the weight of the dry solid.

In pharmacy, the term *loss on drying*, commonly referred to as LOD, is an expression of moisture content on a wet-weight basis, which is calculated as follows:

$$\% \text{ LOD} = \frac{\text{Weight of water in sample}}{\text{Total weight of wet sample}} \times 100 \quad \ldots (6)$$

The LOD of a wet solid is often determined by the use of a moisture balance, which has a heat source for rapid heating and a scale calibrated in percent LOD. A weighed sample is placed on the balance and allowed to dry until it is at constant weight. The water lost by evaporation is read directly from the percent LOD scale. It is assumed that there are no other volatile materials present.

## *Moisture Content*

Another measurement of the moisture in a wet solid is that calculated on a dry-weight basis. This value is referred to as *moisture content*, or MC:

$$\% MC = \frac{\text{Weight of water in sample}}{\text{Weight of dry sample}} \times 100 \quad \ldots (7)$$

If exactly 5 g of moist solid is brought to a constant dry weight of 3 g:

$$MC = \frac{5-3}{3} \times 100 = 66.7\%$$

$$\text{where, as } LOD = \frac{5-3}{5} \times 100 = 40\%$$

LOD values can vary in any solid-fluid mixture from slightly above g% to slightly below 1gg%, but the MC values can change from slightly above g% and approach infinity. Thus, a small change in LOD value, from 80% to 83%, represents an increase in MC of 88%, or a 22% increase in the amount

of water that must be evaporated per pound of dry product. Thus, percent MC is a far more realistic value than LOD in the determination of dryer load capacity.

## Classification of Solids

Solids may be classified into two major categories on the basis of their drying behavior, namely (1) granular- or crystalline-type solids and (2) amorphous solids.

### *Granular-or Crystalline-Type Solids*

The water in crystalline solids is held in shallow and open surface pores as well as in interstitial spaces between particles that are easily accessible to the surface. Typical pharmaceuticals of the this category are calcium sulfate, zinc oxide, and magnesium oxide.

The moisture in crystalline solids is lost with little hindrance by either gravitational or capillary forces. The constant rate period is the major portion of the drying curve, and this period continues until the material has virtually no free water. The falling rate period is much shorter. Materials in this category are usually inorganic substances and consequently are not affected by heat, unless the temperature is high enough to change any hydrate forms that the chemical may manifest. Equilibrium moisture contents for these materials are close to zero.

### *Amorphous Solids*

In these solids, the moisture is an integral part of the molecular structure as well as being physically entrapped in fine capillaries and small interior pores. Materials with fibrous, amorphous, or gelatinous structures such as starch, casein, yeast, insulin, and aluminum hydroxide fall into this category.

Moisture movement is slow in amorphous solids and the liquid diffuses through structural obstacles caused by the molecular configuration. The drying curves of these amorphous materials have short constant-rate periods, ending at high critical moisture contents. The first falling rate period, the period of water unsaturation on the surface, is relatively short. The second drying rate period is longer, as it depends on the diffusion rate of the water through the solid. The equilibrium moisture content is high, because most of the water remains intimately associated within the molecular interstitial spaces of the substance. The structure and physiologic activity of many of these substances are affected by high temperatures. All of the amorphous solid materials are more difficult to dry than granular or crystalline solids.

The drying of these materials often requires the use of lower temperatures, reduced pressure, and increased air flow.

## DRYERS

Dryers may be classified in several different ways depending on the criteria used. Two useful classifications are based on either the method of heat transfer or the method of solids handling. Classification according to the type of heat transfer, i.e. conduction, convection and radiation, is important in demonstrating gross differences in dryer design, operation, and energy requirements. Classification by the method of solids handling is more suitable when special attention must be given to the nature of the material to be dried.

When dryers are classified according to their method of solids handling, the major criterion is the presence or absence of agitation of the material to be dried. A dryer that produces excessive agitation is contraindicated when the dried material is friable and subject to attrition. On the other hand, if the dried product is intended to be pulverized, then the drying time can be reduced, and the process made more efficient, by the use of a dryer that produces intense agitation during the drying cycle.

Classification based on the method of solids handling is shown schematically in Fig. 4.4. Dryers in this classification scheme are divided into the following types:

1. *Static-bed dryers*—systems in which there is no relative movement among the solid particles being dried, although there may be bulk motion of the entire drying mass. Only a fraction of the total number of particles is directly exposed to heat sources. The exposed surface can be increased by decreasing the thickness of the bed and allowing drying air to flow through it.
2. *Moving-bed dryers*—systems in which the drying particles are partially separated so that they flow over each other. Motion may be induced by either gravity or mechanical agitation. The resultant separation of the particles and continuous exposure of new surfaces allow more rapid heat and mass transfer than can occur in static beds.
3. *Fluidized-bed dryers*—systems in which the solid particles are partially suspended in an upward-moving gas stream. The particles are lifted and then they fall back in a random manner so that the resultant mixture of solid and gas acts like a boiling liquid. The gas-solid contact is excellent and results in better heat and mass transfer than in static and moving

beds.

4. *Pneumatic dryers*—systems in which the drying particles are entrained and conveyed in a high-velocity gas stream. Pneumatic systems further improve on fluidized beds, because there is no channeling or short-circuiting of the gas flow path through a bed of particles. Each particle is completely surrounded by an envelope of drying gas. The resultant heat and mass transfer are extremely rapid, and thus, drying times are short.

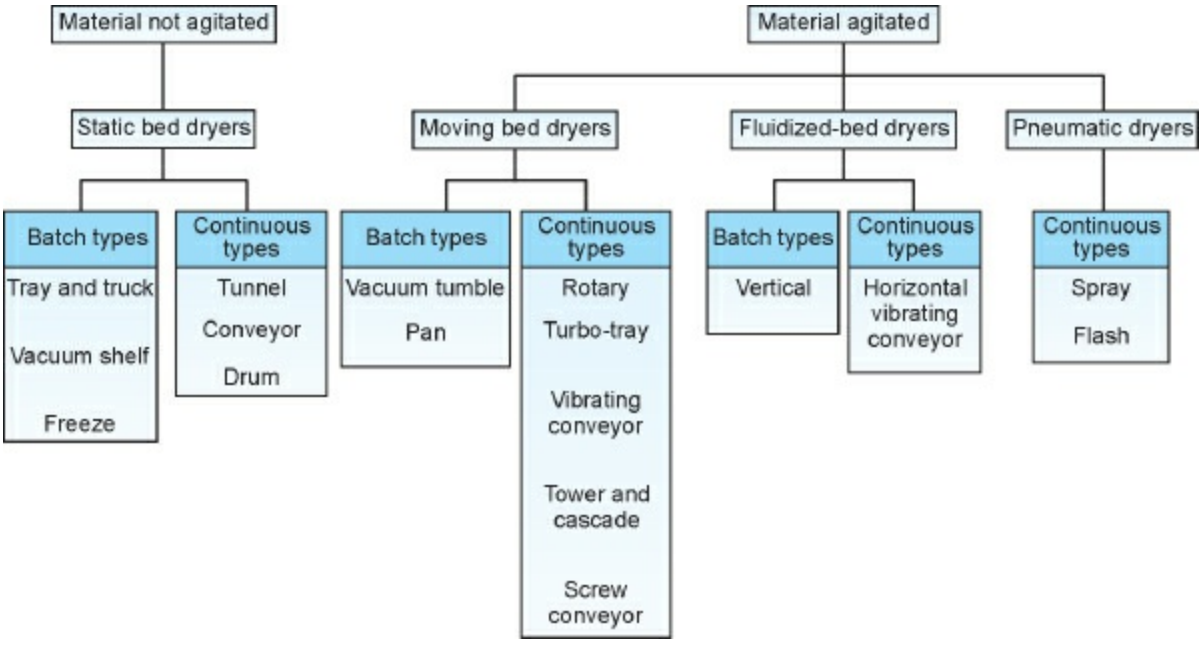



**Fig. 4.4:** Classification tree of dryers, based on the methods of solid handling

Because of the great variety of available drying equipment, it is impossible to describe all types of dryers. Attention is devoted to those that find ready application to the production of pharmaceuticals. These dryers are grouped according to their method of solids handling.

## Static-Bed Systems

### *Tray and Truck Dryers*

The dryers most commonly used in pharmaceutical plant operations are tray and truck dryers. An example of a tray dryer is illustrated in Fig. 4.5. Tray dryers are sometimes called shelf, cabinet, or compartment dryers. This dryer consists of a cabinet in which the material to be dried is spread on tiers of trays. The number of trays varies with the size of the dryer. Dryers of laboratory size may contain as few as three trays, whereas larger dryers often hold as many as twenty trays. Despite its versatility and the low relative capital investment required for tray-drying, it provides a low rate of drying and the loading and unloading of trays is a labour-intensive process. Tray-drying has become less popular in comparison to other more efficient, reproducible, and well-defined drying procedures such as fluid bed and vacuum-tumble drying.

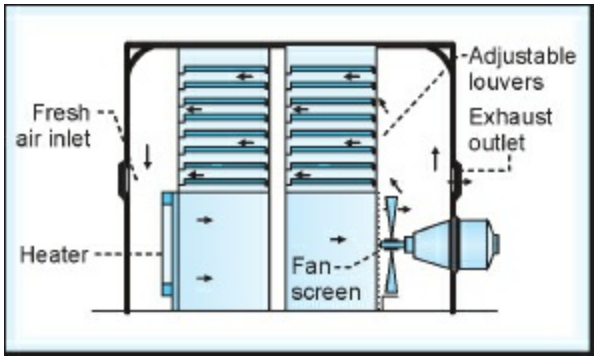



**Fig. 4.5:** Tray dryer

A truck dryer is one in which the trays are loaded on trucks (racks equipped with wheels), which can be rolled into and out of the drying cabinet. In plant operations, the truck dryer is preferred over the tray dryer because it offers greater convenience in loading and unloading. The trucks usually contain one or two tiers of trays, with about 18 or more trays per tier. Each tray is square or rectangular and about 4 to 8 square feet in area. Trays are usually loaded from 0.5 to 4.0 inches deep with at least 1.5 inches clearance between the surface and the bottom of the tray above.

Drying in tray or truck dryers is a batch procedure, as opposed to continuous drying as performed in a moving belt dryer. Batch drying is used

extensively in the manufacture of pharmaceuticals for several reasons (1) each batch of material can be handled as a separate entity, (2) the batch sizes of the pharmaceutical industry are relatively small (500 or less pounds per batch) compared with the chemical industry (2,000 or more pounds per hour) and (3) the same equipment is readily adjusted for use in drying a wide variety of materials.

Tray dryers may be classified as direct or indirect. Most tray dryers used in the pharmaceutical industry are of the direct type, in which heating is accomplished by the forced circulation of large volumes of heated air. Indirect tray dryers utilize heated shelves or radiant heat sources inside the drying chamber to evaporate the moisture, which is then removed by either a vacuum pump or a small amount of circulated gas. Further discussion in this section is confined to the direct (convection-type) dryer. Vacuum dryers are described separately later in this chapter.

The trays used have solid, perforated, or wire-mesh bottoms. The circulation of drying air in trays with a solid base is limited to the top and bottom of the pan, whereas in trays with a perforated screen, the circulation can be controlled to pass through each tray and the solids on it. The screen trays used in most pharmaceutical drying operations are lined with paper, and the air thus circulates across rather than through the drying material. The paper is used as a disposable tray liner to reduce cleaning time and prevent product contamination.

To achieve uniform drying, there must be a constant temperature and uniform airflow over the material being dried. This is accomplished in modern dryers by the use of a well-insulated cabinet with strategically-placed fans and heating coils as integral parts of the unit. The air circulates through the dryer at 200 to 2,000 feet per minute. The use of adjustable louvers helps to eliminate nonuniform airflow and stagnant pockets.

The preferred energy sources for heating the drying air used on pharmaceutical products are steam or electricity. Units fired with coal, oil, and gas produce higher temperatures at lower cost, but are avoided because of possible product contamination with fuel combustion products, and explosion hazards when flammable solvents are being evaporated. Steam is preferred over electricity, because steam energy is usually cheaper. If steam is not readily available, and drying loads are small, electric heat is used.

## *Vacuum Shelf*

During vacuum drying, the total pressure surrounding the pharmaceutical material is reduced to levels below the saturation pressure of the solvent at the interface between the wet and dry layers causing the generation of vapour. Vacuum tray dryers differing only in size from the laboratory vacuum oven, offer an alternative method for drying small quantities of material. Because of massive construction and higher cost, vacuum tray dryers are used only when a definite advantage over the hot air oven is secured, such as low-temperature drying of thermolabile materials (penicillin preparation) and for applications requiring solvent recovery or extremely low residual solvent levels (ethanol extractives).

With suitable vacuum levels, drying can be cost-effective at relatively low product temperatures. The exclusion of oxygen may also be advantageous or necessary in some operations. Heat is usually supplied by passing steam or hot water through hollow shelves. Drying temperatures can be carefully controlled and, for the major part of the drying cycle, the material remains at the boiling point of the wetting liquid under the operating vacuum. Radiation from the shelf above may cause a significant increase in temperature at the surface of the material if high drying temperatures are used. Drying times are long, usually about 12 to 48 hours.

## *Freeze Dryers*

Many products of pharmaceutical interest lose their viability in the liquid state and readily deteriorate if dried in air at normal atmospheric pressures. These pharmaceutical materials may be heat-sensitive or may react readily with oxygen, so that in order to be stabilized, they must be dehydrated to a solid state. The material to be dried is first frozen and then subjected under a high vacuum to heat (supplied by conduction or radiation, or by both) so that the frozen liquid sublimes leaving only the solid, dried components of the original liquid. Such materials as blood serum, plasma, antibiotics, hormones, bacterial cultures, vaccines, and many foodstuffs are dehydrated by *freeze drying*, also referred to as *lyophilization*, *gelsiccation* or *drying by sublimation*. The dried product can be readily redissolved or resuspended by the addition of water prior to use, a procedure referred to as *reconstitution*.

*Principle:* Freeze drying depends on the phenomenon of *sublimation*,

whereby water passes directly from the solid state (ice) to the vapour state without passing through the liquid state. As shown in the schematic pressure-temperature diagram for water (Fig. 4.6), sublimation can take place at pressures and temperatures below the triple point, 4.579 mm Hg absolute (4,579 µm) and 0.0099. C. The water in pharmaceutical products intended for freeze drying contains dissolved solids, resulting in a different pressure-temperature relationship for each solute. In such cases, the pressure and temperature at which the frozen solid vaporizes without conversion to a liquid is referred to as the *eutectic point*. Freeze drying is carried out at temperatures and pressures well below this point to prevent the frozen water from melting, which would result in frothing, as the liquid and frozen solid vaporize simultaneously. In actual practice, freeze drying of pharmaceuticals is carried out at temperatures of 100 to 2,000 and at pressures of 2,000 to 100 µm.

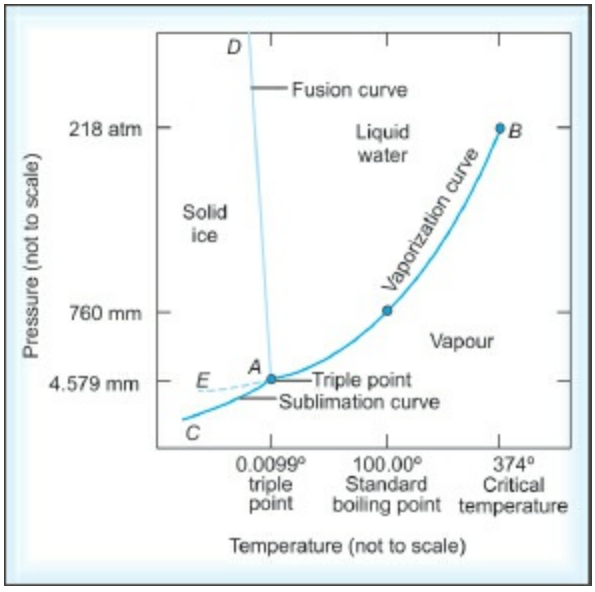



**Fig. 4.6:** Schematic pressure-temperature diagram for water, showing the conditions for various phases

*Requirements:* Freeze drying must meet the three basic requirements for all types of drying, despite this unusual approach to drying.

1. The vapour pressure of the water on the surface of the material being dried must be higher than the partial pressure of the enveloping atmosphere, i.e. there must be a positive vapour pressure driving force.
2. The latent heat of vaporization must be introduced to the drying solid at

such a rate as to maintain desirable temperature levels at both the surface and interior.

3. Provision must be made for the removal of evaporated moisture.

*Components:* Freeze dryers are composed of four basic components (Fig. 4.7) (1) a chamber for vacuum drying, (2) a vacuum source, (3) a heat source and (4) a vapor-removal system. The chamber for vacuum drying is generally designed for batch operation and thus can be compared to the vacuum shelf dryer. Special inlet and outlet mechanisms have been designed in some drying chambers to achieve a continuous drying operation. Vacuum is achieved by pumps, steam ejectors, or a combination of the two. Heat is provided by conduction or radiation, or both. Three different methods for the removal of water vapour are employed: condensers, desiccants, and pumps. The water vapour is removed from the drying chamber and condensed in the form of a thin layer of ice on a heat-transfer surface in the condenser. The ice is removed intermittently by melting it with a heated fluid that is circulated through the condenser, or in the case of a continuous operation, by means of scraper blades. Liquid or solid desiccants are often employed in the initial vapour removal to enhance the efficiency of the pumps removing the water vapor. In general, scraper blades and desiccants are used for freeze drying large-volume biologicals (e.g. serum, penicillin), and usually are not used for preparing pharmaceutical dosage forms.

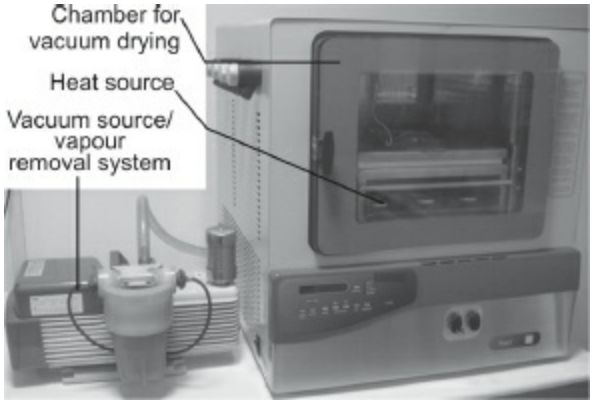



**Fig. 4.7:** Schematic representation of a Freeze dryer

## *Stages of Freeze Drying*

1. *Pre-freezing:* Material is frozen by keeping the material below or at −20°C. Pre-freezing the material before application of vacuum avoids

foaming.

2. *Vacuum:* Rotary pumps on small scale, and ejector pumps on large-scale, are used to reduce the pressure sufficiently.
3. *Primary drying:* During primary drying the latent heat of sublimation must be provided and the vapour removed. Primary drying stage by sublimation can remove the unbound water.
4. *Secondary drying:* It is used to remove bound water or traces of water left after primary drying. The temperature is raised (up to 50°C) or desiccant is used to carry secondary drying.

## *Tunnel and Conveyor Dryers*

Tunnel dryers (Fig. 4.8) are adaptations of the truck dryer for continuous drying. The trucks are moved progressively through the drying tunnel by a moving chain. These trucks are loaded on one side of the dryer, allowed to reside in the heating chamber for a time sufficiently long to effect the desired drying, and then discharged at the exit. The operation may be more accurately described as *semicontinuous*, because each truck requires individual loading and unloading before and after the drying cycle. Heat is usually supplied by direct convection, but radiant energy may also be used.

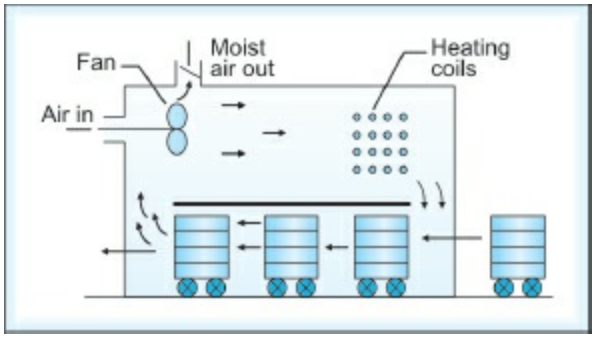



**Fig. 4.8:** Tunnel dryer

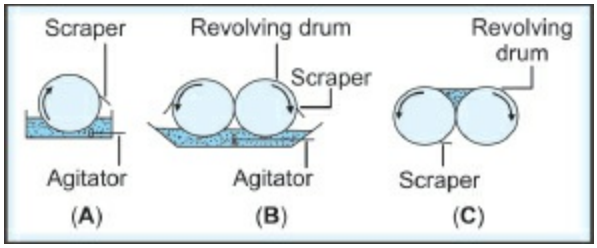
Conveyor dryers are an improvement over tunnel dryers because they are truly *continuous*. The individual trucks of the tunnel are replaced with an endless belt or screen that carries the wet material through the drying tunnel. Conveyor dryers provide for uninterrupted loading and unloading, and are thus more suitable for handling large volumes of materials.

The drying curve characteristic of the material in batch drying is altered

considerably when continuous-type dryers are used. As the mass is moved along its drying path in a continuous operation, this mass is subjected to drying air, the temperature and humidity of which are continually changing. As a consequence, the "constant rate" period is not constant, but decreases as the air temperature decreases, although the surface temperature of the wetted mass remains constant. Thus, drying rate curves for batch drying are not equally applicable to continuous drying procedures.

## *Drum Dryers*

The drum dryer consists of one or two slowly rotating, steam-heated cylinders, coated with solution or slurry by means of a dip feed, illustrated in Figs 4.9A to C. The lower portion of the drum is immersed in an agitated trough of feed material or, in the case of some doubledrum dryers, by feeding the liquor into the gap between the cylinders as shown in Fig. 4.9. Spray and splash feeds are also used. When dip feeding is employed, the hot drum must not boil the liquid in the trough. Drying takes place by simple evaporation rather than by boiling. The dried material is scraped from the drum at a suitable point by a scraper.




**Figs 4.9A to C:** Drum dryers

With the double-drum dryer, the drying rate is influenced by drum speed, drum temperature and feed temperature, whereas the gap between the cylinders determines the film thickness.

Drum dryers, like spray dryers, are relatively expensive in small sizes and their use in the pharmaceutical industry is largely confined to drying thermolabile materials where a short contact time is advantageous. The heat treatment to which the solid is subjected is greater than in spray drying and the physical form of the product is often less attractive. Drums are normally fabricated from stainless or chrome-plated steel to reduce contamination. The drum dryers can handle a variety of materials either as solution or suspension

and is used in the drying of kaolin or zinc oxide suspension.

## Moving-bed Systems

### *Vacuum Tumble Dryer*

The limitations of static dryers, particularly with respect to the long drying times, have, where possible, promoted the design and application of tumble dryers. The most common shape of the tumble dryer is the double cone vacuum dryer as shown in Fig. 4.10. Operating under vacuum, this dryer provides controlled low-temperature drying, the possibility of solvent recovery, and increased rates of drying. With correct operation, a uniform powder should be obtained, as distinct from the cakes produced when static beds are dried. A typical rotation speed of the vacuum dryer is 6–8 rpm. Vacuum is supplied by conventional pumps, blowers, or steam jets. Heat is supplied to the tumbling charge by contact with the heated shell and by heat transfer through the vapor. Heating fluid circulates through a jacket and enters and exits through dynamic seals along the axis of rotation. Indirect methods rely on contact between the wet material and the jacketed walls of the dryer to supply energy and the drying rate can be heat transfer-limited. When applied to drying tablet granules, periods of 2 to 4 h replace the 18 to 24 h obtained with hot air ovens. The ratio of jacket area to working volume tends to decrease with increasing size, so larger models often require additional internal plates or pipe coils to increase available area for heat transfer. Some materials, such as waxy solids, cannot be dried by this method because the tumbling action causes the material to aggregate into balls. Vacuum drying can be readily incorporated into high-shear granulation designs to permit multiple processing steps to be completed in a single piece of equipment.

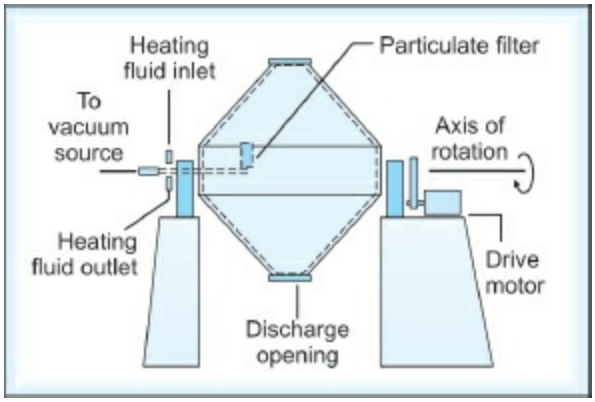


**Fig. 4.10:** Rotating double-cone vacuum dryer

## *Pan Dryer*

Pan dryers are moving-bed dryers of the indirect type that may operate under atmospheric pressure or vacuum, and are generally used to dry small batches of pastes or slurries. The dryer consists of a shallow, circular jacketed pan having a diameter of 3 to 6 feet and depth of 1 to 2 feet, with a flat bottom and vertical sides. Heat is supplied by steam or hot water. There is a set of rotating plows in the pan that revolve slowly, scraping the moisture-laden mass from the walls and exposing new surfaces to contact with the heated sides and bottom. Atmospheric pan drying allows the moisture to escape, whereas in vacuum dryers in which the pan is completely enclosed, solvents are recoverable if the evacuated vapours are passed through a condenser. The dried material is discharged through a door on the bottom of the pan.

## *Rotary Dryer*

The rotary dryer is the modified form of the tunnel dryer in which the particles are passed through a rotating cylinder, counter-current to a stream of heated air (Fig. 4.11). Due to the rotation of the cylinder, the material is turned over and drying takes place from individual particles and not from a static bed. The cylindrical shell is mounted with a slight slope so as to discharge the material and make the operation continuous. Baffles or flights in the shell may increase the rate of drying.

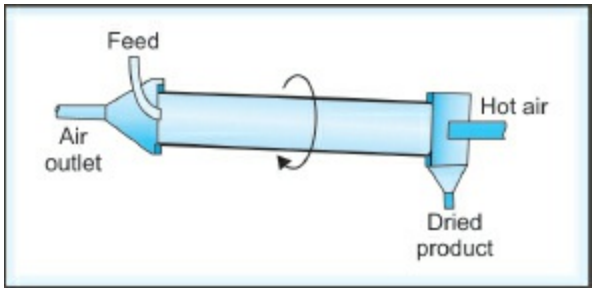



**Fig. 4.11:** Rotary dryer

## *Turbo-tray Dryers*

The turbo-tray dryer, illustrated in Fig. 4.12, is a continuous shelf, moving-bed dryer. It consists of a series of rotating annular trays arranged in a

vertical stack, all of which rotate slowly at 0.1 to 1.0 rpm. Heated air is circulated over the trays by turbo-type fans mounted in the center of the stack. Wet mass fed through the roof of the dryer is leveled by a stationary wiper. After about seven-eighths of a revolution, the material being dried is pushed through radial slots onto the tray below, where it is again spread and leveled. The transfer of mass from one shelf to the next is complete after one revolution. The same procedure continues throughout the height of the dryer until the dried material is discharged at the bottom. Because the turbo-tray dryer continuously exposes new surfaces to the air, drying rates are considerably faster than for tunnel dryers.

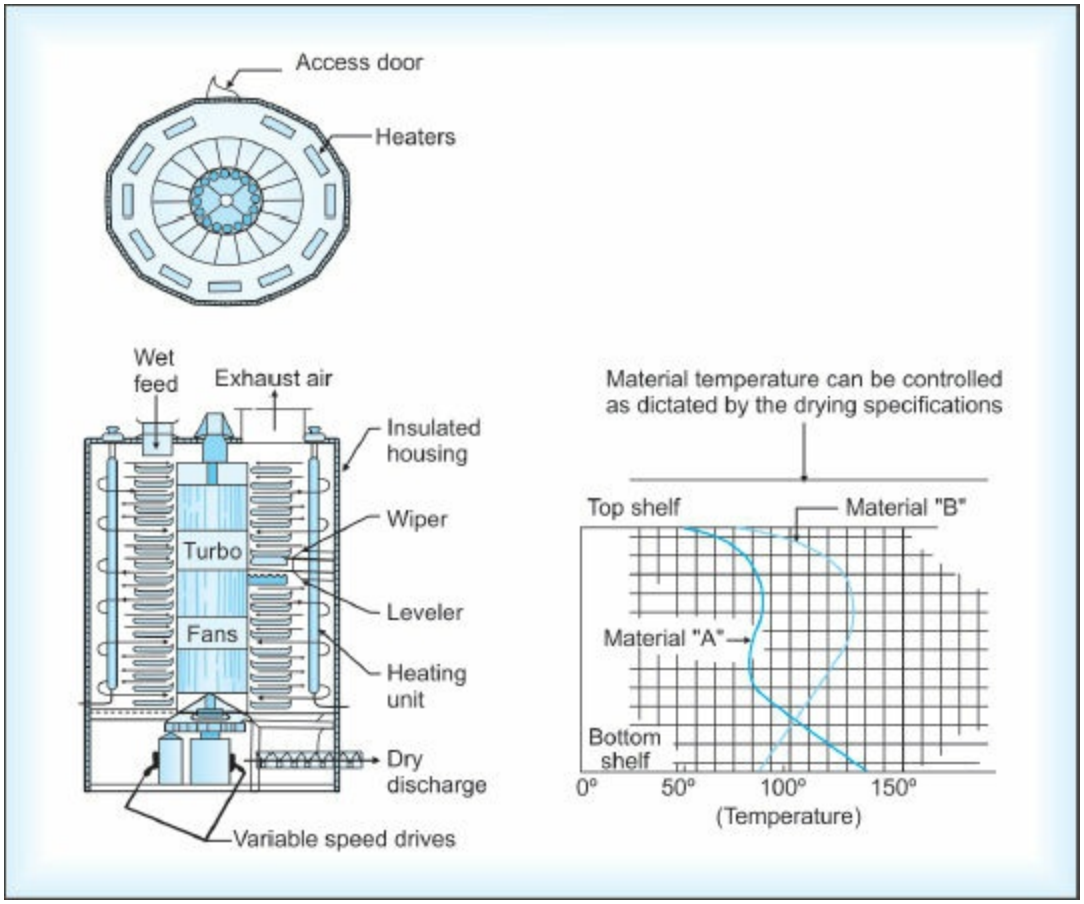



**Fig. 4.12:** Turbo-tray dryer

## Fluidized-bed Dryer

If a gas is allowed to flow upward through a bed of particulate solids at a velocity greater than the settling velocity of the particles and less than the velocity for pneumatic conveying, the solids are buoyed up and become partially suspended in the gas stream. The resultant mixture of solids and gas behaves like a liquid, and the solids are said to *be fluidized*. The solid particles are continually caught up in eddies and fall back in a random boiling motion. The gas-solids mixture has a zero angle of repose, seeks its own level, and assumes the shape of the vessel that contains it. The fluidization technique is efficient for the drying of granular solids, because each particle is completely surrounded by the drying gas. In addition, the intense mixing between the solids and gas results in uniform conditions of temperature, composition, and particle size distribution throughout the bed. Fluidized-bed drying has been reported to offer distinct advantages over conventional tray drying for tablet granulations. In general, tablet granulations have the proper particle sizes for good fluidization. The only requirements are that the granules are not so wet that they stick together on drying, and that the dried product is not so friable as to produce excessive amounts of fine particles through attrition. It was found that the fluidized-bed dryer showed a two-fold to sixfold advantage in thermal efficiency over a tray dryer. The fluidized-bed dryer was also shown to be significantly faster in both drying and handling time than the tray dryer. To avoid electrostatic charge build-up and the resultant explosion hazards, fluid beds are provided with static charge grounding devices.

### *Types of Fluidized-bed Dryers*

The fluidized-bed dryers available for use in the pharmaceutical industry are of two types, *vertical* and *horizontal*. The design features of a vertical fluidized-bed dryer are shown in Fig. 4.13.

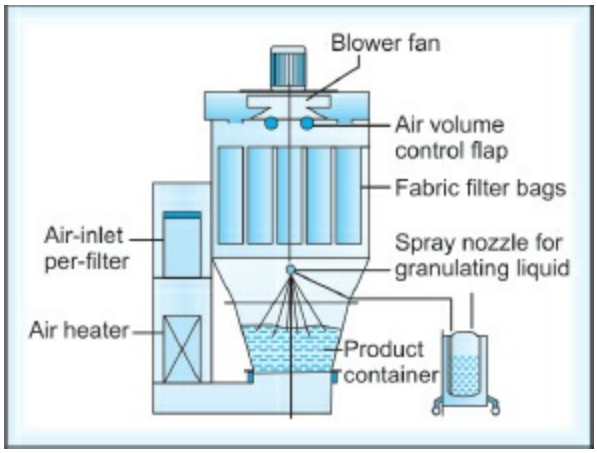



**Fig. 4.13:** Fluidized-bed granulator-dryer

## *Vertical Fluidized-bed Dryers*

The fluidizing air stream is induced by a fan mounted in the upper part of the apparatus. The air is heated to the required temperature in an air heater and flows upward through the wet material, which is contained in a drying chamber fitted with a wire mesh support at the bottom. The air flow rate is adjusted by means of a damper, and a bag collector filter is provided at the top of the drying chamber to prevent carryover of fine particles. The unit described is a batch-type dryer, and the drying chamber is removed from the unit to permit charging and dumping. Dryer capacities range from 5 kg to 200 kg and the average drying time is about 20 to 40 min. Because of the short drying time and excellent mixing action of the dryer, no hot spots are produced, and higher drying temperatures can be employed than are used in conventional tray and truck dryers.

The unit shown in Fig. 4.13 is designed for the direct preparation of tablet granulations as well as for the drying of conventionally produced wet granulations. When the unit is used as a granulator, the dry ingredients are placed in the chamber and fluidized while the granulating liquid is sprayed into the bed, causing the particles to agglomerate into granules. At the end of the granulation cycle, the granules are dried by heating the fluidizing air.

## *Horizontal Vibrating Conveyor*

A continuous dryer is more suitable than a batch type for the drying of larger volumes of materials. A fluidized-bed dryer of this type, which is suitable for pharmaceutical use, is a horizontal vibrating conveyor dryer, shown in Fig.

4.14. The heated air is introduced into a chamber below the vibrating conveying deck and passes up through the perforated or louvered conveying surface, through the fluidized-bed of solids, and into an exhaust hood. A fluidized-bed of uniform density and thickness is maintained in any given drying zone by the vibration. Residence time in any drying zone is controlled by the length of the zone, frequency and amplitude of the vibration, and use of dams. The dryer can be divided into several different zones with independent control of airflow and temperature, so that drying can take place at the maximum desirable rate in each stage without sacrificing efficiency or damaging heat-sensitive materials. Dryers vary in width from 12 to 57 inches and in length from 10 to 50 feet, with bed depths of 3 inches. Dryer capacity is limited only by the retention time produced by conveying speeds, which range from 5 to 25 feet per minute. In pharmaceutical operations, capacities range as high as 1 to 2 tons per hour.

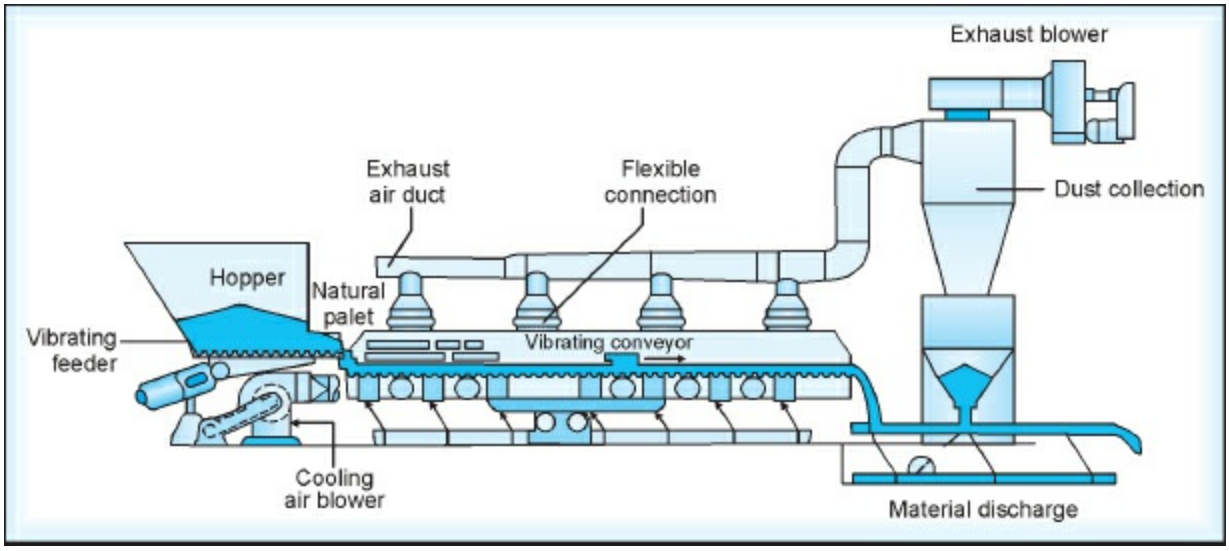



**Fig. 4.14:** Horizontal vibrating-conveyor fluidized-bed dryer

## Pneumatic Dryers

### *Spray Dryers*

Spray dryers differ from most other dryers in that they can handle only fluid materials such as solutions, slurries, and thin pastes. The fluid is dispersed as fine droplets into a moving stream of hot gas, where they evaporate rapidly before reaching the wall of the drying chamber. The product dries into a fine powder, which is carried by the gas current and gravity flow into a collection system.

*Principle:* When the liquid droplets come into contact with the hot gas, they quickly reach a temperature slightly above the wet-bulb temperature of the gas. The surface liquid is quickly evaporated, and a tough shell of solids may form in its place. As drying proceeds, the liquid in the interior of the droplet must diffuse through this shell. The diffusion of the liquid occurs at a much slower rate than does the transfer of heat through the shell to the interior of the droplet. The resultant build-up of heat causes the liquid below the shell to evaporate at a far greater rate than it can diffuse to the surface. The internal pressure causes the droplet to swell, and the shell becomes thinner, allowing faster diffusion. If the shell is nonelastic or impermeable, it ruptures, producing either fragments or budlike forms on the original sphere. Thus, spray-dried material consists of intact spheres, spheres with buds, ruptured hollow spheres, or sphere fragments.

There are many types of spray dryers, each designed to suit the material being dried and the desired product characteristics. One example is shown in Fig. 4.15. All spray dryers can be considered to be made up of the following components: feed delivery system, atomizer, heated air supply, drying chamber, solids-gas separator, and product collection system.

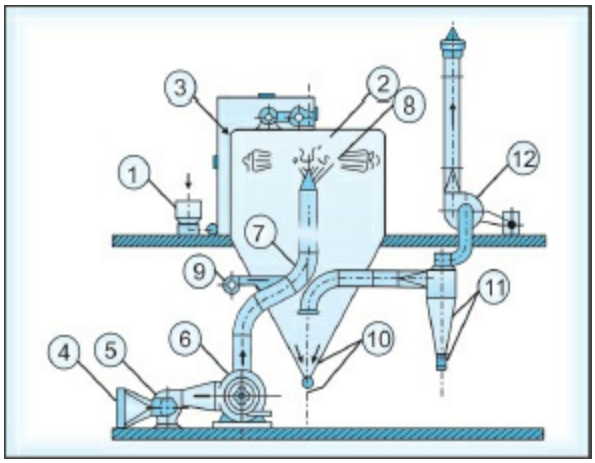



**Fig. 4.15:** Spray dryer: 1. feed tank; 2. centrifugal atomizer; 3. drying chamber; 4. inlet air filter; 5. air supply fan; 6. air heater; 7. triple inlet duct; 8. adjustable air disperser; 9. cooling air fan; 10. chamber product collection; 11. cyclone product collection; 12. exhaust fan

*Working:* The feed is delivered to the atomizer by gravity flow or by the use of a suitable pump. The rate of feed is adjusted so that each droplet of sprayed liquid is completely dried before it comes into contact with the walls of the drying chamber, and yet the resultant dried powder is not overheated in the drying process. The proper feed rate is determined by observation of the outlet air temperature and visual inspection of the walls of the drying chamber. If the inlet air temperature is kept constant, a drop in the liquid feed rate is reflected by a rise in the outlet temperature. Excessive feed rates produce a lowering of the outlet temperature, and ultimately, a build-up of material on the walls of the chamber.

*Construction.* Spray dryer atomizers are of three basic types: *pneumatic atomizers*, *pressure nozzles*, and *spinning disc atomizers*. In the pneumatic atomizer (also called two-fluid or gas-atomizing nozzle), the liquid feed is broken up into droplets by a high-velocity jet of air or other gas. Pneumatic atomizers are used to produce small particles and for spraying more viscous liquids than can be handled by pressure nozzles. The pneumatic atomizer, however, requires more power than other type of atomizers to achieve the same fine spray. The liquid feed is delivered by pressure nozzles under high pressure (up to 7,000 pounds per square inch) and is broken up on coming into contact with the air or by impact on another jet or fixed plate. In spinning disc atomizers, the liquid is fed to the center of a rapidly rotating disc (3,000

to 50,000 rpm), where centrifugal force breaks up the fluid into droplets. Spinning disc atomizers find wide utility in the spray drying of pharmaceuticals because of their ability to handle all types of liquid feeds, including high-viscosity liquids and slurries of particles that would clog other atomizers.

Hot and dry air is supplied by the blowing of air over a heat exchanger. The heat may be supplied by steam, or by direct- or indirect-fire heaters. The usual heat source in laboratory units is electricity or gas. *Steam* or *indirect-fire heaters* are preferred in the spray drying of pharmaceuticals because their use avoids product contamination with combustion products.

Separation of the solid product from the effluent gas is usually accomplished by means of a cyclone separator. It is referred to as the *primary collector*. The dried product collected at this point is referred to as *cyclone product*. Any dust remaining in the air may be removed by means of a filter bag collector or a wet scrubber to avoid air pollution. Product that reaches the walls of the drying chamber, referred to as *chamber product*, is removed at the bottom of the chamber. This chamber product is usually coarser in size and subjected to heat longer (because of increased retention time) than the cyclone product. The final dried product is *usually* a mixture of both the chamber and cyclone products.

*Spray drying of pharmaceuticals:* Spray drying finds great utility in the pharmaceutical industry because of the rapidity of drying and the unique form of the final product. There are three major uses for the spray drying processes (1) drying heat-sensitive materials, (2) changing the physical form of materials for use in tablet and capsule manufacture and (3) encapsulating solid and liquid particles.

Spray drying can be used to dry materials that are sensitive to heat and/or oxidation without degrading them, even when high temperature air is employed. The liquid feed is dispersed into droplets, which are dried in seconds because of their high surface area and intimate contact with the drying gas. The product is kept cool by vaporization of the enveloping liquid, and the dried product is kept from overheating by rapid removal from the drying zone.

Spray drying is valuable in the modification of materials for use in tablet and capsule formulations, because the drying process changes the shape, size, and bulk density of the dried product. The spherical particles produced

usually flow better than the same product dried by conventional procedures, because the particles are more uniform in size and shape, with fewer sharp edges. The spherical shape has the least possible surface area, thus minimizing air entrapment between the particles. The improvement in flow and reduction of air entrapment make the spray-dried material suitable for use in the manufacture of tablets and capsules. The spherical particle shape is obtained by spray drying either a solution of the material or a slurry of particles in a saturated solution of the same material. In the latter case, the configuration of the suspended particle is rounded out by deposition of the material in solution. An example of a spray dried material that is commonly used as a tablet excipient is spray dried lactose.

Spray drying has proved extremely useful in the coating and encapsulation of both solids and liquids. Solid particles are coated by spray drying a suspension of the material in a solution of the coating agent. As the solvent is evaporated, the coating material envelops the suspended particle. The coating provides such valuable characteristics as taste and odour masking, improvement in stability, enteric coating, and sustained release. Oily liquids may be encapsulated by emulsification in water with the aid of a gum such as acacia, or starch, and subsequent spray drying. As the water evaporates, the oil is entrapped in a shell of the gum. This process is used for the preparation of “dry” flavour oils.

## *Spray Congealing*

An alternative to spray for the encapsulation of solid particles is *spray chilling* or *spray congealing*. This process consists of suspending the particles in a molten coating material and pumping the resultant slurry into a spray dryer in which cold air is circulated. The slurry droplets congeal on coming into contact with the air and are collected in the same manner as the spray dried product. The coating agents normally employed are low melting materials such as waxes. The congealing process requires a much higher ratio of coating agent to active material than does spray drying, because only the molten coating agent constitutes the liquid phase. Spray congealed coatings are used mainly for taste masking and sustained-release formulations.

## *Flash Dryers*

In flash drying, the moistened solid mass is suspended in a finely divided state in a high-velocity (3,000 to 6,000 feet per minute), high-temperature 

(300°F to 1,300°F) air stream. The dispersed particles may be carried in the air stream to an impact mill, or the pneumatic flow itself may reduce the particle size of friable material. The resultant attrition exposes new surfaces for more rapid drying. The dried, fine particulate matter passes through a duct with an opening small enough to maintain desired air-carrying velocities. The dried solid is collected by a cyclone separator, which may be followed by a bag collector or wet scrubber. Thus, the flash dryer is an example of a parallel (cocurrent) airflow drying system.

The drying process is referred to as *flash drying*, because the drying time is extremely short. The drying air temperature can drop from 1,300°F to 600°F in two seconds and to 350°F in only four seconds. The temperature of the drying solid can be kept at 100°F or less.

The drying cycle in many flash dryers occurs in one unit (single-stage conveyor). Multi-stage units are employed for drying solids that have high moisture content and contain large amounts of bound water. Figure 4.16 illustrates a two-stage unit wherein partial drying occurs in the first unit, and drying is completed in the second pneumatic conveyor dryer. In other units, some of the dried material is mixed with the wet incoming solid to make it easier to mill.

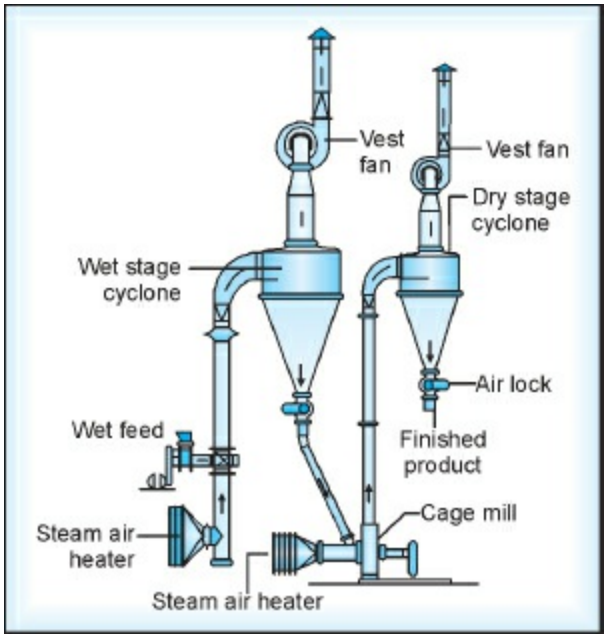



**Fig. 4.16:** Two-stage flash dryer

## Specialized Drying

### *Microwave Drying*

The application of microwave energy to the drying of solids represents a radical departure from conventional means of drying. Instead of applying heat externally to a material, energy in the form of microwaves is converted into internal heat by interaction with the material itself. This permits extremely rapid heat transfer throughout the material, which in turn can lead to rapid drying.

The heating effect is produced by the interaction of a rapidly oscillating electric field (915 or 2450 megahertz) with the polarized molecules and ions in the material. The field imposes order on otherwise randomly oriented molecules. As the field reverses polarity, it relaxes and allows the molecules to return to their random orientation, giving up the stored potential energy as random kinetic energy, or heat. The interaction of the alternating field with ions causes billiard-ball-like collisions with un-ionized molecules, and the impact energy is converted into heat.

A given material's molecular and ionic makeup intimately affects its ability to be dried, as is shown in the power conversion equation for microwave heating:

$$P = kf\,E_2\,\acute{\varepsilon}''\,\tan\delta = kf\,E^2\,\acute{\varepsilon}''\ \ldots (8)$$

| | | |
|---|---|---|
| where, $P$ | = | the power developed (watts/unit volume) |
| $k$ | = | a constant |
| $f$ | = | the frequency |
| $E$ | = | the electric field strength (volts/unit distance) |
| $\acute{\varepsilon}$ | = | the relative dielectric constant of the material being heated |
| $\tan\delta$ | = | the loss tangent, or dissipation factor of the material |
| $\acute{\varepsilon}$ | = | the loss factor of the material, equal to the product $\acute{\varepsilon}\,\delta$ |

In microwave drying, the mass transfer is primarily the result of a pressure gradient due to rapid vapour generation inside the material, that is, most of the internal moisture is vapourized before leaving the sample. Thus, the moisture is mobilized as a vapour rather than a liquid, and its movement

to the surface can be extremely rapid because it does not depend on mass concentration gradients or on slow liquid diffusion rates.

Industrial microwave dryers are usually of the static bed continuous type. Materials to be dried are placed on conveyor belts and conveyed through the microwave applicator. Generally, a stream of hot air is used simultaneously with the microwaves to sweep away the moisture evolving from the surface of the material being dried. Often, the microwave treatment is used in the last stages of hot air drying (the second falling rate period of Fig. 4.1) to remove the last remaining portion of the solvent, reducing total drying time by 50% or more.

Microwave drying can be used for the drying of pharmaceutical materials at low ambient temperatures, avoiding high surface temperatures, case hardening, and solute migration. Microwave vacuum drying at low pressure (1 to 20 mm Hg) and moderate temperature (30 to 40°C) can be used for drying thermolabile materials such as vitamins, enzymes, proteins, and flavours.

The rising cost of energy has generated a great deal of interest in microwave drying. The microwaves couple directly into the solvent and no energy is used to heat the air, the walls of the dryer, the conveyor, or the trays. This results in extremely efficient energy utilization, and energy savings of as much as 70% have been realized in industrial installations.

---

* Method of determining rate of drying: The difference in moisture content between any two measurements divided by the time period between measurements represents the rate of drying for this time period. This value is plotted against the midpoint of the time period for a drying rate versus time curve, or against the midpoint of the moisture content values for a drying rate versus moisture content curve.

# 5:
# Clarification and Filtration

Clarification may be defined as the process that involves the removal or separation of a solid from a liquid, or a fluid from another fluid. The term "fluid" encompasses both liquids and gases. Clarification can be achieved using either filtration or centrifugation techniques. Filtration is mainly required to remove unwanted solid particles from a liquid product or from air and centrifugation is normally used to separate fluid from another fluid or to collect the solid as the product.

*Filtration* is defined as the process in which particles are separated from a liquid by passing the liquid through a permeable material. The permeable medium is a porous material that separates particles from the liquid passing through it and is known as a *filter*. Thus, filtration is a unit operation in which a mixture of solids and liquid, the *feed, suspension, dispersion, influent* or *slurry,* is forced through a porous medium, in which the solids are deposited or entrapped. The solids retained on a filter are known as the *residue*. The solids form a *cake* on the surface of the medium, and the clarified liquid known as *effluent* or *filtrate* is discharged from the filter. If recovery of solids is desired, the process is called *cake filtration*. The term *clarification* is applied when the solids do not exceed 1.0% and filtrate is the primary product.

There are numerous applications of filtration in pharmaceutical processing which mainly include (i) clarification of products to improve their appearance, i.e. to give them 'sparkle' or 'brightness', (ii) removal of potential irritants e.g. from eye drop preparations or solutions applied to mucous membranes, (iii) filtration for recovery of desired solid material from a suspension of slurry, e.g. to obtain a drug or excipient after a crystallization process, (iv) production of water of appropriate quality for pharmaceutical use, (v) meeting sterility specification (removal of microorganisms) required for some products using *sterile filtration or aseptic filtration,* (vi) sterilization of solutions and suspensions that are chemically or physically unstable under heating conditions, (vii) detect ion of microorganisms present in liquids by

analyzing a suitable filter on which the bacteria are retained and (viii) assessment of the efficacy of preservatives. Recently, techniques such as nanofiltration, ultrafiltration, and microfiltration have been used to recover colloidal delivery systems from mother liquor.

## MECHANISMS OF FILTRATION

Four different mechanisms of filtration according to the way in which the suspended material is trapped by the filter medium are as follows:

## Surface Straining

In surface straining, any particle that is larger in size than the pores of the medium deposits on the surface, and stays there until it is removed. Particles that are smaller in size than the pores pass quickly through the medium. This is the main operating mechanism for bar screens, and plain woven monofilament plastic or wire mesh. Filters employing this mechanism are used where the contaminant level is low or small volumes need to be filtered.

## Depth Straining

Depth straining is also governed by particle size or shape. For filter media that are relatively thick in comparison with their pore diameters, particles will travel along the pore until they reach a point where the pore narrows down to a size too small for the particles to go any further, so that they become trapped.

## Depth Filtration

Depth filtration involves a complex mixture of physical mechanisms. The path followed by the fluid through a bed is extremely tortuous. Violent changes of direction and velocity occur as the system of pores and waists is traversed. As the flowing fluid passes filter medium, fluid flow pattern is disturbed. The suspended particles are first brought into contact with the pore wall (or very close to it), by inertial or hydraulic forces, or by Brownian (molecular) motion (*impingement*). They then become attached to the pore wall, or to another particle already held by means of van der Waals and other surface forces (*entanglement*). In depth filtration, the particles becomes entrapped in the depth of the medium, even though they are smaller in diameter, and possibly much smaller, than the pore at that point. This mechanism is important for most media, but especially for high efficiency air filters and for deep bed (sand) filters.

## Cake Filtration

Cake filtration (which is a development of surface filtration) begins with the formation of a layer of particles on the surface of the filter medium, with larger pores bridged by a group of smaller particles. On this layer, a cake of particles accumulates to act as the filter medium for subsequent filtration. Cake filtration in which solid recovery is the goal is an important pharmaceutical process. The most common industrial application is the filtration of slurries containing a relatively large amount of suspended solids, usually 3% to 20%. The physical properties of the cake largely determine the methods employed. Often, washing and partial drying or dewatering are an integral part of the process. Effective discharge of the cake completes the process. The solids, the filtrate, or both may be required.

These definitions emphasize that the mechanisms of filtration may result in the trapping of far smaller particles than might be expected from the size of the pores in the medium. The actual mechanism or combination of mechanisms in any specific instance is dependent on the characteristics of both the medium and the suspension being filtered. If the filter medium is thick with respect to the diameters of the fibers or particles of which it is composed, then the separation of the suspended particles from the fluid will occur by both depth straining and depth filtration and it will not be possible to distinguish between them. Depth filtration mechanism is also well suited for the process of clarification with its ability to remove particles considerably smaller than the average pore diameter of the filter medium.

## THEORY OF FILTRATION

Even today, filtration is more of an art than science. The filtration theory, with all its mathematical models, has a deficiency. The deficiency is its preoccupation with resistance to flow, almost to the exclusion of considerations of filtrate quality. It is possible to estimate the resistance to flow of a clean filter medium but impossible to estimate with comparable accuracy what the resistance will be as the filter begins to trap solids. The mathematical models do provide a means of showing apparent relationships between variables in a process and may be valuable decision-making tools in the selection of apparatus and techniques for a particular filtration application.

The mathematical models for flow through a porous medium, cake filtration, and granular bed filtration may differ, but all follow this basic rule: The energy lost in filtration is proportional to the rate of flow per unit area.

The flow of liquid through a filter follows the basic rules that govern flow of any liquid through a medium offering resistance. The rate of flow may be expressed as:

$$\text{Rate} = \frac{\text{Driving force}}{\text{Resistance}} \qquad \ldots (1)$$

The rate may be expressed as volume per unit time and the driving force as a pressure differential. The apparent complexity of the filtration equations arises from the expansion of the resistance term. Resistance is not constant since it increases as solids are deposited on the filter medium. An expression of this changing resistance involves a material balance as well as factors expressing permeability or coefficient of resistance of the continuously expanding cake.

The rate concept as expressed in modifications of *Poiseuille's equation* is prevalent in engineering literature and is given by:

$$\frac{dV}{dT} = \frac{AP}{\eta(\alpha\frac{W}{A} + R)} \qquad \ldots (2)$$

where, $V$ = volume of filtrate
$T$ = time
$A$ = Filter area
$P$ = total pressure drop through cake and filter medium

| | | |
|---|---|---|
| η | = | filtrate viscosity |
| $\alpha$ | = | average specific cake resistance |
| $W$ | = | weight of dry cake solids |
| $R$ | = | resistance of filter medium and filter |

Any convenient units may be used in this equation, since inconsistencies are absorbed in the cake and filter resistances.

The practical limitation of this equation is that the constants must be determined on the actual slurry being handled. There is no crossover application of data, and the majority of filters are selected on the basis of empiric laboratory or pilot plant tests. Equation (2) has been integrated under various assumptions, and these integrated forms may be used to predict the effects of process changes and to evaluate test work.

Consequently, it is desirable to attempt to use an equation in which the resistance of the bed may be expressed in terms of those characteristics of the bed that affect the resistance. Also, since the actual crosssectional area of flow is not known, it is necessary to replace this by the area of the bed. One such form of relationship was established by *Kozeny* which may be expressed as:

$$\frac{dV}{dT} = \frac{A\Delta P}{K\eta LS^2} \times \frac{\varepsilon^3}{(1-\varepsilon)^2} \qquad \text{... (3)}$$

| | | |
|---|---|---|
| where, $A$ | = | cross-sectional area of porous bed |
| $\varepsilon$ | = | porosity of bed |
| $V$ | = | volume of filtrate |
| $T$ | = | time |
| $\eta$ | = | filtrate viscosity |
| $S$ | = | surface area per unit particle volume |
| $L$ | = | bed thickness in direction of fluid flow |
| $K$ | = | permeability of bed |

The constant, *K*, generally ranges in value from 3 to 6. The *Kozeny equation* finds its greatest limitation in complex systems such as filter paper, but provides excellent correlation in filter beds composed of porous material.

In applying Poiseuille’s law to filtration processes, one must recognize that the capillaries found in the filter bed are highly irregular and non-

uniform. Therefore, if the length of the capillary is taken as the thickness of the bed or medium and the correction factor for the radius is applied, the flow rate is more closely approximated. These factors have been taken into account in the formation of the *Darcy's equation:*

$$\frac{dV}{dT} = \frac{KA\,\Delta P}{\eta L} \qquad \text{... (4)}$$

where, $K$ is the permeability coefficient and depends on the nature of the precipitate to be filtered and filter medium itself.

It is convenient to summarize the theoretic relationship as:

$$\text{Rate of filtration} = \frac{(\text{Area of filter}) \times (\text{Pressure difference})}{(\text{Viscosity}) \times (\text{Resistance of cake and filter})} \qquad \text{... (5)}$$

Interpretation of the basic equations, however, leads to a general set of rules:

1. Pressure increases usually cause a proportionate increase in flow unless the cake is highly compressible. Pressure increases on highly compressible, flocculent, or slimy precipitates may decrease or terminate flow.
2. An increase in area increases flow and life proportional to the square of the area since cake thickness and thus resistance, are also reduced.
3. Cake resistance is a function of cake thickness, therefore, the average flow rate is inversely proportional to the amount of cake deposited.
4. Particle size of the cake solids affects flow through effect on the specific cake resistance, $a$. A decreased particle size results in higher values of $R$ and proportionally lower filtration rates.
5. The filter medium resistance, $R$, usually negligible in cake filtration, is the primary resistance in clarification filtration. In the latter case, flow rate is inversely proportional to $R$.
6. The filtrate flow rate at any instant is inversely proportional to viscosity so, filtration efficiency also may be affected by changes in temperature. The viscosities of most liquids decrease with increase in temperature. Increasing the temperature of heavy pharmaceutical syrups lowers the viscosity and increases filtration rates. Most liquids must be maintained at

a high temperature during filtration to prevent the formation of crystals. The filtration of cosmetic products at low temperatures, approximately 5°C, is also common. The consequent reduction in flow rate is tolerated, since the goal is reduction in solubility of contaminants or perfume oils, resulting in their more effective removal. Equation (6) represents the relationship of the coefficient of viscosity with temperature.

$$\eta = Ae^{E/RT} \quad \text{... (6)}$$

where, $\eta$ = coefficient of viscosity of the liquid
$E$ = activation energy
$R$ = ideal gas constant
$T$ = absolute temperature
$A$ = pre-exponential factor

7. The permeability coefficient may be examined in terms of its two variables: porosity and surface area. Evaluation of the term $\varepsilon^3/(1-\varepsilon)^2$ shows that the permeability coefficient is a sensitive function of porosity. When slurry is filtered, the cake porosity depends on the way in which particles are deposited and packed. A fast deposition rate, given by concentrated slurries or high flow rates, may give a higher porosity because of the greater possibility of bridging and arching in the cake. Although theoretically the particle size has no effect on porosity, a broad particle size distribution may lead to a reduction of porosity if small particles pack in the interstices created by larger particles. Surface area, unlike porosity, is markedly affected by particle size and is inversely proportional to particle diameter. Hence, a coarse precipitate is easier to filter than a fine precipitate even though both may pack with the same porosity.

Most clarification problems can be resolved empirically by varying one or more of these factors. A broader understanding of filtration theory is required only if cake filtration applications are under consideration.

The membrane filters are highly porous. A number of methods are used for establishing the pore size and pore size distribution. Most methods are derived from the interfacial tension phenomenon of liquids in contact with the filter structure. Each pore in the filter acts as a capillary. For a non-wetting fluid, the following equation was established by Poiseuille:

$$p = \frac{-2\gamma \, \mathrm{Cos}\theta}{r} \qquad \ldots (7)$$

where, $p$ = applied pressure
$\gamma$ = liquid surface tension
$\theta$ = contact angle between the liquid and solid
$r$ = radius of the pore

## FILTER MEDIA

The surface upon which solids are deposited in a filter is called the *filter medium*. For the pharmacist selecting this important element, the wide range of available materials may be bewildering. The selection is frequently based on past experience, and reliance on technical services of commercial suppliers is often advisable. The ideal filter material should have the following characteristics:

- A medium for cake filtration must retain the solids without plugging and without excessive bleeding of particles at the start of the filtration. In clarification applications in which no appreciable cake is developed, the medium must remove all particles above a desired size.
- It should offer minimum resistance and the resistance offered by the medium itself will not vary significantly during the filtration process.
- It allows easy discharge of cake.
- It should be chemically and physically inert.
- It should not swell when it is in contact with filtrate and washing liquid.
- It should have sufficient mechanical strength to withstand pressure drop and mechanical stress during filtration.

There are a variety of different depth filter and membrane filter materials used in pharmaceutical processes. Depth filters are mainly polymeric fibrous materials. The filter fabrics are commonly woven from natural fibers such as cotton and from synthetic fibers and glass. The properties of these fibers and glass applicable for media selection are tabulated in Table 5.1.

**Table 5.1:** Fiber properties for filter media selection

| Fiber | Temperature recommended safe limit (°F) | Wet breaking tenacity (g/denier) | | Acid resistance | Alkali resistance | Price ratio to cotton |
|---|---|---|---|---|---|---|
| Cotton | 210 | 3.3 | 6.4 | Poor | Fair | 1 |
| Polyester (dacron) | 300 | 6.0 | 8.2 | Very good | Good | 2.7 |
| Dynel modacrylic | 200 | 3.0 | – | Excellent | Excellent | 3.2 |
| Glass (spun) | 750 | 3.0 | 4.6 | Excellent | Fair | 6.0 |
| Glass (filament) | 550 | 3.9 | 4.7 | Excellent | Fair | 2.2 |
| Nylon | 250 | 2.1 | 8.0 | Fair | Excellent | 2.5 |
| Acrylic (orion) | 300 | 1.8 | 2.1 | Excellent | Fair | 2.7 |
| Polyethylene | 165 | 1.0 | 3.0 | Excellent | Excellent | 2 |
| Polypropylene | 175 | 3.5 | 8.0 | Excellent | Excellent | 1.75 |
| Saran | 160 | 1.2 | 2.3 | Excellent | Excellent | 2.5 |
| Teflon | 475 | 1.9 | – | Excellent | Excellent | 25.0 |
| Polyvinyl chloride | 165 | 1.0 | 3.0 | Good | Excellent | 2.7 |
| Wool | 210 | 0.76 | I 1.6 | Very good | Fair | 3.7 |
| Rayon and acetate | 210 | 1.9 | 3.9 | Poor | Fair | 1 |

*Filter cloth,* a surface type medium, is woven from either natural or synthetic fiber or metal. Cotton fabric is the most common and is widely used as a primary medium, as backing for paper or felts in plate and frame filters, and as fabricated bags for coarse straining. Nylon is often superior for pharmaceutical use, since it is unaffected by molds, fungi, or bacteria, provides an extremely smooth surface for good cake discharge, and has negligible absorption properties. Both cotton and nylon are suitable for coarse straining in aseptic filtrations, since they can be sterilized by autoclaving. Monofilament nylon cloth is extremely strong and is available for openings as small as 10 μm. Teflon is superior for most liquid filtration, as it is almost chemically inert, provides sufficient strength, and can withstand elevated temperatures.

*Woven wire cloth,* particularly stainless steel, is durable, resistant to plugging, and easily cleaned. Metallic filter media provide good surfaces for cake filtration and are usually used with filter aids. As support elements for disposable media, wire screens are particularly suitable, since they may be cleaned rapidly and returned to service. Wire mesh filters also are installed in filling lines of packaging equipment. Their function at this point is not clarification, but security against the presence of large foreign particles.

*Non-woven filter media* include felts, bonded fabrics, and kraft papers. A *felt* is a fibrous mass that is free from bonding agents and mechanically

interlocked to yield specific pore diameters that have controlled particle retention. High flow rate with low pressure drop is a primary characteristic. Felts of natural or synthetic material function as depth media and are recommended where gelatinous solutions or fine particulate matter are involved. *Bonded fabrics* are made by binding textile fibers with resins, solvents, and plasticizers. These materials have not found wide acceptance in dosage form production because of interactions with the additives. *Kraft* paper is a pharmaceutical standard. Although limited to use in plate and frame filters and horizontal-plate filters, it offers controlled porosity, limited absorption characteristic, and low cost. The latter is important since concern over cross-contamination makes a disposable medium attractive to pharmacy. White papers are preferred, and they may be crinkled to produce greater filtration area. A support of cloth or wire mesh is necessary in large filter presses to prevent rupture of the paper with pressure.

*Porous stainless steel filters* are widely used for the removal of small amounts of unwanted solids from liquids (clarification) such as milk, syrup, sulfuric acid, and hot caustic soda. Porous metallic filters can be easily cleaned and repeatedly sterilized.

*Membrane filter media* are the basic tools for microfiltration, ultrafiltration, nanofiltration and reverse osmosis. They are used commonly in the preparation of sterile solutions. Membrane filters, classified as surface or screen filters, are made of various esters of cellulose or from nylon, Teflon, polyvinyl chloride, polyamide, polysulfone, or silver. The filter is a thin membrane, about 150 μm thick, with 400 to 500 million pores per square centimeter of the filter surface. The pores are extremely uniform in size and occupy about 80% of filter volume. The high porosity permits flow rates at least 40 times higher than those obtained through other media of comparable particle retention capability. The fragility of membrane filters is partially overcome by the use of monofilament nylon as a supporting web within the membrane structure. Because of surface screening characteristics, prefiltration is often required to avoid rapid clogging of a membrane. The selection of a membrane filter for a particular application is a function of the size of the particle or particles to be removed. An approximate pore size reference guide is given in Fig. 5.1 and Table 5.2.

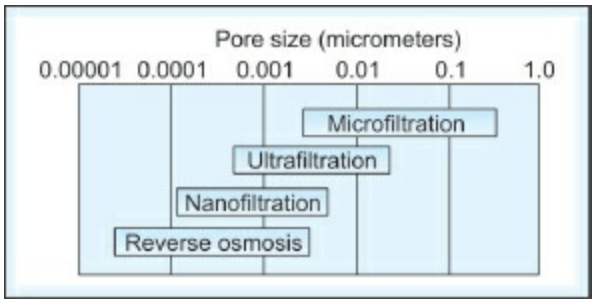



**Fig. 5.1:** Pore size reference guide

**Table 5.2:** Tabular representation of pore size reference guide

| Pore size (µm) | Particle removed |
|---|---|
| 0.2 (0.22) | All bacteria |
| 0.45 | All coliform group of bacteria |
| 0.8 | All airborne particles |
| 1.2 | All nonliving particles considered dangerous in IV fluids |
| 5 | All significant cells from body fluids |

The distinction between nanofiltration, ultrafiltration and microfiltration lies in the nature of the filter medium (Fig. 5.1). Nanofiltration and ultrafiltration membranes contain pores of relatively narrow size distribution of $10^{-4}$ to $10^{-2}$ µm and are formed by etching cylindrical pores into a solid matrix. Most commercial ultrafiltration and nanofiltration membranes are polyamides, such as nylon, polyethersulfon (PESU), or polyvinyldienefluoride (PVDF). Track-etched microfiltration membranes are made from polymers such as polycarbonate and polyester, wherein electrons are bombarded onto the polymeric surface creating sensitized tracks. The resulting irradiated film is placed in an etching bath (such as a basic solution), in which a damaged polymer in the tracks is preferentially etched from the film, thereby forming cylindrical pores.

Synthetic and natural fibers, cellulose esters and fiberglass, fluorinated hydrocarbon polymers, nylon, and ceramics are employed for the manufacture of disposable cartridges. These cartridges are economical and

convenient when used to remove low percentages of solids ranging in particle size from 100 μm to less than 0.2 μm. Porous materials for cleanable and reusable cartridges use stainless steel, monel, ceramics, fluorinated hydrocarbon polymers, and exotic metals.

*Surface-type cartridges* of corrugated, resin-treated paper are common in hydraulic lines of processing equipment, but are rarely applied to finished products. Ceramic cartridges have the advantage of being cleanable for reuse by back-flushing. Asbestos and porcelain filter candles are acceptable for some sterile filtrations along with membrane filters. Sintered metal or woven-wire elements are also useful, but fine-wire mesh lacks strength. The metallic-edge filters overcome this problem by allowing liquid to pass between rugged metal strips, which are separated by spacers of predetermined thickness. ***Depth-type cartridges*** consist of fibrous media, usually cotton, asbestos, or cellulose.

## FILTER AIDS

Usually, the resistance to flow due the filter medium itself is very low, but increases as a layer of solids builds up, blocking the pores of the medium and forming a solid, impervious cake. Poorly flocculated solids offer higher resistance than do flocculated solids or solids providing high porosity to the cake. In the case of cake filtration, the rate varies with the square of the volume of liquid. When the volume of the filter cake solids per unit volume of filtrate is low, the solids deposited on the filter medium may penetrate the void space, thus making the filter medium more resistant to flow. At a higher concentration of solids in a suspension, the bridging over of openings over the void space, rather than blinding of the openings, seems to predominate. Slimy or gelatinous materials, or highly compressible substances, form impermeable cakes with a high resistance to liquid flow. The filter medium becomes plugged or slimy with the accumulation of solids, and the flow of filtrate stops. A filter aid acts by reducing this resistance.

Filter aids are a special type of filter medium. Ideally, the filter aid forms a fine surface deposit that screens out all solids, preventing them from contacting and plugging the supporting filter medium. Usually, the filter aid acts by forming a highly porous and noncompressible cake that retains solids, as does any depth filter. The duration of a filtration cycle and the clarity attained can be controlled as density, type, particle size, and quantity of the filter aid are varied. The quantity of the filter aid greatly influences the filtration rate. If too little filter aid is used, the resistance offered by the filter cake is greater than if no filter aid is used, because of the added thickness to the cake. On the other hand, if high amounts of filter aid are added, the filter aid merely adds to the thickness of the cake without providing additional cake porosity. Figure 5.2 is a typical plot of filter aid concentration versus permeability. In the figure, flow rate and permeability are directly proportional to each other. At low concentrations of filter aid, the flow rate is low because of low permeability. As the filter aid concentration increases, the flow rate increases and peaks off. Beyond this point, the flow rate decreases as the filter aid concentration is increased.

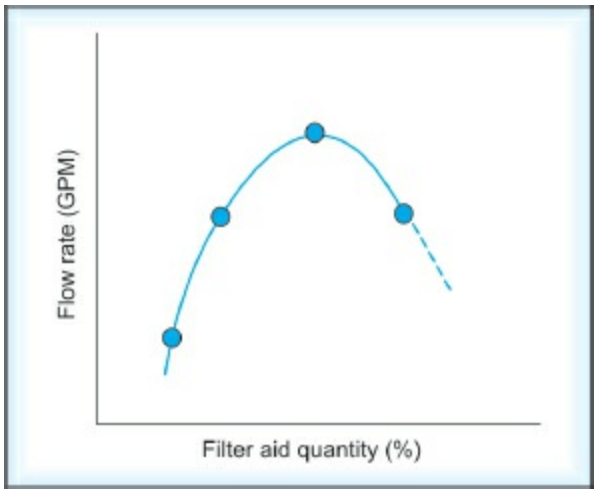



**Fig. 5.2:** Experimental determination of flow rate as a function of filter aid quantity discloses correct operating level

The ideal filter aid performs its functions physically or mechanically and no absorption or chemical action is involved in most cases. The important characteristics for filter aids are the following:

1. It should have a structure that permits formation of pervious cake.
2. It should have a particle size distribution suitable for the retention of solids, as required.
3. It should be able to remain suspended in the liquid.
4. It should be free of impurities.
5. It should be inert to the liquid being filtered.
6. It should be free from moisture in cases where the addition of moisture would be undesirable.

The particles must be inert, insoluble, incompressible, and irregularly shaped. Filter aids are classified from low flow rate (fine: Mean size in the range of 3 to 6 μm) to fast flow rate (coarse: Mean size in the range of 20 to 40 μm). Clarity of the filtrate is inversely proportional to the flow rate, and selection requires a balance between these factors. Filter aids are considered to be equivalent in performance when they produce the same flow rate and filtered solution clarity under the same operating conditions when filtering a standard sugar solution. Table 5.3 lists the advantages and disadvantages of filter aid materials.

**Table 5.3:** Advantages and disadvantages of filter aid materials

| Material | Chemical composition | Advantages | Disadvantages |
|---|---|---|---|
| Diatomaceous earth | Silica | Wide size range available; fines reduced by calcination; can be used for very fine filtration. | Slightly soluble in dilute acids and alkalies. |
| Expanded perlite | Silica and aluminosilicates | Wide size range available; not capable of finest retention | More soluble than diatomites in acids and of diatomites. alkalies; may give highly compressible cakes. |
| Asbestos | Aluminosilicate | Usually used in conjunction with diatomites; very good retention on coarse screens. | Chemical properties similar to perlite. |
| Cellulose | Cellulose | Used mainly as a coarse precoat; high purity; excellent chemical resistance, slightly soluble in dilute and strong alkalies, none in dilute acids. | Expensive |
| Carbon | Carbon | May be used for filtering strong alkaline solutions. | Available in coarser grades only; expensive. |

*Diatomite* (diatomaceous earth) is the most important filter aid. Processed from fossilized diatoms, it has irregularly shaped porous particles that form a rigid incompressible cake. Since diatomite is primarily silica, it is relatively inert and insoluble. *Perlite*, an aluminum silicate, forms filter cakes that are 20 to 30% less dense than diatomaceous cakes. Perlite does not comprises porous incompressible particles, but it has an economic advantage over diatomite.

*Cellulose*, *asbestos*, and *carbon* filter aids are also commercially available. Cellulose is highly compressible and costs two to four times more than diatomite or perlite. It is reserved for applications where the liquids may be incompatible with silica compounds. Cellulose is used as a coarse precoat. It is available as high-purity material and has excellent chemical resistance. Asbestos has good retention on coarse screens, but has limited application because of its high cost, and leaching of fibers into the filterate that might be toxic. Asbestos filters may be used in pharmaceutical industry if their application is followed by membrane filtration. Nonactivated carbons that are not suitable for decolorization or absorption are rarely used in pharmaceutical applications because of cleanliness problems. They may be used for filtering strong alkaline solutions. Commercial blends of various filter aids are

common, and their specialities must be considered in selection.

New, high performance filter aids with self flocking (SF) property provide low tortuosity and fine particle filtration with high flow rates. These filter aids are compounded calcined rice hulls that coagulate extremely fine particles into large, rigid, permeable, flocculated particles. *MaxFlo SF* is valued for its formation of superior, high-density filter cakes at low operating costs, whereas *ProFix SF* sequesters metals and creates a stabilized filter cake.

Filter aids may be applied by *precoating* or *body-mix* techniques. Precoating requires suspending the filter aid in a liquid and recirculating the slurry until the filter aid is uniformly deposited on the filter septum. The quantity varies from 5 to 15 pounds per 100 square feet of filter area, or that sufficient to deposit a cake 1/16 to 1/8 inches thick. The liquid is preferably a portion of the feed or retained filtrate from a prior cycle, since the physical properties of the precoat liquid must approximate those of the material to be filtered. Precoating should proceed at the same flow rates and pressures to be used in final filtration, and the transition from precoat liquid to regular feed must be rapid to prevent disruption of the cake. Body mix (direct addition of filter aid to the filter feed) is more common in batch pharmaceutical operations. The filter aid, 1 to 2 pounds per pound of the contaminant, or 0.1 to 0.5% of the total batch weight, is added to the feed tank. This slurry is recirculated through the filter until a clear filtrate is obtained; filtration then proceeds to completion. The body-mix method minimizes equipment requirements and cross-contamination potentials.

Often, a filter aid performs its function not physically or mechanically, but chemically, by reacting with the solids. These chemicals may cause the solids depositing in a filter bed to adhere more strongly to the filter medium. Water-soluble polymers such as flocculating agents are often used as filter aids. The polymers may be derived from vegetable or animal sources, or they may be produced synthetically. Water-soluble polymers may be classified as nonionic, anionic, or cationic, depending on their property to ionize in water. There are a few commercially available water-soluble cationic polymers. These include acrylamide copolymers, polyethyle-neimine, and derivatives of casein, starch and guar gum.

Filter aids are chosen by trial and error in either laboratory or plant. Within the ranges previously indicated, the filter aid is usually selected to

give acceptable filtrate at the highest flow rate; however, in pharmaceutical operations in which quality is a primary consideration, the selection usually favours the fine grades, which yield low flow rates.

The most important pharmaceutical factor is inertness. A filter aid may have such extensive absorption properties that desired coloured substances and active principles are frequently removed. The total quantity of any ingredient absorbed may be small, but it may be a considerable portion of the original concentration.

## FILTRATION EQUIPMENT

Commercial filtration equipment is classified by the end product desired (filtrate of cake solids), by the method of operation (batch or continuous), by type of operation (non-sterile filtration, sterile filtration, centrifugation filtration, centrifugation sedimentation), but most importantly by the type of driving force (gravity, vacuum, pressure and centrifugation) as depicted in Fig. 5.3.

The clarification demands of pharmaceutical processes are usually met by batch pressure units. Compatibility with a wide range of products restricts the materials of construction to stainless steel, glass, and inert polymers.

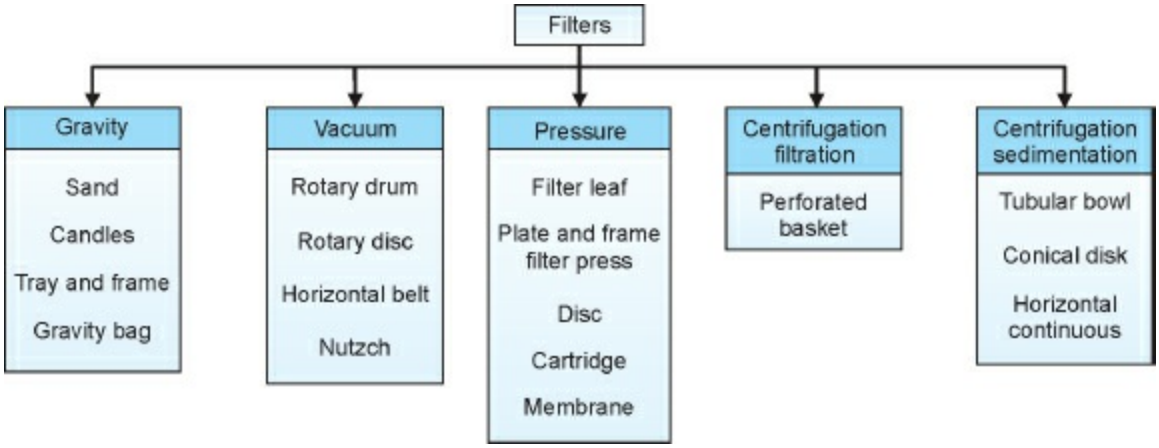



**Fig. 5.3:** Classification of filters and based on the driving force

## Gravity Filters

Gravity filters rely on gravity generated low operating pressure (usually less than $1.03 \times 10^4$ N/m$^2$) and give low filtration rates unless very large surfaces areas are used, which limits their use on a large scale. However, these are simple and cheap, and are frequently used in laboratory filtration where volumes are small and low filtration rate is relatively insignificant. Gravity filters employing thick, granular beds are common in water treatment, where clarification of water is done prior to deionization or distillation. The filtering medium may consist of sand or cake beds, or for special purposes, a composition containing asbestos, cellulose fibers, activated charcoal, diatomaceous earth, or other filter aids. *Sand filters* are the most preferred for water treatment and are of two main types that operate with gravity flow downwards through a deep bed-the "slow" and "rapid" sand filters. *Slow sand filters* operate with a velocity of 0.1–0.2 m/h down through the bed. They actually function largely by straining through a biological layer that forms on the surface of the bed. They are cleaned occasionally by the reasonably complete removal of this layer, without disturbing the rest of the bed. *Rapid sand filters* utilize a velocity of 5–15 m/h and function by depth filtration within the bed. They are cleaned frequently by cessation of process flow, followed by a reverse upward flow of wash water at such a rate that the bed expands and releases the trapped dirt particles. This cleaning flow may be augmented by some form of agitation, such as injecting compressed air below the bed or hydraulic jets impinging on the surface. Small-scale purification of water may use porous ceramics as a filter medium in the form of hollow "*candles*". The fluid passes from the outside through the porous ceramics into the interior of the hollow candles. *Tray and frame filters* are best adapted for slow, difficult filtrations and for exceptionally soft- or fine-grained precipitates, which clog the filters under the slightest pressure or pass through the openings of a cloth. *Gravity bag filters* are applied to the concentration of magmas, such as milk of magnesia. More efficient methods, however, particularly with respect to space requirements, are available. The *gravity nutzch* is a false-bottom tank or vessel with a support plate for filter media. Usually, the bottom is dressed with a cloth. The slurry is added and the material filters under its own hydrostatic head. The filtrate is collected in the sump beneath the filter. Thorough washing is possible either by simple displacement and diffusion or by resuspending the solids in a wash liquid and

refiltering. The nutzch is comparatively difficult to empty, and labour costs are high. Various new gravity filter systems are available commercially such as cylindrical gravity filters, rectangular gravity filters, and hydro-clear gravity filters which utilize granular particles in a basin. Fluid streams pass through the basin and particles are physically and/or chemically captured by the media.

## Vacuum Filters

These are employed on a large scale, but are rarely used for the collection of crystalline precipitates or sterile filtration. Continuous vacuum filters can handle high dirt loads, and on a volume basis, are cheap in terms of cost per gallon of the filtered fluid. Vacuum filters are simple and reliable machines, and therefore have gained wide acceptance in the chemical, food and pharmaceutical industries. Except nutzch filter, all have continuous operation. For large-scale operations, continuous vacuum filters are the most widely used.

### *Rotary Drum Filter*

The rotary drum vacuum filter is divided into sections, each connected to a discharge head (Fig. 5.4). Each filter unit is rectangular in shape with a curved profile so that a number can be joined up to form a drum. Each unit has a perforated metal surface and is covered with filter cloth. The slurry is fed to a tank in which solids are held in suspension by an agitator. As the drum rotates, each section passes through the slurry and vacuum draws filtrate through the filter medium at the drum surface (*pick-up zone*). The suspended solids deposit on the filter drum as a cake, and as rotation continues, vacuum holds the cake at the drum surface. This is followed by washing and further drainage in the drying zone. As the cake moves towards the discharge point, it may be scraped from the drum or it may be supported by strings until it breaks free under gravitational forces (cake removal zone). The cake discharge may be done through a scraper, belt, roll or a string. Scraper discharge mechanisms will suit cakes that could be scraped readily and roller discharge mechanism are better for thixotropic cakes. A *string discharge* filter is useful in the manufacture of antibiotics, when the removal of cake of mould mycelia is necessary.

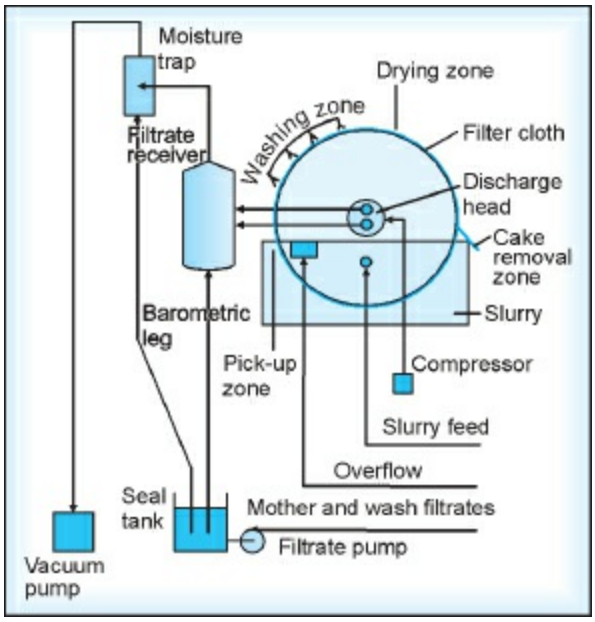



**Fig. 5.4:** Rotary drum filter

For solids that tend to block the filter cloth, a precoat of filter aids such as diatomaceous earth, perlite and cellulose is deposited on the drum prior to the filtration process. These materials serve as a filter medium in an analogy to the filter cloth on a conventional drum filter. The scraper on a precoat filter, called as Doctor Blade, shaves-off the blinding layer of the contaminants together with a thin layer of the precoating material thus exposing a fresh layer of the precoat surface. Precoat filters are generally used where a very high degree of clarity is required and solids content is very low (less than 2–3%) or where solids are sticky and would otherwise clog the filter cloth.

## *Rotary Disc Filter*

It consists of several discs, up to 15 in the larger machines, each made up from sectors which are clamped together to form the disc (Fig. 5.5). Each sector is connected to a vacuum system, compressed air, and appropriate receivers, in the correct sequence, by means of special rotating valve. The operation sequence of a disc filter is similar to a drum filter. The main feature of disc filter is less floor space and the lowest cost of filtration when compared to other vacuum filters.

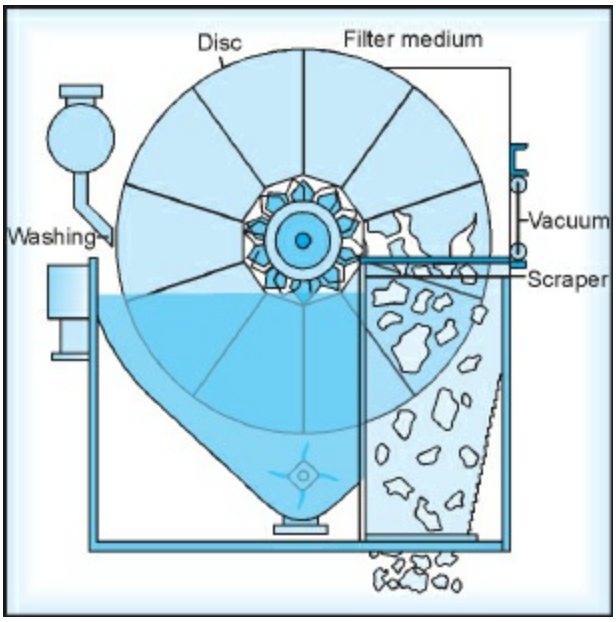



**Fig. 5.5:** Disc filter

## *Horizontal Belt Filters*

Another adaptation of rotary vacuum filters is top feed horizontal belt filters illustrated in Fig. 5.6. Horizontal belt filters incorporate an endless and thick rubber belt of a complex design to support the cake retained by the filter cloth. Once the belt reaches the end of the vacuum box the cake drying portion of the cycle terminates and the cloth leaves the rubber belt. The cloth continues moving, changes direction over the discharge roll and the cake drops through a chute for further handling.

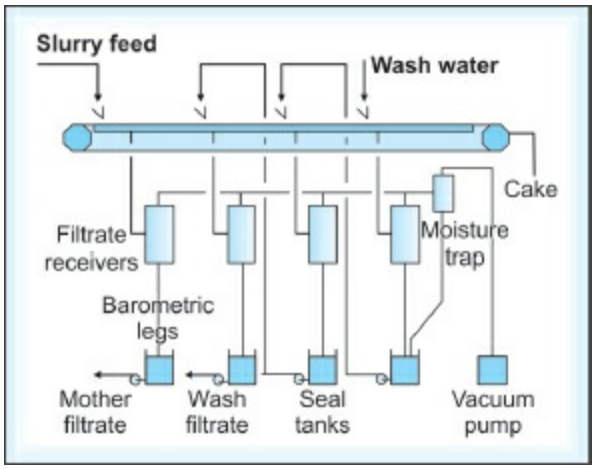



**Fig. 5.6:** Schematic representation of working of horizontal belt filter

A pneumatic or electrical tracking mechanism controls the filter cloth from slipping sideways by guiding it to the left or right. These filters are flexible in operation and suitable to handle large throughputs. To fulfill the demand for more area and higher throughputs, *tilting pan filters* were designed (Fig. 5.7). The operation of tilting pan filters is based on a series of horizontal independent trapezoidal pans mounted on a circular supporting structure that rotates under vacuum during the filtration cycle and invert under blow to discharge the cake.

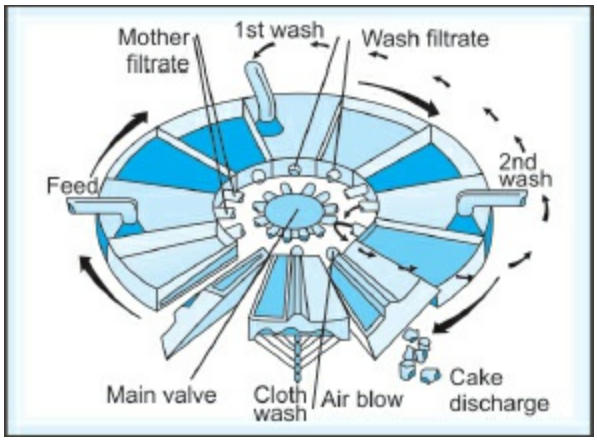



**Fig. 5.7:** Schematic representation of tilting pan filter

## *Nutzch Filter*

The nutzch filter is the industrial version of the well known laboratory scale, *Buchner* funnel except that it is designed to operate under either vacuum or pressure (Fig. 5.8). Nutzch filters are well suited for handling flammable, toxic, corrosive and odour-noxious materials since they are autoclaved and designed for use in hazardous and exproof environments when extremely safe operation is required. The filter consists of a densely perforated plate sufficiently strong to hold the cake weight and the pressure that is exerted on the cake's surface. The filters are available with the paddles consisting of two arms with slanted blades that rotate in one direction to re-slurry the cake during washing and discharge it at the end of the cycle. The filter is charged with slurry and pressure is applied to displace the filtrate leaving the cake retained over the filter medium.

**Fig. 5.8:** Schematic representation of Nutzch filter

## Pressure Filters

Most of the pressure filters are batch operated but continuous filters are also available. However, owing to the difficulty in removing the cake they are mechanically complex and expensive so mainly used where the added value to the product is high. The filtration rate is influenced, in broad terms, by the properties of the slurry. The trend is that the rate goes up with increased pressure, coarser particles, particle distribution with high uniformity, non-slimy or non-gelatinous solids, noncompressible cakes, lower liquid viscosity and higher temperatures.

### *Filter Leaf*

The filter leaf, as depicted in Fig. 5.9, is probably the simplest form of filter, consisting of a frame enclosing a drainage screen or grooved plate, the whole unit being covered with filter cloth. The outlet of the filtrate connects to the inside of the frame which represents a vertical section through the leaf. The frame may be of circular, square or rectangular shape. In use, the filter leaf is immersed in the slurry and the vacuum system is connected to the filtrate outlet. The method has the advantage that the slurry can be filtered from any vessel and cake can be washed simply by immersing the filter in a vessel of water. An alternative method is to enclose the filter leaf in a vertical vessel. A typical example of this kind is *sweetland filter or candle filter*. The slurry is pumped under pressure into a vessel that is fitted with a stack of leaves that serve as filter elements.

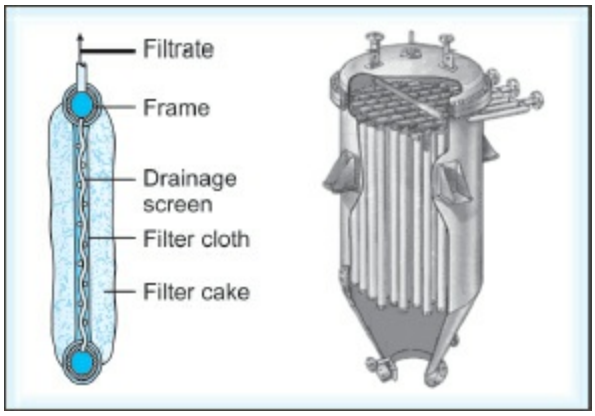



**Fig. 5.9:** Schematic representation of filter leaf

### *Plate and Frame Filter Press*

******ebook converter DEMO Watermarks*******

The plate and frame filter press is the simplest of all pressure filters and is the most widely used (Fig. 5.10). Filter presses are used for a high degree of clarification of the fluid and for the harvesting of the cake. When clarity is the main objective, a “batch” mode of operation is applied. The filter media are supported by structures in a pressure vessel. When an unacceptable pressure drop across the filter is reached during the filtration process, the filter media are changed. Methods of supporting the filter media include horizontal plates, horizontal or vertical pressure leaf, and plate and frame. As the name implies, the plate and frame filter press is an assembly of hollow frames and solid plates that support filter media. When assembled alternately with a horizontal or a vertical unit, conduits permit flow of the slurry into the frames and through the media. One side of the plate is designed for the flow of the feed. After passing the filter media, the filtrate is accommodated on the other side. The solids collect in the frames, and filtrate is removed through place conduits. In cake filtration, the size of the frame space is critical, and wide sludge frames are used.

The filter press is the most versatile of filters since the number and type of filter sheets can be varied to suit a particular requirement. It can be used for coarse to fine filtrations, and by special conduit arrangements, for multistage filtration within a single press. The filter press is the most economical filter per unit of filtering surface, and its material of construction can be chosen to suit any process conditions. Labour costs in assembly and cleaning are a primary disadvantage, and leakage between the plates may occur through faulty assembly. The normal range of flow is three gallons per minute per square foot of the filter surface at pressures of upto 25 psi.

Figure 5.10 illustrates the flow schematics through a plate. Fluid flows from the outside along the thin flow channel in the plate. The filtrate flows along similar channels in the bottom plate, and then to the inside circumference. This type of filter is intended only for clarification operations. Flow rates are similar to plate and frame presses at operating pressures of upto 50 psi. Pulp packs or *filter-masse* may be used instead of disc sheets for high-polish filtrations, but flow rates are then appreciably lower. Maximum filtrate recovery by air displacement of liquid is usually possible with a disc filter. Pressure leaf filters utilize the rotation of a pressure leaf to partially remove the cakes and extend the life of the filter media.

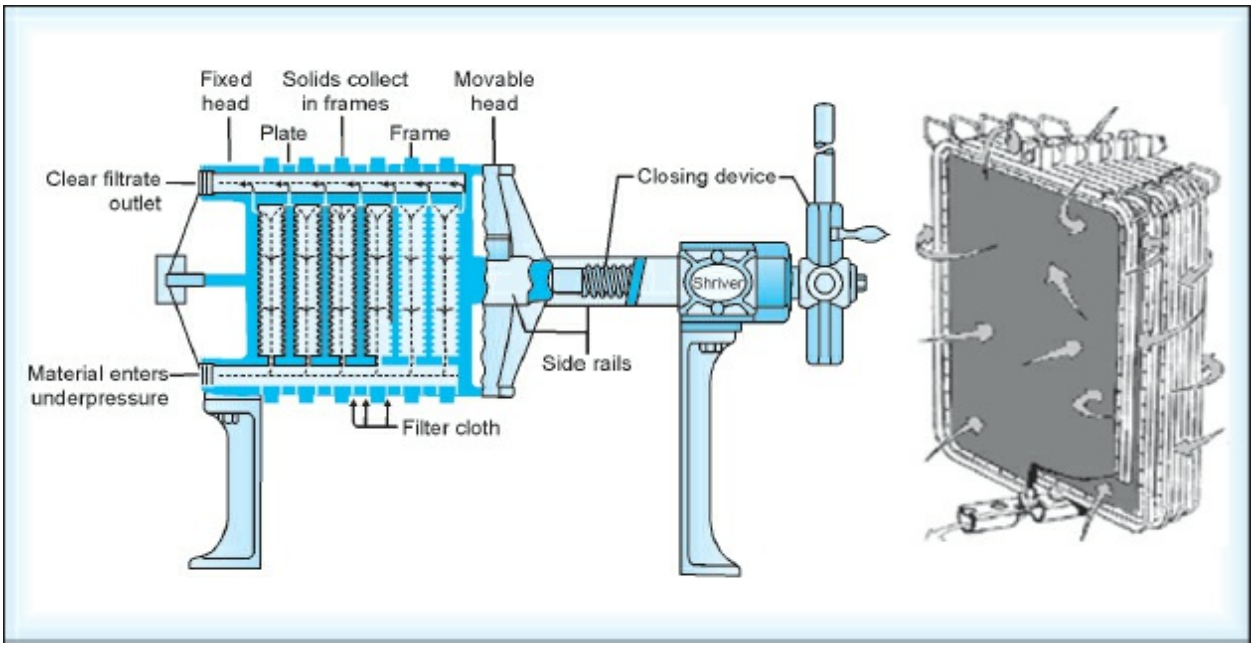



**Fig. 5.10:** The plate and frame filter-press may have 10 to 100 filtering surfaces and may be filled with pumps, sanitary fittings, sludge frames, or dividing plates for serial filtration

When filter aids are required, a plate and frame press with sludge frames is generally acceptable, but disposal of cake and cleaning becomes time-consuming. The *precoat pressure filter* is designed to overcome this objection. It consists of one or more leaves, plates, or tubes upon which a coat of filter aid is deposited to form the filtering surface. The filter area is usually enclosed within a horizontal or vertical tank, and special arrangements permit discharge of spent cake by backflush, air displacement, vibration, or centrifugal action. This type of filter is desirable for high-volume processes. Two or more units can be used alternatively, or surge tanks for clear filtrate may permit intermittent operation of a single unit.

## *Disc Filters*

The term disc filter is applied to assemblies of felt or paper discs sealed into a pressure case (Figs 5.11A and B). The discs may be preassembled into a self-supporting unit, or each disc may rest on an individual screen or plate. Single plate or multiples of single plates may be applied. The flow may be from the inside out wards or outside in wards. The disc filter overcomes some deficiencies of the filter press. Compactness, portability, and cleanliness are obvious advantages for pharmaceutical batch operations.

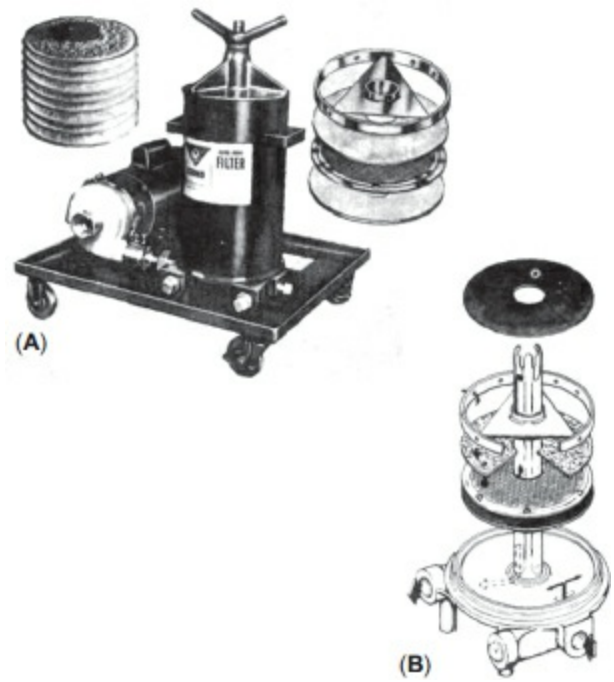



**Fig. 5.11:** Disc filter: (A) Pre-compressed cartridges or disc media; (B) Assembled disc

## *Cartridge Filters*

Most types of filter media are also available as cartridge units. These cartridges are economical and convenient when used to remove low percentages of solids ranging in particle size from 100 μm to less than 0.2 μm. The cartridge may be a surface or depth filter and consists of a porous medium integral with plastic or metal structural hardware. Synthetic and natural fibers, cellulose esters and fiberglass, fluorinated hydrocarbon polymers, nylon, and ceramics are employed for the manufacture of disposable cartridges. Porous materials for cleanable and reusable cartridges use stainless steel, monel, ceramics, fluorinated hydrocarbon polymers, and exotic metals.

*Surface-type cartridges* of corrugated, resin-treated paper are common in hydraulic lines of processing equipment, but are rarely applied to finished products. Ceramic cartridges have the advantage of being cleanable for reuse by back-flushing, and porcelain filter candles are acceptable for some sterile filtrations along with membrane filters in cartridge form. Sintered metal or woven-wire elements are also useful, but fine-wire mesh lacks strength. The metallic-edge filters overcome this problem by allowing liquid to pass

between rugged metal strips, which are separated by spacers of predetermined thickness. *Depth-type cartridges* consist of fibrous media, usually cotton, asbestos, or cellulose. The cartridge may be formed by felting or by resin-bonding fibers about a mandrel. Effective units are also manufactured by winding yarn around a central supporting screen. The depth cartridge is always a disposable item since cleaning is not feasible.

Cartridge filters have an integral cylindric configuration made with disposable or cleanable filter media and utilize either plastic or metal structural hardware. With the discovery of strong membranes such as cellulose nitrate, polyamide, polyvinylidene chloride, PTFE, and nylon, cartridge filters have revolutionized the filtration industry. Cartridge filters provide maximum filtration area in the smallest possible package, allow quick changeout of the media, and save time and money. Cartridge filters of different shapes, structures, forms, and sizes for different applications in the pharmaceutical industry are now available in disposable and non/disposable forms. The housings for cartridge filters come in a wide variety of configurations for both micron and submicron filtration. The major differences in various housings are in the design, materials of construction, seals that are used to install the cartridge in the housing, and the application for which they are used in the pharmaceutical industry. The housing for cartridge filters is described in terms of the height of the cartridge and number of cartridge receptacles in the base end of the housing. When a user purchases housing from one manufacturer, he is usually not “locked in” to that manufacturer’s cartridges. Adaptors are available that allow the cartridge filter of one manufacturer to fit into virtually any other manufacturer’s housing.

Filter media can be formed into cartridge form by tubular-wound, string-wound, or pleated formation. Alternate layers of filter media and separator material are rolled into a spiral configuration, and by potting the ends of the cartridge, form the “dead-ended” or “cross-flow” type of flow channels. String-wound cartridges are the most commonly used and inexpensive filters available. Pleated cartridges are modified tubular configurations with a large filtration area. A single knife-edge flat gasket may be a satisfactory seal for cartridge filters with 1.0-pm or larger pore size. For submicron filtration, the most satisfactory seal is an O-ring.

*Disposable or permanent cartridge filters* are used for fluid clarification

or sterilization. Standard elements for nonsterile filtration may be interchanged between cartridge holders offered by several companies. Increases in capacity result from multi-element holders, and 12 element units are usually adequate for batches of 500 to 1,000 gallons. The cost of disposable elements is offset by labour savings inherent in the simplicity of assembly and cleaning of cartridge clarifiers.

The *metallic edge filters*, particularly those with self-cleaning devices (Fig. 5.12), are excellent security filters for suspensions that may plug or blind conventional wire mesh. A cleaning blade combs away accumulated solids, which fall into a sump in the filter casing. A quick-coupling *metal cartridge filter* with construction that prevents short-circuiting of the filter element is also available. The special design permits rapid disassembly as well as interchange of reusable filter media (Fig. 5.13). Metal elements permit particle retention as low as 1.5 μm. Duo filters, two units connected in parallel, are recommended where uninterrupted service is required. A high-frequency vibrator, acting only on the element, assists in filtration of slurries that have blinding tendencies.

Vendors of membrane filters offer cartridge units in single-and multiple-element configurations. These cartridges have become the unit of choice for high-volume, sterile filtrations and are ideal for in-line, final polish prior to bottling of bulk parenterals. Cartridge filters having absolute ratings of 0.04 μm are also available. The latter units have 5 to 10 square feet of effective filtering area per cartridge of 10 inch height, and some can also be steam-sterilized.

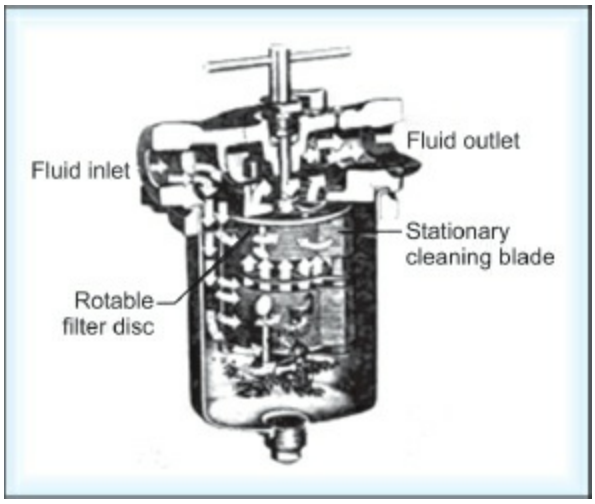


**Fig. 5.12:** An edge filter with automatic cleaning device may be automated by replacing handle with motor

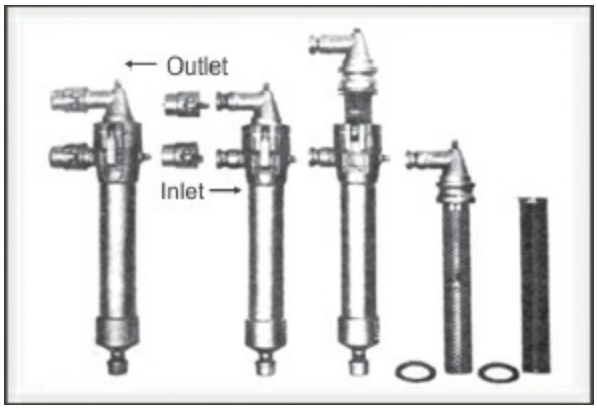



**Fig. 5.13:** Quick-coupling cartridge filters for metallic media is readily cleaned

## *Membrane Filters and Housings*

The use of membrane cartridge filters and housings has been discussed extensively in the previous section. The following section deals mainly with disc membranes and holders.

*Membrane filter holders* accept membranes from 13 to 293 mm in diameter. A useful rule of thumb for membrane media and holder sizes for various volumes of a low-viscosity liquid is shown in Table 5.4. Although 90 or 142 mm units are suitable for moderate volumes, the 293 mm membrane holder is the usual production choice for small-batch sizes. Stainless steel holders for the sterilizing filter have sanitary connections, and the support screens are faced with Teflon to permit autoclaving with the membrane in place. Special compatibility problems may require polyvinyl chloride holders with stainless steel supports, or units that have only Teflon and polypropylene contact parts.

**Table 5.4:** Membrane disc filter sizes for various volumes

| Volumes | Filter |
|---|---|
| 10–100 ml | 13 or 25 mm discs |
| 100–300 ml | 47 mm discs |
| 300–5,000 ml | 90 mm discs |

| | |
|---|---|
| 5,000–10,000 ml | 142 mm discs |
| 20–1,000 L | 293 mm discs |
| 1,000 L and up | cartridges |

Serial filtration is often desired to fractionate the particulates in a fluid. A membrane of large pore size may often be used as a prefilter for a final downstream membrane filter of a smaller pore size.

Pressure drop across the filter media is often observed. This pressure drop may be contributed by either the filter media, holder, or housing. In a properly designed system, the pressure drop due to housing should usually be insignificant except for high-flow liquids or gases.

## Laboratory Filtration Equipment

Laboratory equipment catalogs offer a wide choice of funnels and flasks adaptable to pharmaceutical filtration studies. Although a Buchner funnel test permits analysis of the major difficulties in a filtration problem, development laboratories should have additional procedures and apparatus that produce the qualitative conditions expected in large-scale production. This requirement can be met with a nominal capital investment.

For gravity filtration, conventional glass percolators are applicable, in which case the bottom tube is covered with fibrous material. The filtering funnel is the most common of all laboratory filter devices. Filter paper is used with funnels. Sometimes, a plug of fibrous material may be used instead. Filter bags for laboratory use are made of fabric and are mounted for gravity filtration. The uncertainty of adequate clarification with *glass beads* or *sand* has restricted their use as gravity filters for certain operations in the laboratory.

Suction filters are greatly utilized in the laboratory. Usually, a conical funnel and the Buchner funnel are used for suction filtration, as are immersion and suction-leaf filters. Immersion filter tubes, also known as filter sticks, are generally used for small-scale laboratory operations.

Small-laboratory pressure filters have been used substantially in recent years for both sterile and nonsterile filtration operations. Gravity and suction filters are used mostly for non-sterile filtration. For the pressure filtration of small amounts of material, the filter medium may be mounted in a filter tube, with the liquid poured in and pressure applied to the upper surface of the liquid.

Filter paper in circular form is the most common medium for laboratory filtrations. Filter papers are available in a wide variety of textures, purities, and sizes and are available for different uses. They may be circular (1 to 50 cm in diameter), folded, or arranged in sheets or rolls. Some of the special types of laboratory filter papers for pharmaceutical industry are:

1. Filter papers impregnated with activated carbon for the adsorption of colours and odours in pharmaceutical liquids.
2. Filter paper impregnated with diatomaceous earth for the removal of colloidal haze from liquids with low turbidity.

Minimum laboratory equipment includes a plate and frame press, a membrane filter holder, and a single-element housing for disposable cartridges. A 6- or 8-inch, stainless steel filter press with four to eight filter surfaces and sludge frames is adequate. This covers the flow range from 8 to 200 gallons per hour with minimum filtrate holdup in the press. Stainless steel construction permits autoclaving for sterile operations. Auxiliary equipment for mixing filter aid and feeding the press (10-to 20-gallon tanks, agitators, and centrifugal pump) should also be available. A 90-mm, stainless steel membrane filter holder processes 1 to 15 gallons of sterile solutions per hour. The support plate should be Teflon-lined to permit autoclave sterilization with membranes in place and the gaskets should also be Teflon. Integrity testing apparatus and a stainless steel pressure vessel of 1-to 5-gallons capacity are essential auxiliaries. The same pressure assembly may be used in cartridge filter tests. A broad selection of media should be on hand for each unit.

More flexibility is obtained by adding a metal cartridge filter and a small, manually operated, self-cleaning, edge filter. If processing of high-volume cosmetic products is expected, a single-leaf, precoat pressure filter should be available. Units can be obtained with capacities as low as 1V gallons.

## Centrifugation Filtration and Sedimentation

*Filtering centrifuges and centrifugal sedimentors* are another general class of solids recovery devices. In filtering centrifuges, centrifugal force is used to affect the passage of the liquid through the filter medium. This type of filtration is particularly advantageous when very fine particles are involved. Whenever solids recovery is the primary goal, filtering centrifuges must be considered as an alternative to filtration. In centrifugal sedimentors, the separation is due to the difference in the density of two or more phases. Centrifugal sedimentors are used for complete separation of solid-liquid mixtures and liquid-liquid mixtures. If, however, the separation is incomplete, there will be a gradient in the size of the dispersed phase within the centrifuge due to the faster radial velocity of the larger particles. Operated in this way, the centrifuge becomes a classifier.

## Filtering Centrifuges

The filtration principles discussed previously can be directly applied to filtering centrifuges, although theoretical predictions of spinning time and filtration rate are uncertain. The advantages of the process are effective washing and drying. Residual moisture after centrifugation is far less than in cakes produced by pressure or vacuum filtration. By this method the moisture content of a cake of coarse crystals can be reduced to as low as 3%. This facilitates the drying operation which normally follows. Enclosure of the centrifuge is easy so that toxic and volatile materials can be processed. The process is widely used for separating granular products from liquors, but is less effective for concentrated slurries containing smaller particles.

### *Perforated Basket*

As shown in Fig. 5.14, the device consists of a perforated metal basket mounted on a vertical axis by means of which it can be rotated at a speed of 20 to 25 revolutions per second. The cloth used to retain solids is often supported on a metal screen and the outer casing collects the liquid thrown out from the perforated basket by centrifugal force. Baskets mounted are emptied by shoveling the cake. If, however, top suspension is used, the cake can be more easily withdrawn through traps in the base of the basket. In batch operation, considerable time is lost during machine acceleration and deceleration. Machines operating with continuous discharge of solids are used for separating coarse solids during large-scale operations. Such machines are commonly constructed with a horizontal axis of rotation.

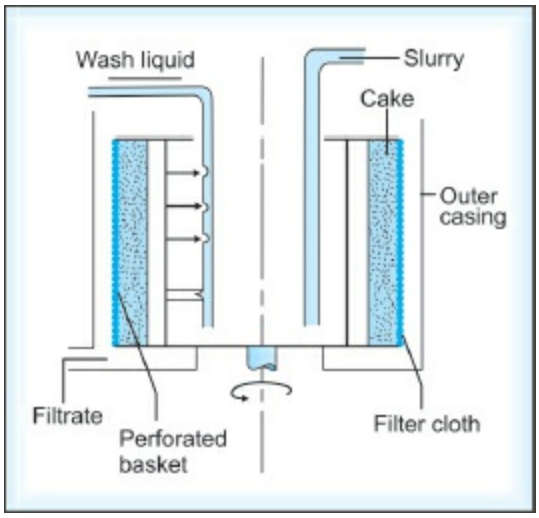


**Fig. 5.14:** Perforated basket type batch centrifugal filter

## Centrifugal Sedimentors

Rate of settling of particles in a liquid is described by Stokes' equation. If particle diameter is *d*, the rate, *u*, at which particle settles by gravity in a liquid of viscosity $\eta$ and density $\rho$ is given by equation:

$$u = \frac{1}{18} d^2 \frac{\rho_p - \rho}{\eta} g \qquad \text{... (8)}$$

where, *g* is the acceleration due to gravity and $\rho_p$ is the particle density.

In the centrifuge the gravitational force causing separation is replaced by a centrifugal force. The centrifugal force of a particle having mass *m*, moving at an angular velocity $\omega$ in a circle of radius *r*, is $\omega^2 r (m\text{-}m_1)$, where $m_1$ is the mass of the displaced liquid. The value of the ratio of the centrifugal and gravitational forces ($\omega^2 r{:}g$) can exceed 10,000. The separation is, therefore, quicker, more complete, and more effective in systems containing very fine particles that will not sediment by gravity because of Brownian movement. Expressing the particle mass in terms of its volume and effective density, we can write the centrifugal force as:

$$F = \frac{\pi}{6} d^2 (\rho_s - \rho) \omega^2 r \qquad \text{... (9)}$$

In streamline conditions the opposing viscous force, given by equation 1.22, is $3\pi d\eta$, *u* being the terminal velocity of the particle. Equating these expressions, we get:

$$u = \frac{1}{18} d^2 \left( \frac{\rho_s - \rho}{\eta} \right) \omega^2 r \qquad \text{... (10)}$$

The sedimentation rate is proportional to the radius of the basket and the square of the speed at which it rotates.

*Tubular bowl centrifuge* is the *Sharpies supercentrifuge* (Fig. 5.15) which operates at upto 15,000 rpm or, in turbine-driven laboratory models, upto 50,000 rpm. For operation at very high speeds, the centrifuge bowl is tubular with a length-to-diameter ratio from 4 to 8. The machine is an effective clarifier when the concentration of solids is very low. It also gives continuous discharge of two separated liquids and is widely used in emulsion separation. Some important uses include cleaning fats and waxes, blood fractionation, and virus recovery.

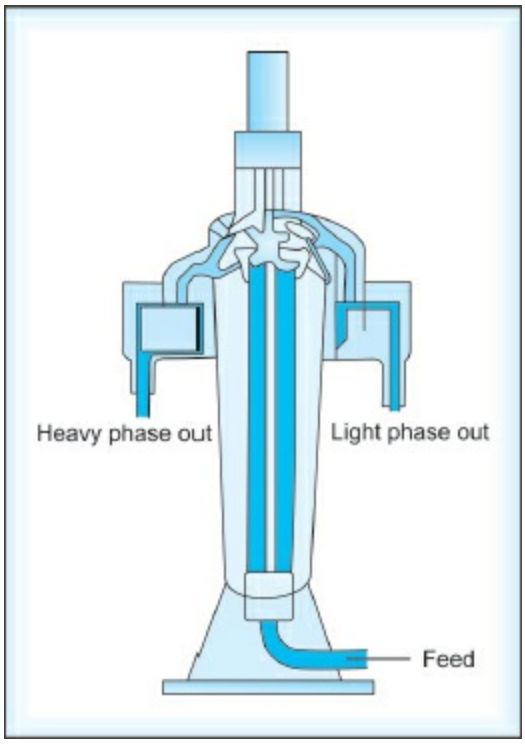



**Fig. 5.15:** Tubular bowl type centrifugal sedimentor

*Conical disk-type centrifuges* introduce a series of conical discs into the bowl. The length-to-diameter ratio is usually much smaller than in tubular bowl centrifuges and operational speeds are lower. The feed enters through a concentric tube surrounding the central drive shaft and flows into the spaces between the discs (Fig. 5.16). As the centrifuge rotates, the heavier liquid or solid moves underside and the lighter liquid moves to the upper side of the discs. Small separating distance, i.e. only the space between the discs, increases the efficiency of the equipment.

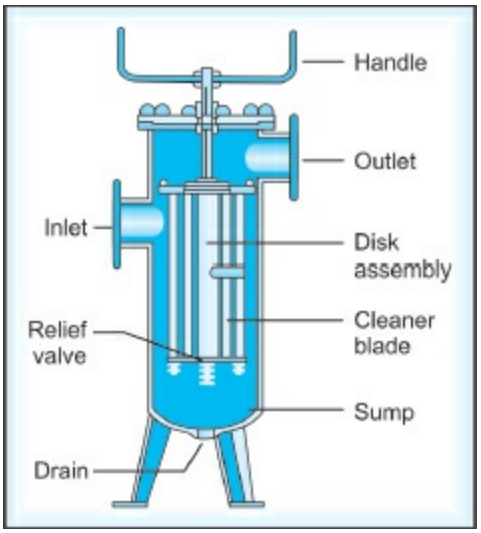



**Fig. 5.16:** Conical disk-type centrifugal sedimentor

*Horizontal continuous centrifuge* are preferred when higher proportions of solids are involved. This type of machine consists of a conical bowl, mounted horizontally (Fig. 5.17). The slurry is introduced through the shaft and liquid separates to the wider portion of the bowl. In continuous models, a conveying scroll, operating at a slightly different speed from the basket, plows the solids to one end and discharges the material as a damp powder. These centrifuges are capable of handling solids with a wide particle size range and in slurries with concentration ranging from 0.5 to 50%.

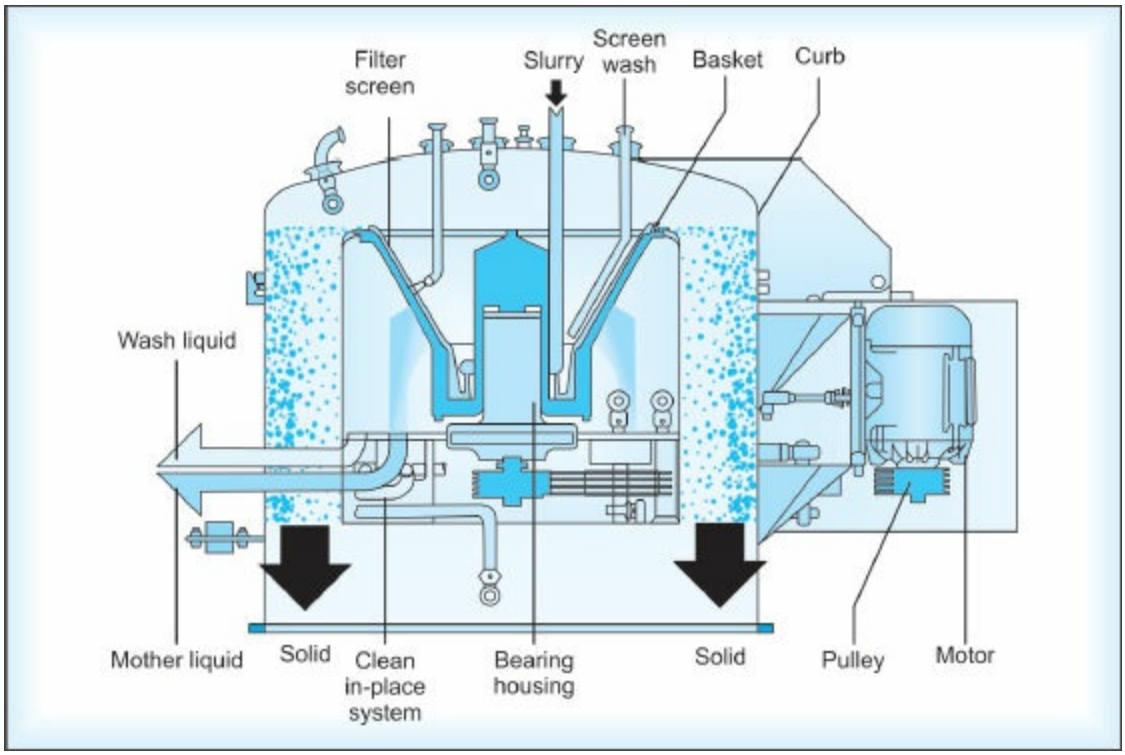



**Fig. 5.17:** Horizontal continuous centrifuge

# SPECIALIZED FILTRATION

## Sterile/Aseptic Operations

Filtration may be used to clarify and sterilize pharmaceutical solutions that are heat-labile. Until the introduction of membrane media, unglazed porcelain candles and asbestos pad were the accepted standards. The candle requires extensive cleaning and is a fragile medium. High flow rates are attained only through the use of multiple-element manifolds. The asbestos pad has significant absorption and adsorption properties, and chemical prewash and pH adjustment are required to prevent interaction with products. Failure to achieve sterility may occur with asbestos pads owing to blow-through and channeling of medium of organisms when critical pressures are exceeded. Both asbestos and porcelain are migratory media as; fragments of a candle or asbestos fibers may be found in the filtrate unless serial filtration through secondary media is used. Since membrane filters do not have these disadvantages, porcelain candles and asbestos pads are no longer considered the media of choice for sterile filtration.

Membrane filters have become the basic tool in the preparation of sterile solutions and have been officially sanctioned by the United States Pharmacopoeia and the USFDA. The available materials permit selection so that absorption effects are negligible and ionic or particulate contamination need not occur. The membrane requires no pretreatment and may be autoclaved or gas-sterilized after assembly in its holder.

A sterility requirement imposes a severe restraint on filter selection. All sterility tests are presumptive, and one must rely upon total confidence in the basic process, and economics becomes a secondary factor. Membranes with porosity ratings of 0.2 or 0.45 μm are usually specified for sterile filtrations. In this porosity range, membrane filters may clog rapidly, and a prefilter is used to remove some colloidal matter to extend the filtration cycle. The FDA allows the use of 0.45 μm filters only in cases of colloidal solutions in which 0.2 μm filters have been shown to clog very rapidly.

Most pharmaceutical liquids are compatible with one or more of the membrane niters now available. High viscosity or abnormal contaminant levels are the primary restraints to the use of membranes, since an extremely large filtration area is needed for practical flow rates. Oil and viscous

aqueous menstruums are therefore heat-sterilized whenever possible. These solutions are usually clarified through coarser, nonsterilizing membranes, preferably prior to heat sterilization. Paraffin oils, however, may be successfully filtered through 0.2 micron membranes after heating to reduce viscosity.

Simple formulations such as intravenous solutions, ophthalmics, and other aqueous products may be filtered directly through membranes in an economical manner. Heat-labile oils and liquids containing proteins require pretreatment, e.g. centrifugation or conventional filtration, prior to sterilization filtration. The objective is removal of gross contamination that would rapidly plug the finer membranes. Difficult materials, such as blood fractions, demand serial filtration through successively finer membranes. The cost of multiple filtration may seem excessive, but it is often the only way to achieve sterility.

In selecting a filtration system for sterilization of any growth-supporting medium, the following precautions must be kept in mind:

1. Identify the potential sources of adverse biochemical and chemical contamination at each point of the system.
2. Identify the control points necessary to eliminate possible contamination and decrease cost.
3. Identify the hazards associated with each control point, i.e. airborne contamination and protein denaturation.
4. Establish a protocol for monitoring the hazards at control points of the system.

Figure 5.18 illustrates the basic filtration system for non-sterile filtration of serum, water, and salts to reduce the microbiologic and particulate matter, followed by final filtration through the sterile membrane.

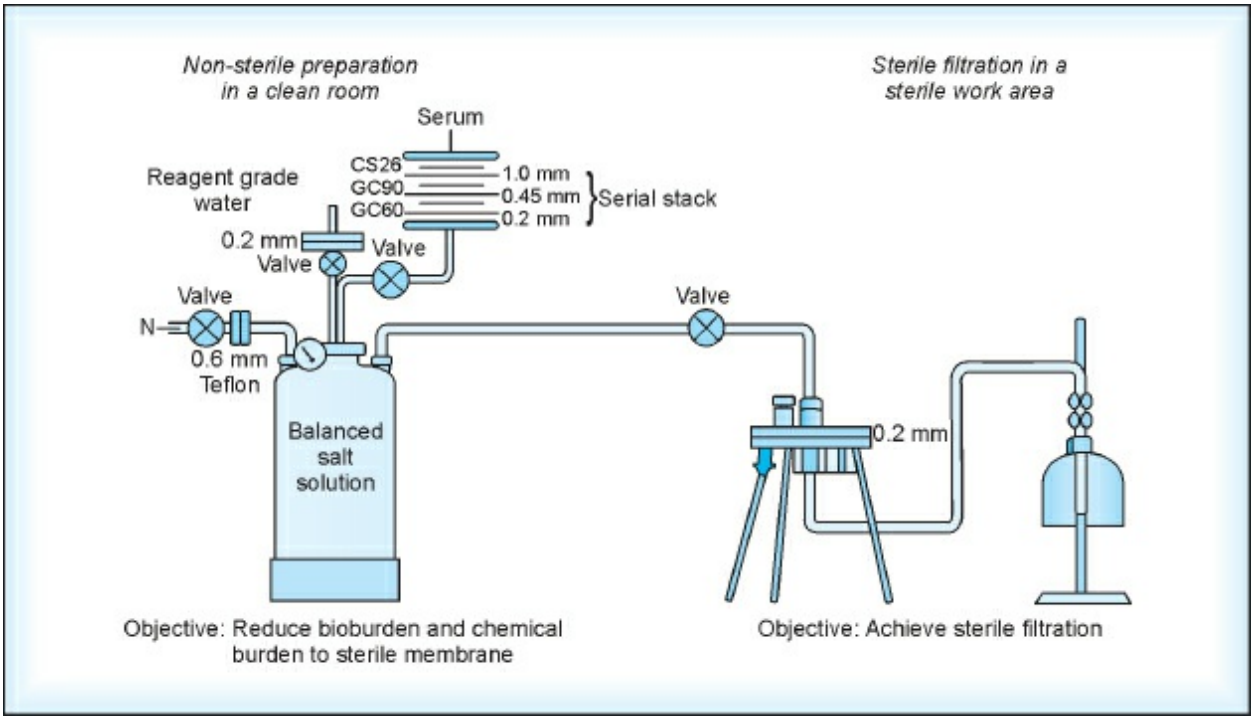



**Fig. 5.18:** Schematic representation of operational sequence

The use of filtration to remove bacteria, particulate matter from air, and other gases such as nitrogen and carbon dioxide is widespread in the pharmaceutical industry. The following are some common applications employing initial gas filtration:

- Vent filtration
- Compressed air used in sterilizers
- Air or nitrogen used for product and inprocess solution transfers and at filling lines
- Air or nitrogen used in fermentation.

When sterile and under ideal conditions, traditionally packed fiberglass or cotton filters provide vent protection. The use of hydrophobic membrane filters is increasing. These filters guarantee bacterial removal in wet and dry air and do not channel, unload, or migrate the medium. These filters may need to be heated by jacketing. Restrictions of airflow through the vent filter can result in pump damage or tank collapse.

Manufacturers of membrane filters provide extensive application data and detailed directions for assembly, sterilization, and use of their filters. The basic elements of any sterile operation must be followed. All apparatus

should be cleaned and sterilized as a unit. Filtration should be the last step in processing, and the filter should be placed as close as possible to the point of use of final packaging. In serial filtrations, only the final unit needs to be sterile, but minimal contamination in prior steps increases the reliability of the total process. Sterile filtration should always be a pressure operation; a vacuum is undesirable since bacteria may be drawn in at leaky joints and contaminate the product.

After the successful introduction of a new filtration process, manufacturing tolerances allow reasonable changes in flow rate so long as quality is met. Therefore, the most common production problem is complete plugging of filter media resulting in no productivity. Subtle changes in raw material quality are often at fault. The level of an impurity needs change only slightly to create problems with the fine porosity media used in polishing operations. For example, iron contamination in an alkaline product can lead to colloidal precipitates, which blind the media. Raw material problems should always be suspected when synthesis procedures have been altered or when the vendor of a purchased commodity has changed.

## Membrane Ultrafiltration

Membrane ultrafiltration has become a commercially feasible unit operation in the past decade. Unlike conventional filtration, ultrafiltration is a process of selective molecular separation. It is defined as a process of removing dissolved molecules on the basis of membrane size and configuration by passing a solution under pressure through a very fine filter. Ultrafiltration membrane retains most macromolecules while allowing smaller molecules and solvent to pass through the membrane, even though the membrane is not rated as absolute. The difference between microfiltration and ultrafiltration is significant. The former removes particulates and bacteria; the latter separates molecules. Application of hydraulic pressure reverses the normal process of osmosis, so that the membrane acts as a molecular screen through which only those molecules below a certain size are allowed to pass.

Separation of a solvent and a solute of different molecular size may be achieved by selecting a membrane that allows the solvent, but not the solute, to pass through. Alternatively, two solutes of different molecular sizes may be separated by choosing a membrane that allows the smaller molecule, to pass through but holds back the larger one (Fig. 5.19). Ultrafiltration is similar in process to reverse osmosis as both filter on the basis of molecular size. Ultrafiltration is different from reverse osmosis in the sense that it does not separate on the basis of ionic rejection. Dialysis and ultrafiltration are similar in the sense that both processes separate molecules, but ultrafiltration is different in that it does involve the application of pressure.

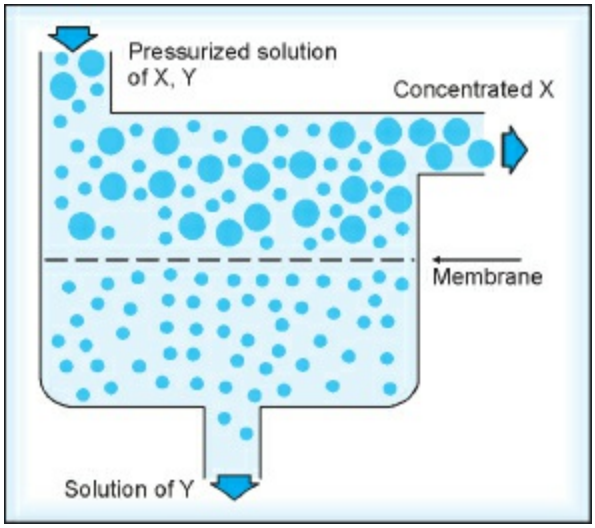


**Fig. 5.19:** Schematic diagram of membrane ultrafiltration process

The selectivity and retentivity of a membrane are characterized by its molecular weight cut off. It is difficult to characterize the porosity of an ultrafiltration membrane by means of precise molecular weight cut off. The configuration of the molecule and its electrical charge may also affect the separation properties of the membrane. Ultrafiltration membranes are therefore rated on the basis of nominal molecular weight cut off. The shape of the molecule to be retained plays a major role in retentivity. Many of the same techniques that are used in microfiltration to increase flow rate and throughput are also used for ultrafiltration. Ultrafiltration membranes are available as flat sheets, pleated cartridges, or hollow fibers. The hollow fibers have the selective skin on the inside of the fiber.

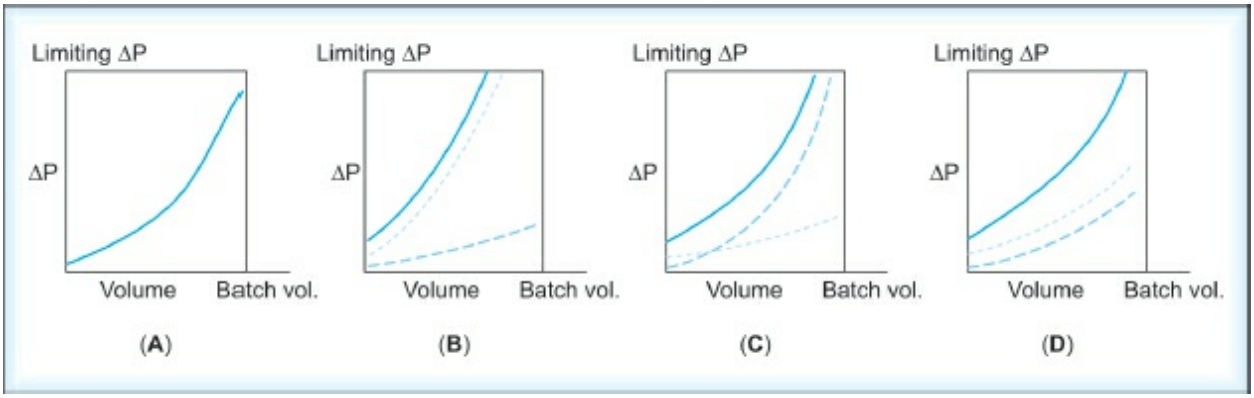
Industrial use of this procedure has followed the development of anisotropic polymer membranes in a variety of biologically inert, noncellulosic materials. These membranes are fragile structures, however, and usually require a backing plate of porous material to withstand operational pressure. During the processing of a solution, a region of high solute concentration also develops at the surface of the membrane, resisting further passage of solvent. Providing essential support for the membrane and overcoming concentration polarization through shear effects have resulted in a wide variety of commercial apparatus, including tangential-flow cassette systems, process ultrafiltration cartridges, hollow fiber beakers, and collodion bags. Since the technology continues to change rapidly, reliance on technical expertise of the manufacturer is advisable.

Applications in the pharmaceutical industry are predominantly in the concentration of heatlabile products, such as vaccines, virus preparations, and immunoglobulins. Ultrafiltration also has been used to recover antibiotics, hormones, or vitamins from fermentation broths, to separate cells from fermentation broth, to clarify solutions, and to remove low-molecular-weight contaminants prior to using conventional recovery techniques. The most important application of ultrafiltration is the removal of pyrogens.

# FILTRATION PROCESS OPTIMIZATION

## Prefilter Optimization

The optimum system often requires the use of a series of filters in a single multilayered filter containing layers of various pore sizes or a prefilter followed by a final filter. Optimum performance is obtained when the filters in a series exhaust their dirt-holding capacities at the same time. When the flow resistance across each filter in the series approaches the limiting pressure drop, the dirt-holding capacity of the system is considered expended. Figures 5.20A to D illustrate the prefilters with adequate and inadequate dirt holding capacity. In Fig. 5.20, the coarse prefilter does not provide sufficient retention efficiency, thus causing the poorly protected final filter to clog prematurely. Too fine a filter, on the other hand, has enough retention efficiency but insufficient dirtholding capacity, and plugs very quickly, as illustrated in Fig. 5.20C. As shown in Fig. 5.20D, both filters—the final filter and the "correct" prefilter—will have almost expended their dirt-holding capacities as the last of the batch is filtered. A final filter that is not protected by prefilter has a short filter life. When a prefilter is used in combination with a final filter, the efficiency of the prefilter is maximum. In these cases, it is important that the O-ring seal sits directly on the membrane itself and not on the prefilter. Therefore, the diameter of the disc prefilter selected should be somewhat smaller than the diameter of the final filter.




**Figs 5.20A to D:** Filtration system with: (A) Ideal properties; (B) Inadequate prefilter—too coarse; (C) Adequate retentive prefilter but inadequate dirt-holding capacity; (D) Adequate prefilter

Table 5.5 lists the diameter of the filter and the diameter of the prefilter when used in combination. Seating the O-ring on the prefilter often fails to produce a seal, thus causing the filtration system to leak. This leakage may result in the filtrate being exposed to contamination.

**Table 5.5:** Diameter of filter and corresponding prefilter when used in combination

| Filter size (mm) | Prefilter size (mm) |
|---|---|
| 25 | 22 |
| 47 | 35 |
| 90 | 75 |
| 142 | 124 |
| 293 | 257 |

## Filter Media Optimization

Although filtration analysis can be sophisticated, pilot plant studies are usually basic. The common problems are to select the media, determine the time required, and if possible, estimate when a semicontinuous cycle should be terminated for cleaning.

For nonsterile polish filtrations, the quality level must be established prior to choice of media. Particulate matter above 30 to 40 μm particles may be noticeable. Most pharmaceutical filtrations therefore aim for removal of particles of 3 to 5 μm or less. A nephelometer, an instrument that measures the degree of light scattering (Tyndall effect) in dilute suspensions, is an excellent tool for assessing effectiveness in this range.

The nephelometer gives a quantitative value to the formulator's quality specification of "sparkling clear." This value may be used to compare results using different filtration media. Figure 5.21 shows a typical curve obtained from filtration of an elixir through disposable cartridges and standard kraft paper. If an existing process is to be shifted from paper on a filter press to cartridges, this curve permits selection of an element that gives comparable performance. The technique also may be applied to the assessment of filter aid effectiveness by determining transmittance as a function of filter-aid type, quantity, or method of use.

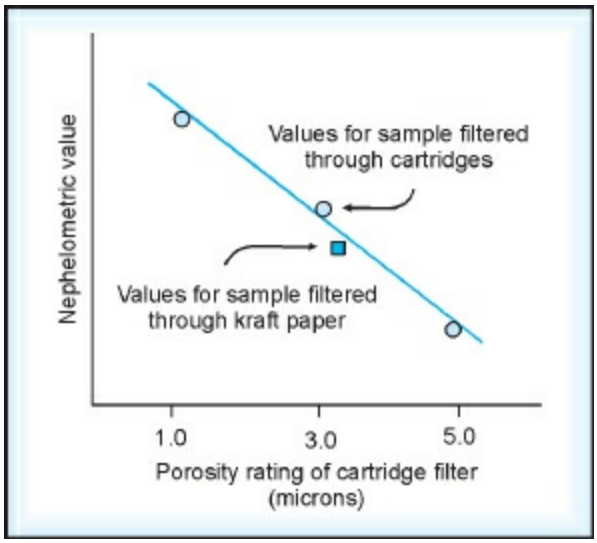



**Fig. 5.21:** A nephelometric readings of a filtrate provides data that may be used to compare performance of different media

## Filter Aid Optimization

In addition to improving clarity, filter aids are used to increase flow rates. Figure 5.2 indicates a typical flow rate pattern as the amount of filter aid is increased. Exceeding an optimum quantity can frequently lead to decreased flow rate without improving clarity. The filter aid quantity can be expressed as a percentage of cake solids, a percentage of filter aid in body mix, or the weight applied as a precoat per unit of filter area. Flow rate should be determined for each case at constant pressure and after a uniform time interval. The maximum filter aid level used in laboratory tests must be within the cake capacity of the projected or existing plant equipment.

## Filtration Cycle Optimization

The question of time for a filtration cycle is resolved by determining the total volume versus time during a test run at pressures approximating normal operating conditions. Flow rate decreases with time as the media plugs or as the cake builds up. Plotting log total volume per unit area versus log time usually gives a straight line suitable for limited extrapolation (Fig. 5.22). If the filter area of production equipment is fixed, the time to filter a given batch size may be estimated. Alternately, the filter area required to complete the process within an allotted time period may be established. Similar flow decay studies can also be performed during sizing of a filtration system for sterile operations.

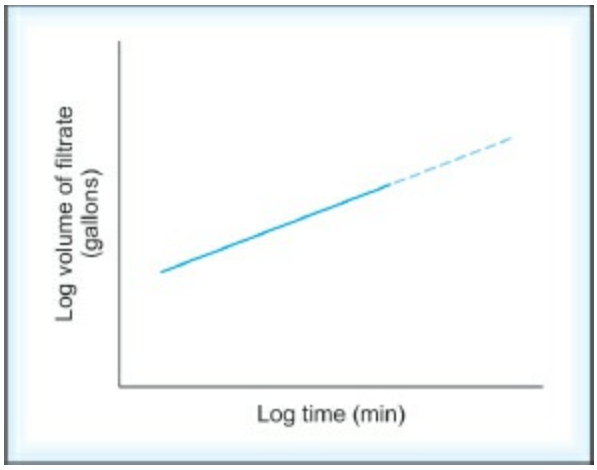



**Fig. 5.22:** Extrapolation of filtrate volume produced in a given time can be made from log-log plots of experimental data

In semicontinuous operations, decisions must be made on length of the cycle prior to shutdown for replacement of media. If the goal is maximum output from the filter per unit of overall time, the graphic approach of Fig. 5.23 is applicable. During productive time, T, the filter discharges a clear filtrate at a steadily decreasing rate. Nonproductive time, *T*, is required to clean the filter and replace media. For graphic analysis, nonproductive time, *T*, is plotted to the left of the origin of volume, *V*, versus time curve. When a line is drawn from *T* tangent to the curve, the value of *V* and *T* at the point of tangency indicates where the filtration should be stopped. The time lost in cleaning is offset by a return to high filtration rates associated with the new media. This point also can be calculated from theoretic relationships for

constant pressure or constant volume filtration. Data from laboratory equipment can be applied to production units since the analysis is independent of filter area.

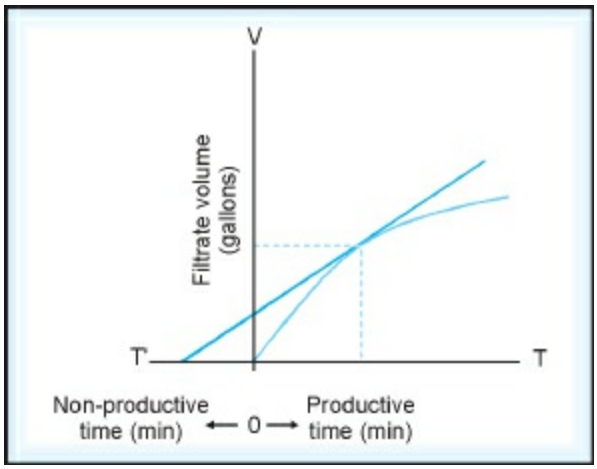



**Fig. 5.23:** The optimal filtration cycle prior to cleaning can be determined by a graphic technique

## Straining Operation Optimization

The evaluation of coarse straining operations is limited to sizing a filter that will not have excessive pressure drop. The amount of impurity is usually small, and continued operation does not significantly decrease filter capacity. The metal cartridge filters, either woven-mesh or edge-type, and porous sintered stainless steel, have replaced cheesecloth in most pharmaceutical applications. Straining suspensions containing gums or other viscous ingredients can be accomplished with self-cleaning edge filters. These suspensions frequently bridge the media, and cleaning devices are needed to maintain adequate flow.

## FILTER SELECTION

In designing or selecting a system for filtration, the specific requirements of the filtration problem must be defined. The following questions should be answered before any assistance is requested from the manufacturers of filtration equipment:

1. What is to be filtered—liquid or gas?
2. Which liquid or gas is to be filtered?
3. What is the pore size required to remove the smallest particle?
4. What is the desired flow rate?
5. What will the operating pressure be?
6. What are the inlet and outlet plumbing connections?
7. What is the operating temperature?
8. Can the liquid to be filtered withstand the special temperature required?
9. What is the intended process—clarification or filtration?
10. Will the process be a sterilizing filtration?
11. Will the process be a continuous or batch filtration?
12. What is the volume to be filtered?
13. What time constraints will be imposed, if any?

Once the purpose of the process has been determined, the selection of the filter medium can be made. For example, for a sterilizing filtration, a 0.2 μm pore size is used; for clarification, a plate and frame filter or woven-fiber filter may be used. In general, a pore size smaller than the smallest particle to be removed is selected. The filter medium should be compatible with the liquid or gas to be filtered. It is advisable to check the chemical compatibility charts provided by the vendors for selection of filter type. Filter type, cellulose, polytetrafluoroethylene (PTFE), fiber, metal, polyvinylidene difluoride, nylon, or poly-sulfones may be selected based on the chemical resistance to the most aggressive ingredient in the liquid. For vent filters or gaseous filtration, a hydrophobic filter medium should be chosen.

Filtration surface area is calculated after the filter media, pore size, required flow rate, and pressure differentials are established. For a liquid

having a viscosity significantly different from that of water (1 cp), the clean water flow rate is divided by the viscosity of the liquid in centipoises to obtain the approximate initial flow rate for the liquid in question. For gaseous filtration at elevated temperature and exit pressures, the standard flow rate (20°C, 1 atmosphere) must be corrected by equation (11), the *gaseous filtration flow rate formula:*

$$F = F_0\left(\frac{293}{273+t}\right)\left(\frac{P+\Delta P/2}{14.7+\Delta P/2}\right) \qquad \ldots (11)$$

where, $F$ = corrected flow rate
$F_0$ = standard flow rate from chart (20°C, 1 atmosphere)
$t$ = temperature of air or gas (°C)
$P$ = exit pressure (psi)
$\Delta P$ = pressure drop through the system (psi)

If the pressures are expressed in kg/cm$^2$, the term 14.7 in equation (11) becomes 1.03.

The broad span of pharmaceutical requirements cannot be met by a single type of filter. The industrial pharmacist must achieve a balance between filter media and equipment capabilities, slurry characteristics, and quality specifications for the final product. The choice is usually a batch pressure filter, which uses either surface or depth filtration principles.

# 6:
# Pharmaceutical Rheology

Pharmaceutical fluid preparations are recognized as materials that pour and flow, and have no ability to retain their original shape when not confined. The semisolids are a more nebulous grouping. They essentially retain their shape when unconfined but flow or deform when an external force is applied. Those materials that readily pour from bottles and form a puddle are clearly fluids. Ointments or pastes, that clearly retain their shape after extrusion from a tube, characteristically are associated with pharmaceutical semisolids. Obviously, a continuum of properties exists between these limits.

*Rheology* (from the Greek words *rheos* meaning flow and *logos* meaning science) is the study of the flow or deformation of matter under the influence of stress. Rheology can be applied to solids (completely-resistant to deformation), liquids (moderately-resistant), and gases (completely non-resistant). In pharmaceutical and allied research and technology, rheologic measurements are involved in (i) pharmaceutical processing operations such as mixing of materials, (ii) filling and (iii) packaging into containers, and thus affect the choice of processing equipment to be used in manufacture. Rheology is also involved in the (iv) removal of product from package, whether by pouring from a bottle, extrusion from a tube or other deformable container, spraying liquids from atomizers or aerosol cans, or passage through a syringe needle. (v) Rubbing the product onto and into the skin is controlled by its rheological properties. It affects (vi) a product's physical stability, and is of extreme importance in both product development and quality assurance to ensure that the desired attributes of body and flow are retained for the required shelf-life of the product. It influences (vii) the dose uniformity and (viii) bioavailability of a product as viscosity has been shown to affect the absorption rate of drugs from the gastrointestinal (GI) tract. (ix) Release of drug from dosage forms and delivery systems is controlled by the rheological characteristics of the formulation matrix. The scope of biorheology is concerned with (x) the circulation of blood and lymph through capillaries and large vessels, (xi) flow of mucus in the respiratory tract, (xii)

bending of joints, (xiii) contraction of muscles and (xiv) transit of luminal contents through the GI tract.

# FUNDAMENTAL CONCEPTS

## Elastic Deformation and Viscous Flow

When force is applied to a body, its deformation can be described by two components viz. elasticity or viscosity. Pure *elasticity* is achieved if the shape of body is restored once the force is withdrawn. Elasticity is the property of solid materials. Hooke's law of elasticity is used to describe the elastic deformation of solids. According to Hooke's law, in the ideal case of a Hookean body, stress is directly proportional to strain, and the body returns to its original shape and size, after the applied stress has been relieved. The proportionality between stress and strain is quantified by the constant known as the modulus of elasticity (units: Pascals) or Young's modulus ($E$).

$$dl = \sigma/E \ldots (1)$$

where, $\sigma$ is the applied stress and $dl$ is the elastic deformation or extension in length caused by the application of stress.

On the other hand, pure *viscosity* or pure viscous flow occurs if there is continuous movement during the applied force, and no restorative motion follows after the removal of deforming force. It is a permanent or irreversible deformation. Viscosity is the property of liquid materials and is explained by Newton's law of viscous flow.

## Newton's Law of Viscous Flow

To best understand the fundamental components of viscous flow, consider Fig. 6.1. Two parallel planes are a distance *dx* apart. The viscous body is confined between the planes. The top, plane A, moves horizontally with a velocity *dv* because of the action of force F. The lower plane B is motionless. As a consequence, there exists a velocity gradient *dv/dx* between the planes. This gradient in velocity over a distance is known as the *rate of shear*, *D (dv/dx)*. The horizontal force per unit area

$$F/A = \alpha\, dv/dx \ldots (2)$$
$$F/A = \eta\, dv/dx \ldots (3)$$
$$S = \eta D \ldots (4)$$

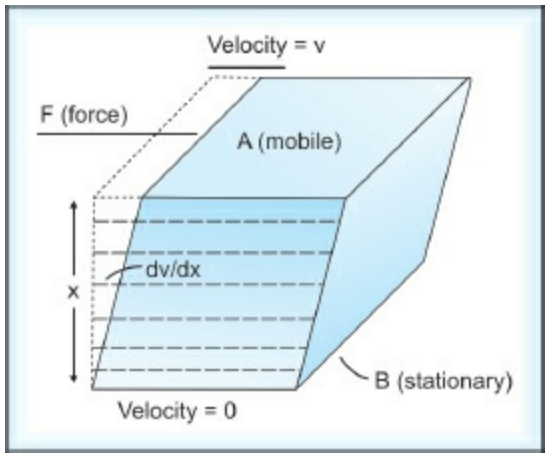



**Fig. 6.1:** Model to demonstrate components of classic viscous flow

where, $\eta$ is the constant of proportionality, known as viscosity or coefficient of viscosity.

*Viscosity* is the internal friction in the fluid, i.e. resistance to the relative motion of the adjacent layers of a liquid. Conventionally, viscosity is represented by $\eta$. Then rearranging Eq. 4, we get:

$$\eta = S/D \ldots (5)$$

Viscosity is defined as the tangential force per unit area, in dynes/cm$^2$, required to maintain a velocity difference of 1 cm/s between two parallel layers of liquid which are 1 cm apart. In the centimeter-gram-second (C.G.S.) system grams per centimeter per second (g cm$^{-1}$ s$^{-1}$) is the dimensional unit

of viscosity. In these units, viscosity is expressed in *poise*, a term used in recognition of the pioneering work of the in the 1840s French scientist, J. L. M. Poiseuille. Thus is the basic unit of viscosity for dilute aqueous solutions, the common unit becomes the centipoise ($10^{-2}$ poise), cp. The viscosity of water is about 1 cp.

The unit of viscosity can be derived as follows:

$$\eta = S/D \quad \ldots (5)$$

Shear stress = $F/A$ = dynes/cm$^2$

Rate of shear, $dv/dx$ = (cm/s)/cm = s$^{-1}$

Therefore, the units of viscosity in C.G.S. system are:

Units of viscosity = (dynes/cm$^2$)s

Since dynes = g cm/s$^2$,

Units of viscosity = {[g cm/s$^2$]s/cm$^2$}

= g cm$^{-1}$ s$^{-1}$ = poise

In the International System of Units (SI system), pascal (Pa) is the unit of stress and has the dimensions of newton/meter$^2$, where the newton is can be expressed as kilogram meter/second$^2$. Hence, the SI unit of viscosity is *pascal.second*. One pascal.second is equal to 10 poise.

*Example 1:* If an oil is rubbed on to the skin with a relative rate of motion between the two surfaces of 15 cm, the film thickness is 0.01 cm, and the oil has the same viscosity as water, then the shear rate and shearing stress can be determined as follows:

$$D = \frac{15}{0.01}\frac{\text{cm/s}}{\text{cm}} = 1500\ \text{s}^{-1}$$

$$\eta = S/D$$

$$1\times 10^{-2}\,\text{poise} = \frac{S}{1500}\,\text{s}^{-1}$$

$$\text{Then, } S = (1500)(1 \times 10^{-2})\ (\text{s}^{-1})\ (\text{poise})$$

$$= 15\ (\text{s}^{1})\ (\text{dyne.s.cm}^{-2})$$

$$= 15\ \text{dyne cm}^{-2}$$

*Example 2:* In S.I. units, the above terms would become:

$$\eta = 1 \text{ mPas}$$
$$D = 1500 \text{ s}^{-1}$$
$$S = 1.5 \text{ Pa}$$

The reciprocal of the viscosity, and is usually designated by the symbol is *fluidity* $\phi$

$$\phi = 1/\eta \ldots (6)$$

*Kinematic viscosity* ($\upsilon$) is the Newtonian viscosity (or absolute viscosity) divided by the density of a liquid at a particular temperature.

$$\upsilon = \eta/\rho \ldots (7)$$

The units of kinematic viscosity are *stoke (s)* and *centistoke (cs)*, in honor of the English scientist who studied problems of gravitational settlement in fluids. Certain fluid flow viscometers give values in the kinematic viscosity scale.

*Example 3:* If the oil from example 1 had a density of 0.82 g/cm$^3$, then the kinematic viscosity would be:

$$v = \frac{\eta}{d} = \frac{1 \times 10^{-2}}{0.82} = 1.22 \times 10^{-2} \text{ stokes}$$
$$= 1.22 \text{ centistokes}$$

## *Temperature Dependence of Viscosity*

Viscosity of liquids falls with a rise in temperature, while that of gases rise with an increase in temperature. The fall in viscosity the in case of liquids is due to a decrease in the intramolecular forces of attraction. The variation of viscosity with temperature is expressed by an equation analogous to the Arrhenius equation of chemical kinetics:

$$\eta = Ae^{Ev/RT} \ldots (8)$$

where, $A$ is a constant depending on the molecular weight and molar volume of the liquid, $E_v$ is the activation energy necessary to initiate flow between the molecules and $R$ is the gas constant (1.9872 cal mol$^{-1}$ K$^{-1}$).

Based on Newton's law of viscous flow, fluids are classified as Newtonian and nonNewtonian. Newtonian fluids are those which follow

Newton’s law of viscous flow, whereas non-Newtonian fluids do not follow it. The classification of fluids based on their rheological behaviour is shown in Fig. 6.2.

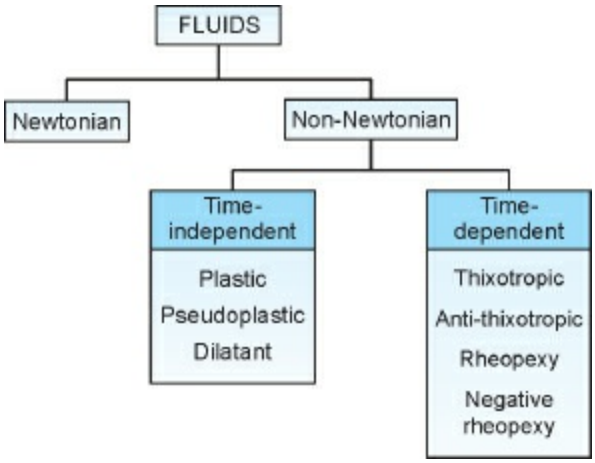



**Fig. 6.2:** Classification of fluids based on rheological behaviour

## NEWTONIAN FLUIDS

Simple liquids, either pure chemicals or solutions of lower molecular weight compounds, are Newtonian fluids, which possess a direct proportionality between shear stress and shear rate for all values of shear. Viscosity of such fluids depends on composition, pressure and temperature, but is independent of the rate of shear. The absolute viscosities of some Newtonian liquids of pharmaceutical interest are given in Table 6.1.

**Table 6.1:** Absolute viscosity of Newtonian liquids at 20°C

| **Liquid** | **Absolute viscosity (at 20°C), cps** |
|---|---|
| Ethyl ether | 0.24 |
| Acetone | 0.34 |
| Chloroform | 0.563 |
| Water | 1.0019 |
| Absolute ethanol | 1.20 |
| Ethanol, 40% (w/w) | 2.91 |
| Olive oil | 100.00 |
| Glycerin, 95% (w/w) | 545.00 |
| Castor oil | 1000.00 |
| Glycerin, anhydrous | 1500.00 |

A plot of rate of shear as a function of shear stress is known as flow curve or *rheogram* and when the rate of shear is plotted as a function of viscosity, the curve is known as a *viscogram*.

For Newtonian fluids, the rheogram is linear and passes through the origin indicating that minimal stress applied will induce shear (Fig. 6.3). The slope of such a curve is known as *fluidity* and the inverse of slope is the *viscosity* of the fluid. Since by definition, shear stress and shear rate are directly proportional, a single viscometric point can characterize the liquid rheology. Viscogram for Newtonian fluids is a straight line parallel to the axis of the rate of shear indicating that Newtonian viscosity is independent of

the rate of shear.

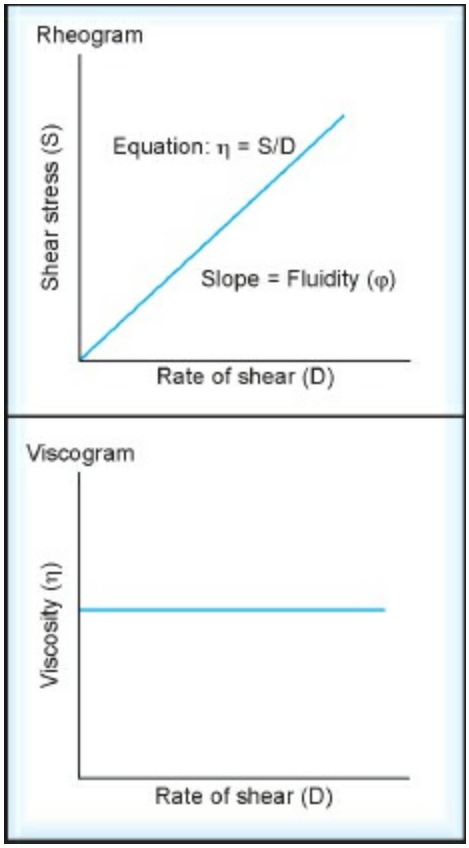



**Fig. 6.3:** Rheogram and viscogram for Newtonian fluids

## NON-NEWTONIAN FLUIDS

Non-Newtonian fluids are those which possess no direct linear relationship between shear stress and shear rate, i.e. they do not follow Newton's law of flow. Most systems of pharmaceutical interest fall into this category. The rheological behaviour of nonNewtonian fluids may be characterized either as *time-independent* or *time dependent* nonNewtonian fluids.

## Time-independent Non-Newtonian Fluids

In these kinds of fluids, the system instantaneously adapts itself to changing shear stress. Time-independent non-Newtonian fluid behaviour can be of three types-plasticity, pseudoplasticity and dilatancy.

### *Plasticity*

Concentrated flocculated suspensions (e.g. concentrated zinc oxide suspension) and semisolid dosage forms, such as gels, creams, and ointments, require an initial finite force, before any rheological flow can start. This initial stress is called *yield value* and such substances are known as plastics or Bingham bodies. At shear stress values below the yield value, such substances behave as elastic solids exhibiting reversible deformation, and above the yield value, they behave as Newtonian systems.

The rheogram of a *Bingham plastic* is represented by a straight line or curve on the stress-shear rate plot, being displaced from the origin by the yield value (Fig. 6.4). The slope of the linear portion is known as the mobility, which is the inverse of plastic viscosity.

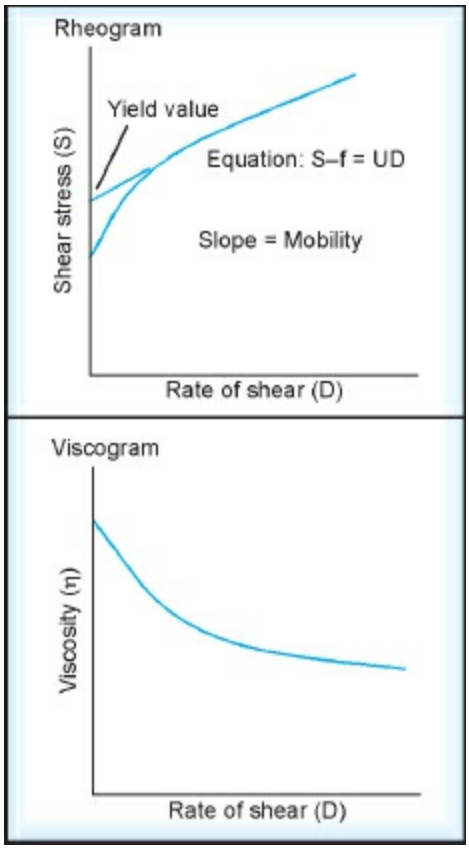



**Fig. 6.4:** Rheogram and viscogram for plastic flow

Thus, for Newtonian behaviour at stresses (*S*) greater than the yield value (*f*) we have:

$$S - f = UD \quad \ldots (9)$$

where, U is the plastic viscosity and *D* is the shear rate. The dimensional units of the yield value must be those of the shear stress. Plastic viscosity is thus defined as the shearing stress in excess of the yield value which has to be applied to induce a unit rate of shear.

Plasticity is often exhibited by concentrated flocculated suspension, where the particles are attracted by van der Waals forces (force of flocculation) giving structure to the system. The shear forces required to break the force of flocculation between the particles contributes to the yield value. The nature of interlocking or interweaving in the structure dictates whether initial flow occurs with difficulty until sufficient structure is lost or whether a sufficient initial force is required to initiate motion, i.e. whether a *yield value* has to be exceeded. In any case, continued shearing breaks further

linkages, so that the apparent viscosity drops with increasing shear. As the yield value is exceeded, the shearing stress and rate of shear become directly proportional. The diagrammatic explanation of plastic behaviour is depicted in Fig. 6.5.

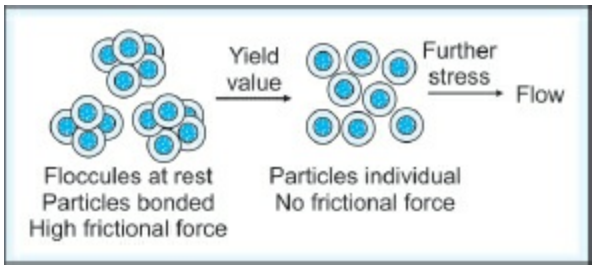



**Fig. 6.5:** Explanation of plastic behavior

Occasionally, semisolid systems, particularly those containing an appreciable solid content or those in which structure is developed through a three-dimensional polymer (silicate system), develop shear planes across which virtually no interactions exist, and hence, flow is easy. These planes of slippage result in flow curves that deviate from straight line at stresses close to the yield value. This phenomenon is called *plug flow* because the material moves in chunks or as a plug rather than by laminar motion. As is discussed later, this results in confusing rheological implications for product performance. In general, such measurements are less reproducible.

## *Pseudoplasticity*

Many polymeric solutions, especially solution of tragacanth, sodium alginate and cellulose derivatives, weakly flocculated suspensions, and semisolid systems containing polymer component, become more fluid the faster they are stirred. This shear-thinning behaviour is often referred to as pseudoplasticity.

The rheogram of pseudoplastic material begins at the origin indicating that the particle-particle bonds are too weak to withstand the applied shear stresses. For pseudoplastic material, the increase in the rate of shear is greater than the corresponding increment in shear stress and the rheogram is concave towards the shear-rate axis (Fig. 6.6). Thus, if the viscosity is calculated at each of a series of shear rate points, then the resultant values decrease with increasing shear rate. The fact that one value of viscosity cannot characterize the viscous behaviour of pseudoplastic materials on the account of their non-

Newtonian flow requires the use of some equation of state. One such equations is the empiric *power law* or *Ostwald-de Waele* equation which is applicable over a wide range of stresses:

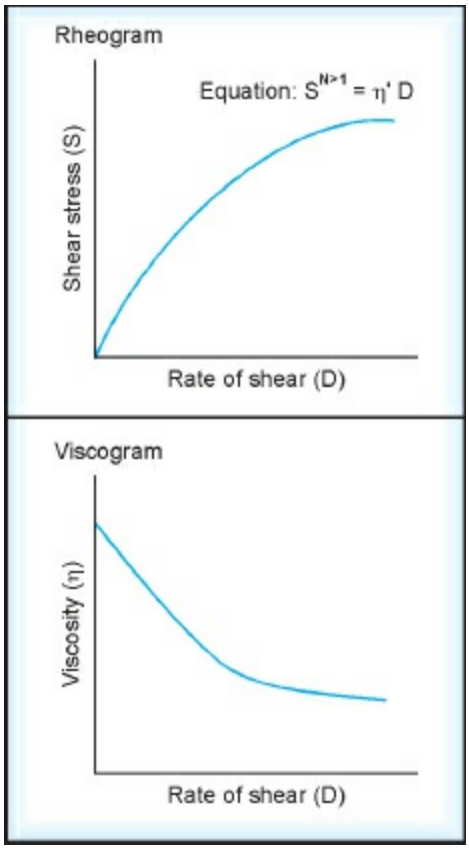



**Fig. 6.6:** Rheogram and viscogram for pseudoplastic flow

$$S^N = \eta' D \ldots (10)$$

where, $S$ and $D$ are the shear stress and shear rate, respectively, $\eta'$ is the apparent viscosity and $N$ is the power index of deviation from Newton's law. In this equation, $N$ is greater than 1 for pseudoplastic materials and less than 1 for dilatant materials. The equation is reduced to Newton's law when $N$ is equal to 1.

When the logarithm of both sides of the equation is taken, the result is:

$$\log D = N \log S - \log \eta' \ldots (11)$$

This is equation for a straight line when log $D$ is plotted as a function of log $S$.

In polymeric solutions, the flexible, long straight-chain macromolecules 

are buffeted constantly by the surrounding water molecules in thermal agitation. In order to attain condition of minimum energy, the macromolecules tend to have some coiling. These coils can develop a degree of interlocking. With branched polymers, the opportunity for frictional interlock the is obviously even greater. In addition, or alternatively, intramolecular bridging may occur by simple hydrogen bonding. This bonding may create innumerable bridges between individual adjacent molecules to create complete cross-linking that is permeated by water. Upon the application of shear, a unidirectional laminar motion is superimposed and the macromolecule chains untangle and align themselves in the direction of flow, as shown in Fig. 6.7.

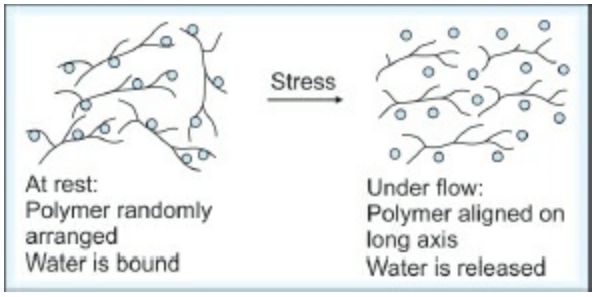



**Fig. 6.7:** Explanation of pseudoplastic behaviour

The imposition of increasing shear rates enables the macromolecule chains to uncoil and reduces the entrapment of water, thereby offering less resistance to flow, and a reduction in viscosity. On the removal of shear stresses, Brownian motion reestablishes the coiled conformation and interparticle links instantaneously, and the system is restored to its high viscosity condition. Thus, the restoration is time-independent.

Pseudoplastic flow is exhibited by weakly flocculated suspensions, in a secondary minimum, which produce a house of card structure. In the case of silica or alumina gel, water may serve alone or with other agents from cross-linking to create a three-dimensional polymer structure. Classically, most suspending agents exhibit similar capability for the development of structure. Depending on the nature of the suspending and crosslinking materials used, these adjuvants may, or may not be in complete homogeneous solution. Instead, they may have a strong affinity for the solvent and yet be partially insoluble. In systems with dispersed solids, these colloid solutions can serve as the interweaving material to hold the whole system together.

## *Dilatancy*

*A dilatant* material is the one in which the increase in the rate of shear is greater than the corresponding increment in shear stress. Thus, if viscosity is calculated at each of a series of shear rate points, then the resultant values *increase* for dilatant materials, and, therefore, such materials are known as shear-thickening systems. Such materials increase in volume, i. e. dilate, when sheared and are hence known as dilatant. On the removal of stress, these materials revert back to their original state of fluidity. The flow properties of such materials are opposite to that of pseudoplastics. The rheogram and viscogram of a dilatant material are illustrated in Fig. 6.8.

The *Ostwald-de Waele* equation used to describe pseudoplasticity is also applicable for dilatant materials.

$$S^N = \eta' D \ldots (10)$$

where, $N$ is less than 1 for dilatant materials. As the degree of dilatancy increases, the value of $N$ decreases.

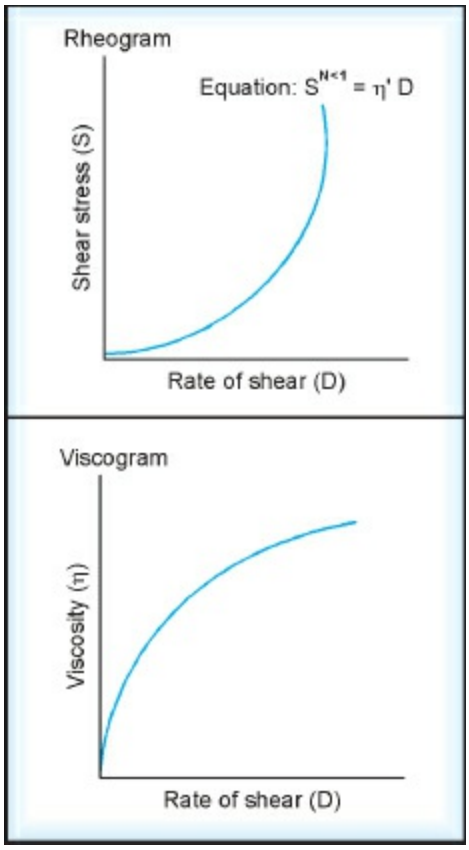


**Fig. 6.8:** Rheogram and viscogram for dilatant flow

Suspensions containing small amount of liquid and high concentration >50 % (w/w) of small, deflocculated particles, exhibit dilatant behaviour. At rest, the deflocculated particles do not tend to aggregate, but are intimately packed with minimum interparticle volume. At low shear rates, the amount of vehicle is sufficient to fill the volume, and to lubricate and permit the particles to slip past each other. At this stage, the material, being fluid, can be poured or stirred. On increasing shear stress, the particles bunch up together, take an open form of packing and develop large voids rather than slipping past each other. The amount of vehicle being constant is unable to completely fill the void spaces and the suspension appears dry, as if the suspended particles had expanded or dilated. With further increase in shear rates the material becomes more viscous, attaining a solid paste-like consistency, and hence known as shear-thickening systems. When shear is removed, the void volume decreases and the viscosity drops and the suspension appear wet again. The diagrammatic explanation of dilatant behaviour is depicted in Fig. 6.9.

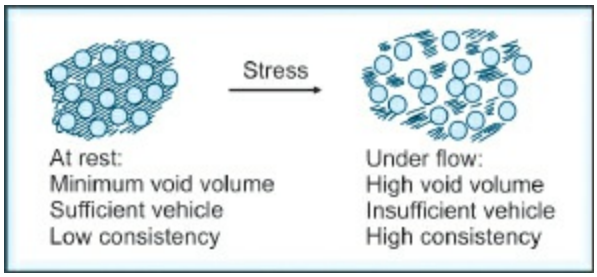



**Fig. 6.9:** Explanation of dilatant behaviour

Among the few systems reported to exhibit dilatant flow are deflocculated suspensions of inorganic pigments in water 30–50% (v/v) titanium dioxide, 12% red iron-oxide and 39% barium sulphate), or in non-polar liquids (18% v/v red iron oxide), and suspension of starch in water aqueous glycerin or ethylene glycol.

## Time-dependent Non-Newtonian Flow

### *Thixotropy*

In the previous discussion, for shear thinning systems, it was assumed that the system adapts itself to changing shear instantaneously, i.e. so fast that the rheograms at increasing or decreasing shear rates are a single curve.

However, if the suspension is viscous or if the suspended particles are large, their Brownian motion is too slow to restore the broken interparticle links instantaneously. If the structure does not immediately recover, the descending rheogram will have lower stress values at each shear rate than the ascending rheograms and the apparent viscosity decreases even while the system is under constant shear. Such a body is said to be *thixotropic*. Thixotropy is therefore, a phenomenon resulting from the time dependency of the breakdown or rebuilding of structure on standing, i.e. a reversible and isothermal transformation of gel to sol.

For example, bentonite sodium 8% (w/w) gel, when stirred above the yield value for few minutes, flows and can be poured. When kept undisturbed, for an hour or two, it reverts back to gel state as the Brownian motion rebuilds the house-of-cards structure.

*Thixotropy in a pseudoplastic system* is shown in Fig. 6.10. Starting with the system at rest (O), in the rheogram for such systems obtained by plotting shear stresses shear rates, two curves are obtained, an up-curve (OAB) when the shear rate is increased and a down-curve (BCO) when the shear rate is reduced. The up-and down-curves are non-super-imposable and the down-curve is displaced lower to the up-curve indicating that the viscosity of the system at any rate of shear is lower on the down-curve than on the up-curve. Thus, the shear stress required to maintain the rate of shear is reduced from $S_1$ to $S_2$ and the apparent viscosity has dropped from $S_1/\eta$ to $S_2/\eta$. This contrasts with the rheogram of pseudoplastic material (Fig. 6.11), where up-curve and down-curve coincide.

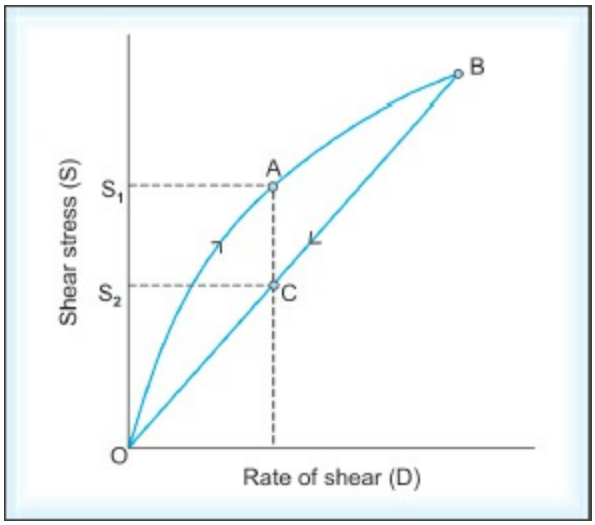



**Fig. 6.10:** Thixotropy in a pseudoplastic system

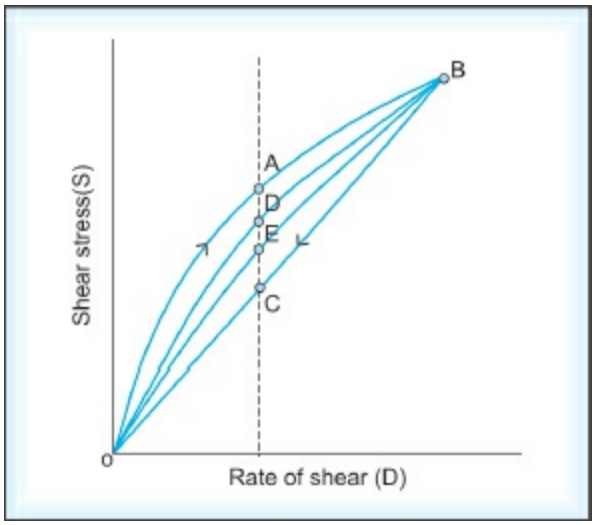



**Fig. 6.11:** Rheogram representing successive shear cycles for a thixotropic pseudoplastic liquid

After subjecting to shear cycle (OABCO), if the thixotropic material is kept at rest for a sufficient time period, it restores its original high consistency. If no rest period is allowed, and the shear cycle is repeated as soon as down curve is completed, the next up-curve is below OAB, say, ODB in Fig. 6.11. A third shear cycle without rest period, may give up-curve, OEB. The down curve, BCO, may be curved or straight. If the buildup of structure is slow, there may be no structure left after third cycle and in that case up-curve coincides with straight down-curve, BCO, and the liquid turns Newtonian. This change is temporary and after a prolonged rest period the

curve BCO, reverts to OAB.

*Thixotropy in a plastic system* is shown in Fig. 6.12. After imposition of one or more shear cycles, the yield value may remain unaltered as in curve A may reduce as in curve B (called as false body behaviour); or it may disappear as in curve C.

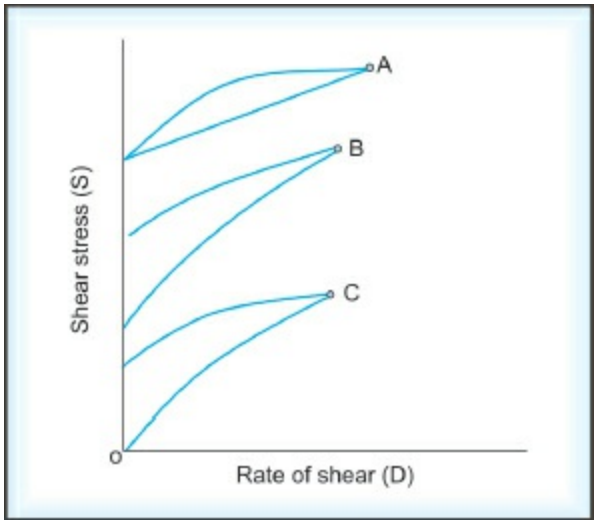



**Fig. 6.12:** Rheogram of plastic systems exhibiting thixotropy

*Hysteresis loop* is the area enclosed by up-curve and down-curve (OABCO of Fig. 6.10) or by up-curve, down-curve and the stress axis (Curves B and C of Fig. 6.12). The area enclosed by the two curves measures the extent of thixotropic breakdown of the system, and with decrease in loop area the structural breakdown decreases. Materials with no structure are Newtonian. The absence of hysteresis in the rheograms of plastic and pseudoplastic system is because the rebuilding of structure by Brownian motion in these systems is as fast as or faster than the shear-induced structural breakdown. The magnitude of difference in the up-curve and down-curve is known as the degree of hysteresis and determines the time taken to reacquire the structure.

## Bulges and Spurs in Thixotropy

Bulges and spurs represent complex hysteresis loops observed in pharmaceutical dispersions (Fig. 6.13). They are obtained when the dispersions are sheared in a viscometer in which the shear rate is increased to a point and then decreased, and then the shear stress is read at each shear rate value to yield the rheograms. Bulge is a characteristic protrusion in the up-curve observed in the hysteresis loops of concentrated bentonite gel, 10–15% (w/w). It is due to the "house-of-cards structure" formed by the crystalline plates of bentonite that causes swelling of bentonite magmas and bulge in the rheograms.


**Fig. 6.13:** Rheogram for a thixotropic material showing bulge and spur in the hysteresis loop

In highly structured systems, for example, procaine penicillin gel formulation, the up-curve develops a spur-like protrusion. Such materials demonstrate a high yield value, known as the Spur Value, *Y*, which traces out a bowed up-curve when the three-dimensional structure breaks in a viscometer. The spur value depicts a sharp point of structural breakdown at low shear rates.

## *Rheopexy*

Once the interparticle links or entanglements among the macromolecules chain have been broken by shear stress, their restoration by Brownian motion

is slow if the particles are large, or the suspension or solution is viscous. In such cases, gentle vibration and shaking (rocking and rolling) may accelerate the restoration of interparticle links between macromolecules. The gentle movements provide mild turbulence, which helps to disperse the particles to acquire a random orientation, and thus re-establish the network. In the case of sheared dispersion of bentonite sodium 8%, (w/w), gentle vibration or rotation speeds up the process of reformation of a gel by rebuilding the house-of-card structure.

## *Negative Thixotropy or Anti-thixotropy*

It is defined as a reversible time-dependent increase in viscosity at a particular rate of shear. In a rheogram of anti-thixotropic system (Fig. 6.14), the down-curve appears upper to the up-curve indicating that the viscosity of the system at any rate of shear is higher on the down-curve than on the up-curve.

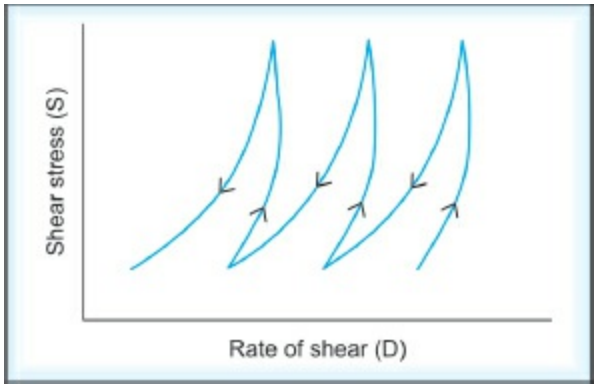



**Fig. 6.14:** Rheogram for magnesia magma showing antithixotropic behaviour

Anti-thixotropic systems are flocculated suspensions containing low solids content (1–10%). In anti-thixotropy there is an increased frequency of collision of the dispersed particles or polymer molecules in suspension, leading to an increase in interparticle bonding with time. Thus, the original state of a large number of individual particles and small floccules is changed to that containing a small number of relatively large floccules. At rest, the large floccules break up and the original state of individual particles and small floccules is restored. Magnesia magma is the classic pharmaceutical example of this behavioural type. In rheological studies, it has been found that on alternately shearing magnesia magma at increasing and decreasing

rates of shear, the magma thickens, but at a decreasing rate. Finally, it reaches equilibrium where there is no increase in consistency on further applying cycles of increasing-decreasing rates of shear. At the final stage, the up-curve and down-curve are coincided.

## *Negative Rheopexy*

It is observed in anti-thixotropic systems where gentle vibration, shaking and mild turbulence speeds up the reformation of solution from the gel state. In this, an antithixotropic system, such as magnesia magma becomes more mobile under the influence of mild turbulence. Figure 6.15 shows the schematic representation of time-dependent non-Newtonian flow.

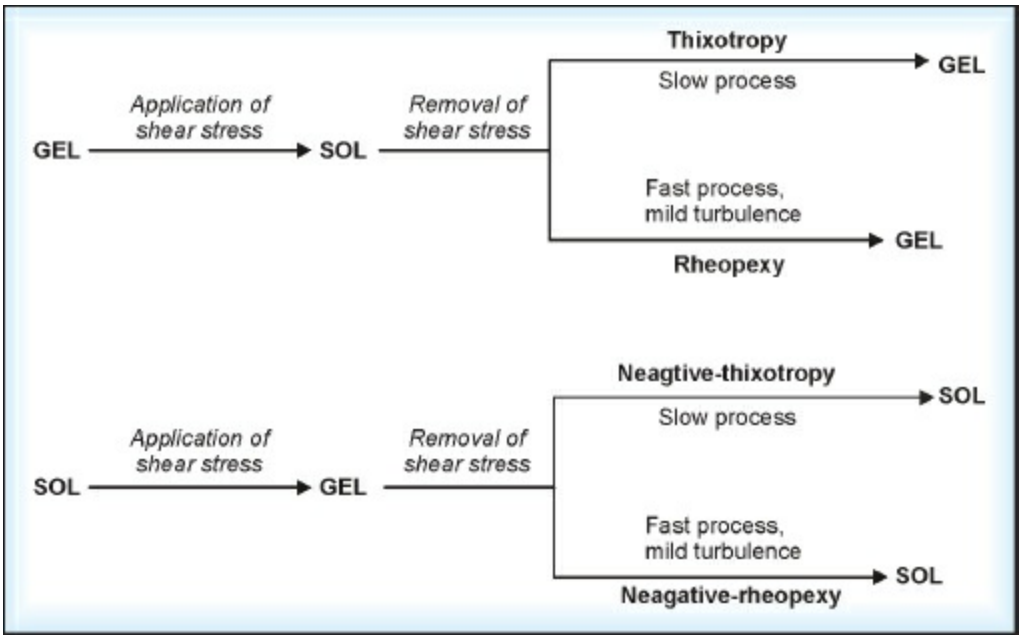



**Fig. 6.15:** Schematic representation of time-dependent non-Newtonian behavior depicting thixotropy, rheopexy, negative-thixotropy and negative-rheopexy

# DETERMINATION OF RHEOLOGICAL PROPERTIES

## Measurement of Viscosity

Selection of the correct instrumental method is important in the successful determination and evaluation of rheological properties. Based on the principle of measuring viscosity, three types of viscometers are available, including:

1. *Capillary viscometers:* Based on the rate of flow of a liquid through an orifice or a fine capillary.
2. *Density-dependent viscometers:* Based on the velocity of a falling object (rising air bubble) through a liquid under the influence of gravity.
3. *Rotational viscometers:* Based on the resistance of a rotating element in contact with or immersed in the liquid.

The choice of instrument in determining rheological properties also depends on the type of material to be analyzed. Based on the type of rheogram obtained, the instruments are classified as:

a. *One-point instruments:* They provide a single point on a rheogram. Extrapolation of a line from this point to the origin yields a complete rheogram. They are suitable only for Newtonian fluids. Examples include capillary viscometers and falling-sphere viscometers.
b. *Multi-point instruments:* They can be used for the non-Newtonian systems, by characterizing the flow properties at variable rates of shear the complete rheogram can be obtained. They can be used for the viscosity determination of Newtonian fluids as well. Examples include cup and bob viscometers, and cone and plate viscometers.

## Capillary or Tube Viscometers

*Principle*. In capillary or tube viscometers, the fluid flows through a round-bore capillary or tube. The fluid in immediate contact with the wall is motionless while that at the center is at the maximum velocity. Between these two limits is then a velocity gradient. The driving force causing liquid to flow is its weight, whereas viscous drag of liquid restrains the flow. Flow of liquid through narrow capillary is depicted by *Poiseuille's equation* and is given as:

$$\eta = \Delta P\, t\, \pi\, r^4 / 8LV \ldots (12)$$

where, $V$ is volume flowing through the capillary per unit time, and $r$ is the radius of the capillary. $L$ is the length of the capillary and $\Delta P$ is the pressure difference across the capillary, which provides the appropriate force to overcome the viscous drag. If $\Delta P$ in the Poiseuille's equation is replaced with hydrostatic pressure, $h\rho g$ of a liquid column of height $h$ and density $\rho$; $g$ is acceleration due to gravity, the equation changes to:

$$\eta = (h\, g\, \pi\, r^4 / 8LV).\ \rho\, t \ldots (13)$$

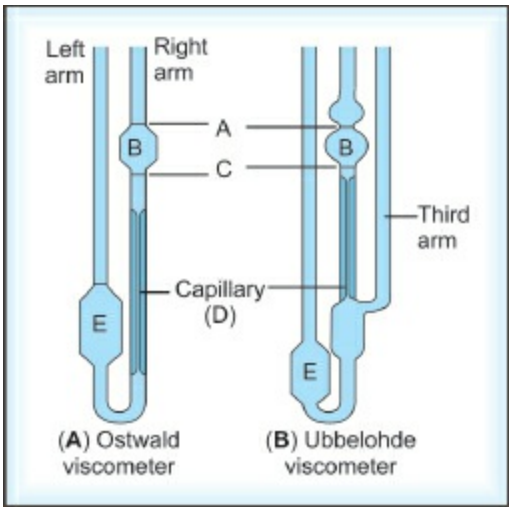
or

$$\eta = K\, \rho\, t \ldots (14)$$

Therefore,

$$K = h\, g\, \pi\, r^4 / 8LV \ldots (15)$$

*Working:* A standard volume of liquid is introduced into viscometer through left arm and is then drawn up from bulb E into the bulb above the mark A, by suction (Fig. 6.16). The efflux time, $t$, required for the liquid level to fall from upper meniscus from line A to line C of the fluid contained in upper reservoir B is measured. The diameter and length of the capillary D control the flow time.




**Figs 6.16A and B:** Capillary or tube viscometers

Calibration of the viscometer is done by determining the constant *K* experimentally with a liquid of known viscosity and density and measuring efflux time.

$$K = \eta/\rho\ t \ldots (16)$$

The viscosity of any unknown liquid is then determined using Eq. 16 as follows:

$$\eta = K\ \rho\ t$$

It is not necessary to determine K, when relative viscosity measurements are made by comparing a standard reference $\eta_R$ and the unknown $\eta_u$. In each case, the volume *V* flowing is the same, but the time for flow are $t_R$ and $t_u$, respectively. It then follows, on substitution into Eq. (16) that:

$$\eta_R/\eta_u = t_R\ \rho_R/t_u\ \rho_u \ldots (17)$$

If the viscosity of reference liquid is not known, a relative viscosity, η, can be determined as defined in the following equation:

$$\eta = \eta_R/\eta_u = t_R\ \rho_R/t_u\ \rho_u$$

If the density of liquid is not known, a kinematic viscosity can be determined using following equation:

$$\text{Kinematic viscosity, } \upsilon = \eta/\rho = K\ t \ldots (18)$$

The form of capillary most frequently used is the glass system typified by the *Ostwald viscometer* (Fig. 6.16A). Such instruments should, of necessity,

be restricted to materials that can readily flow into and out of the apparatus without rheological damage. Hence, these liquids should be essentially Newtonian in behaviour. In general, the use of Ostwald viscometers is restricted to a “one-shear-range” measurement only, and the instrument must be selected for the individual sample under investigation.

Because many research and quality control measurements utilize some variant of the Ostwald viscometer, there are commercial instruments that measure the time for flow between the two indices with high precision by using photoelectric timing. Schott America (New York, USA) markets a system capable of automatic loading of up to 30 samples for sequential measurement by such a timer. Alternate versions of the classic Ostwald viscometer have been developed to minimize inherent problems of the individual practical instruments. *Ubbelohde suspended-level viscometer* consists of an additional third vertical arm attached to the bulb below the capillary part (Fig. 6.16B). The third arm ventilates the liquid below capillary tube and keeps the volume in the middle arm constant. Thus, some instruments depend less on the varying pressure differential, some use either pressure or vacuum as a driving force, some require minimal volume, and some permit successive dilutions to take place in the instrument to follow concentration behaviour.

*Example 4:* The flow time for water in an Ostwald Viscometer was measured at 20°C as 224 s for the average of 5 trials. Similar measurements for an oil of density 0.748 g/cm$^3$ were 426 s. What is the viscosity of the oil? The density of water at 20° is 0.998, and the viscosity is 1.005 cp. The kinematic viscosity of water is

$$v = \frac{\eta}{d} = \frac{1.005}{0.998} = 1.007 \text{ centistokes}$$

Then, by equation (10):

$$\frac{v_{oil}}{v_{water}} = \frac{t_{oil}}{t_{water}}$$

$$\frac{v_{oil}}{1.007} = \frac{426}{224}$$

$$v_{oil} = 1.92 \text{ centistokes}$$

Then

$$\eta_{oil} = d_{oil} v_{oil} = (0.748)(1.92) = 1.43 \text{ cp}$$

Commercially, Ostwald viscometers are purchased by the choice of the

centistokes per second constant desired. The normal expected flow time is 200 to 800 for a given instrument. Lesser times generally utilize some form of external timing. Cannon Instrument Co. (State College, PA) sells a range of Ostwald Viscometers with constants ranging from approximately 0.002 to 20 centistokes per second.

*Example 5:* Quality control personnel wish to purchase an Ostwald Viscometer for routine monitoring of the viscosity of an oil whose density is 0.821 g/cm$^3$ and whose viscosity is about 125 cp. They require the observed flow time to be between 200 and 400 s. What specification instrument should they buy?

$$v = \frac{\eta}{d} = \frac{125\ \text{cp}}{0.821\ \text{g/cm}^3}$$
$$= 152\ \text{centistokes}$$

To have a flow time of 300 s, the constant of the instrument should be approximately:

$$\frac{150}{300}\frac{\text{centistokes}}{\text{s}} = 0.5\,\text{centistokes/s}$$

The Cannon Instrument Co. has size number 350, which has this instrument constant value.

## *Extrusion Rheometer*

Another alternative to Ostwald viscometer is a general instrument containing replaceable capillary tubes and a variable driving force. Such an instrument provides maximum flexibility of use. This system is shown schematically in Fig. 6.17. A sample storage chamber is loaded with the sample to be investigated. The sample is extruded through a capillary tube attached to one end. The chamber contents are forced through this exit capillary by the force of the piston. Two versions are practical. In one, the piston is driven at a constant linear rate.

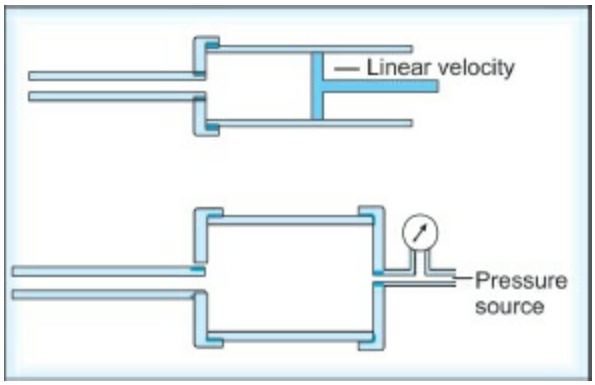



**Fig. 6.17:** The extrusion rheometer. Either a constant linear velocity is imparted to the piston to give a preset volume delivery and the piston force is measured, or a preset force (air pressure) is applied and the rate of volume delivery is determined

The alternative version of Fig. 6.17 requires that a constant force be applied to the piston, and the resultant displacement of piston or sample extruded is measured. Typically, the constant force is applied from gas cylinders with a good pressure regulation system. One commercially available example is the ***CannonManning Pressure Viscometer*** (Cannon Instrument Co., State College, PA) as shown in Fig. 6.18.

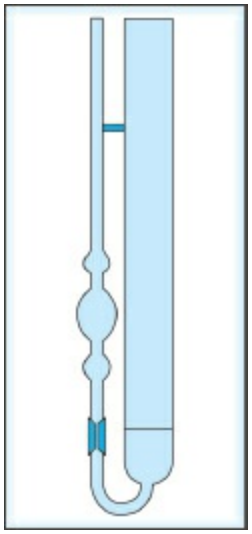


**Fig. 6.18:** Cannon-Manning pressure viscometer

In Cannon-Manning pressure viscometer, the capillary is a metallic fixture between two swage nuts. Pressure is applied through the side arm to the sample (18 ml) contained in the holding vessel. The volume displaced is

measured by either or both of the two glass bulbs. The pressure drop occurs completely within the capillary so that the glass portion of the system is not under any pressure. This instrument is capable of performing rheological studies of most pharmaceutical semisolid formulations: pastes, ointments, and creams.

By suitable choice of the flow tubing, the stress measurement devices, and the timing systems, the extrusion systems are capable of being utilized in virtually any pharmaceutical or cosmetic formulation system, including those in which low shear rates are necessary to define yield value. In these, as in any other measuring device, it is essential that no rheological damage to the sample occurs in loading it for measurement.

## Density-dependent Viscometers

### *Falling Sphere Viscometer*

*Principle*. Stokes law is the basis for falling-sphere viscometers according to which the motion of a body through a viscous medium, is resisted by viscous drag. Initially, the body experiences acceleration due to gravity, but soon this acceleration is balanced by the viscous drag and the body falls with a uniform terminal velocity.

Physically, the falling ball viscometer requires the determination of the time $t$ for a ball of density $\rho_B$ to fall through a fixed distance of a liquid of density $\rho_L$ and of viscosity $\eta$ The equation governing this determination may be developed from first principles if the ball falls freely in an "ocean" of liquid, that is, there is no effect from the container walls. However, a general equation of the following form may be used regardless of whether a free fall, a slide and/or roll down an inclined tube is used:

$$\eta = K(\rho_B - \rho_L)t \ldots (19)$$

The constant $K$, includes wall interaction factors. A liquid of known viscosity is then used to calibrate the instruments for general use in a manner analogous to that with Ostwald and similar capillary viscometers.

*Example 6:* A ball of density 3.12 g/cm$^3$ takes 109 s to fall the fixed distance of an inclined tube viscometer when a calibrating liquid of density 0.94 g/cm$^3$ and viscosity 8.12 poises is used. A) What is the instrumental constant? B) What would be the viscosity of a sample oil of density 0.88 g/cm$^3$ if the fall time under similar conditions is 125 s?

a. $\eta = K(\rho_B - \rho_L)t$

$$K = \frac{\eta}{(\rho_B - \rho_L)t}$$

$$= \frac{8.12}{(3.12 - 0.94)(109)} \frac{\text{poises}}{(\text{grams/ml})(\text{s})}$$

$$= 0.0342 \text{ poises gram}^{-1} \text{ ml s}^{-1}$$

b. $\eta = 0.034(\rho_B - \rho)t$

$$= (0.034)(3.12–0.88)(125)$$

$$= 9.571 \text{ poises}$$

It also is apparent that:

$$\frac{\eta}{\eta_L} = \frac{K(\rho_B - \rho_L)t_L}{K(\rho_B - \rho)t}$$

$$\eta = \frac{(\rho_B - \rho)t}{(\rho_B - \rho_L)t_L}\eta_L$$

$$= \frac{(3.12 - 0.88)(125)}{(3.12 - 0.94)(109)}(8.12)$$

$$= 9.57 \text{ poises}$$

One commercial example is the *Hoeppler viscometer*, a 200 mm tube with an 8 mm radius and a normal 10° tilt from the vertical (Fig. 6.19). Utilizing the appropriate ball, it has a workable range of 0.01 cp to 10,000 poises. However, this is a one-point measurement involving uncertain shear rate and shear stress.

**Fig. 6.19:** Hoeppler falling sphere viscometer

### *Bubble Viscometer*

It is based on a similar principle. A series of sealed standard tubes have calibrated oils covering a range of viscosities. Each has a small air bubble of exact geometry. The unknown sample is placed in an empty tube and stoppered so that it is identical in bubble content to the standards. The unknown and standard tubes are inverted, and the bubble rise times are compared to determine the standard most resembling the unknown.

Certainly, these two gravity procedures have little to offer when compared with modern rotational viscometers.

## Rotational Viscometers

*Principle*. These instruments are based on the fact that a solid rotating body immersed in a liquid is subjected to a retarding force due to the viscous drag, which is directly proportional to the viscosity of the liquid.

### *Cup and Bob Viscometers*

As represented by Fig. 6.20, the cup and bob viscometer comprise two members, a central bob or cylinder and a coaxial or concentric cup. One or both are free to rotate in relation to each other. Between these is the test substance, in the annulus. Three basic configurations have been utilized.

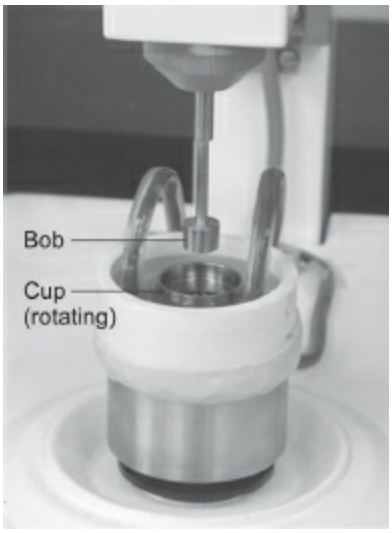



**Fig. 6.20:** Couette type cup and bob viscometer

i. *Couette-type:* Contains rotating outer cup with strain measurement on the central bob. For example, MacMichael viscometer.
ii. *Searle-type:* Contains rotating central bob with strain measurement on the cup. For example, Stormer viscometer and Brookfield viscometer.
iii. Fixed cup with both rotation and strain measured on the bob. For example, Contraves viscometer, Epprecht viscometer and Rotovisko.

The *MacMichael type* viscometer provides cup rotation by the force of weights hanging on a pan that was suspended from a cord wound around the cup. The central bob displacement was restrained by an appropriate spring of known modulus so that the angular displacement could be equated to angular

torque on the bob.

In the *Stormer type* viscometer, the cup is stationary and the bob is driven by weights hanging at the end of pulley to which the shaft of the bob is connected. The shear stress is varied by applying different weights and the corresponding shear rate is measured by the speed of rotation of the bob. The number of revolutions per minute is determined by means of revolution counter.

In operation, the test system is placed in the space between bob and cup. A known weight is placed and time taken by the bob to rotate specific number of times is determined and converted into revolutions per minute (rpm). The procedure is repeated by increasing weights and a rheogram is obtained by plotting rpm vs weight added. By the use of appropriate constants, the rpm value can be converted to shear rates and weights can be transposed into shear stress.

The viscosity of the material may be calculated using the following equation;

$$\eta = K\,(w/\upsilon) \quad \ldots (20)$$

where, $w$ is the weight in grams, $\upsilon$ is rpm and $K$ is instrument constant.

The third classification represents the majority of modern mass-produced instruments. The rotation, usually of the inner member, is provided by a synchronous electric motor, usually with variable ratio drives that provide a series of discrete speeds, with the other member being rigid. An alternative is a motor that provides a continuous variation in speed by an appropriate transmission control or by a variable speed motor drive. A mechanical linkage in the shaft measures angular displacement of the bob from the driving armature. Figure 6.21 exhibits the Brookfield manner of achieving this dual function. More advanced instruments have replaced the measurements of angular displacement with strain gauge linkages of some form to permit direct electronic recording or digital readout.

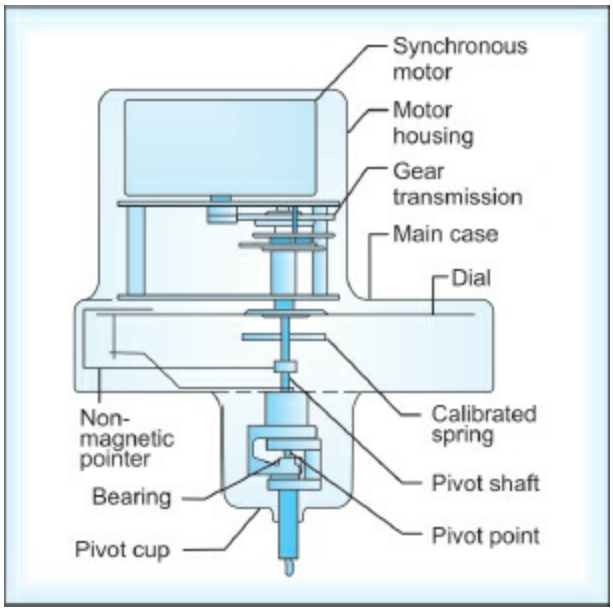



**Fig. 6.21:** Brookfield synchro-lectric viscometer

In the mathematical derivation of the usual expressions for shear stress, S, and shear rate, D, for the cup and bob situation, the existence of a Newtonian fluid is assumed. In this situation, two alternatives may be considered for the average shear stress and shear rate in the gap between the cup and bob. These are the arithmetic (A) and geometric (G) means of the values at the walls of the bob and the cup.

The more customary arithmetic values are:

$$S_A = \frac{T(R_I^2 + R_o^2)}{4\pi h R_I^2 R_o^2} \qquad \text{... (21)}$$

$$D_A = \frac{(R_I^2 + R_o^2)}{(R_o^2 - R_I^2)}\Omega \qquad \text{... (22)}$$

where, $T$ is the measured torque (dyne cm) and usually obtained as the product of a scale reading and the instrumental restoring torque per scale unit of deflection; $R_I$ and $R_o$ are radii of the inner member, i.e. bob, and the outer member, i.e. cup, respectively; $h$ is the effective height or length of the bob; and Ω is the angular velocity of rotation in radians per second.

The values for the geometric mean replace the $(R_I^2 + R_o^2)$ term in the numerators of both Eqs (21) and (22) by 2 $R_I$ $R_o$, yielding:

$$S_G = \frac{T}{2\pi R_I R_o h} \qquad \text{... (23)}$$

$$D_G = \frac{2R_I R_o}{R_o^2 - R_I^2}\Omega \qquad \text{... (24)}$$

Obviously, if the annulus between the cup and bob is very thin, $R_I$ and $R_o$ are virtually identical in magnitude, i.e. these two means are identical. Both sets of equations yield the same value for viscosity:

$$\eta = \frac{S}{D} = \left(\frac{T}{4\pi h\Omega}\right)\left(\frac{R_o^2 - R_I^2}{R_I^2 R_o^2}\right)$$

Commercial instruments supply the values of $R_I$, $R_o$, and $h$, and the instrument constant that converts the observed angular deflection to the restoring torque value, $T$ generated by that deflection.

*Example 7: A* rheometer has a cup with a radius of 2.40 cm and a bob with a radius of 2.28 cm. The effective height of the bob is 6.0 cm. The restoring torque on the rotating bob is 388 dyne cm per instrument scale division. When a sample is measured at a rotational speed of 30 rpm, the scale reading is 42.1 scale divisions. What are the shear stress, the shear rate, and the viscosity?

$$S = \frac{T(R_I^2 + R_o^2)}{4\pi h R_I^2 R_o^2}$$

$$= \frac{(388)(42.1)(2.28^2 + 2.40^2)}{(4)(3.1416)(6.0)(2.28)^2(2.40)^2}$$

$$= 79.3 \text{ dyne cm}^{-2}$$

and

$$D = \frac{R_I^2 + R_o^2)}{(R_o^2 - R_I^2)}\Omega$$

$$= \left[\frac{[(2.28^2 + 2.40^2)]}{[(2.40^2 - 2.28^2)]}\right] \times \left[\frac{30 \text{ rpm}}{60 \text{ s min}^{-1}}\right] \times [2\pi]$$

$$= 61.3 \text{ s}^{-1}$$

Note that the revolutions per minute were converted to radians per second; there being 2π radians per revolution.

$$\eta = \frac{S}{D} = \frac{79.3 \text{ dyne cm}}{61.3} = 1.29 \text{ cp}$$

Note that the same value would have been obtained for the viscosity using Eq. (25).

*Example 8:* A rheometer with a spring constant and cup and bob radii as given in the previous example is used with a known standard viscosity oil to determine the effective height of the bob. If an oil of viscosity 1.98 cp gave a scale deflection of 65.1, what is the effective height of the bob?

$$\eta = \frac{T}{4\pi h\Omega}\frac{R_o^2 - R_I^2}{R_I^2 R_o^2}$$

$$\eta = \frac{T}{4\pi\eta\Omega} = \frac{R_o^2 - R_I^2}{R_I^2 R_o^2}$$

$$= \left(\frac{(65.1)(388)}{(4)(3.1416)(1.98)\left(\frac{30}{60}\right)(2)(3.1416)}\right) \times \left(\frac{(2.40^2 - (2.28)^2}{(2.28)^2(2.40)^2}\right)$$

$$= 6.06 \text{ cm}$$

Often, the end effects cannot be readily calculated for precision geometry. An effective height can then be obtained by calibration in this manner and used with all further work with that cup and bob.

When the ratio of $R_0/R_I$ exceeds 1.1, there are reasons to feel that the use of geometric means of Eqs (23) and (24) is better than use of the arithmetic mean of Eqs (21) and (22) in representing the average condition of the material under shear in the annulus.

Occasionally, in equipment literature and in publications, yet another pair of equations for calculating $S$ and $D$ are used. These result from the assumption that the gap is so wide that $R_o^2$ is much greater than $R_I^2$ and that only the larger term is retained in the summation. The equations are:

$$S = \frac{T}{2\pi h R_I^2} \qquad \ldots (26)$$

$$D = \frac{2R_o^2}{(R_o^2 - R_I^2)}\Omega \qquad \ldots (27)$$

These equations are presented so that they may be recognized when seen in other readings. Surprisingly, the value for the viscosity is still given by Eq. (25) when these are used. There are some rheological empiricists who believe that Eqs (26) and (27) are more suitable for non-Newtonian systems than are Eqs (23) and (24) or (21) and (22). The fundamental practitioners of rheology use a simple equation to establish a Power law relationship and then use the equations developed with that relationship to calculate more proper values of shear stress and shear rate. Such equations are not considered in this discussion.

In addition, it should be recognized that during analysis of plastic material, the shear stress close to the rotating surface is sufficient to exceed the yield value but material away from the rotating surface experience shear stress less than the yield value. Material in this region will remains solid and provides artifactual observations due to formation of plug flow or slippage planes.

If it should become essential to bring values from different cup/bob combinations into absolute coherence, one approach is to use a series of sleeves that fit into a rheometer cup. By conducting measurements with one bob at a series of clearances to the sleeves (cup), one can obtain an estimate of a limiting zero gap by extrapolation to the zero gap condition. This usually permits a smooth transition from one measuring combination to the next as the product changes with age in storage trials. The *Contraves viscometer* has such a set of sleeves available for use in its cups. This versatile instrument, available from the Tekman Company (Cincinnati, OH), is sometimes called the *Epprecht viscometer* after its designer. Another frequently used instrument for pharmaceutical studies is the *Rotovisko* (Haake Inc., Saddle River, NJ). In addition, Brookfield Engineering Inc. (Stoughton, MA) has sets of cup and bob combinations available for its basic instruments.

## *Cone and Plate Viscometers*

In this type, the cone is a slightly beveled plate such that ideally the angle *y* between cone and plate is only a few degrees; even in the cruder forms, it is less than 10° (Fig. 6.22).

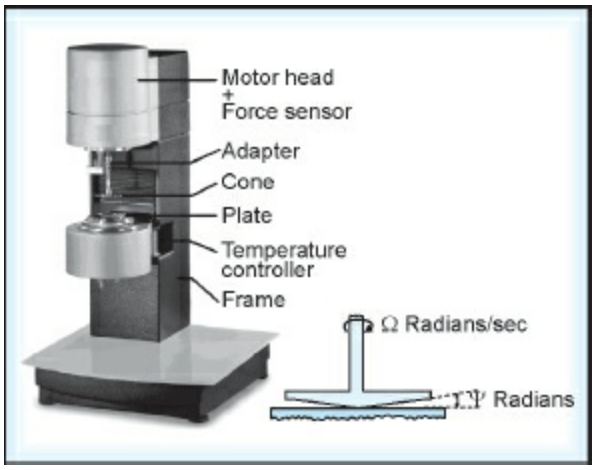



**Fig. 6.22:** Cone and plate viscometer and its basic configuration

The linear velocity at any point on the cone *r* from the apex is rΩ where Ω is the angular velocity while the separation is $r\psi$. Hence, the shear rate is given by Eq. (28):

$$D = \frac{r\Omega}{r\psi} = \frac{\Omega}{\psi} \qquad \ldots (28)$$

Of greatest significance in Eq. (28) is that the shear rate at any point is the same, uniform throughout the gap when the cone angle is small, and therefore plug formation is avoided.

If the cone radius is *R*, then the shear stress *S* is given by:

$$S = \frac{3T}{2\pi R^3} \qquad \ldots (29)$$

where, *T* is the restoring torque measured by the instrument.

*Example 9:* What is the shear stress and shear rate for a cone-plate viscometer measurement using a 5.00 cm radius cone and plate, with the cone rotational speed being 50 rpm, the cone angle 1.8°, and the scale reading 37.6 units? The torque conversion is 1105 dyne cm per scale division.

First, both the rotation and the cone angle must be brought to similar units, in this case, radians:

$$D = \frac{\Omega}{\psi}$$

$$= \frac{(50\ \text{rpm})\left(\frac{1\ \text{min}}{60\ \text{sec}}\right)(2\pi\ \text{radians/revolutions})}{(1.8)\left(\frac{2\pi}{360}\right)(\text{degrees})(\text{radians/degree})}$$

$$= 166.7\ \text{s}^{-1}$$

Alternately, the conversion could be to degrees:

$$D = \frac{(50\ \text{rpm})(360^\circ / \text{rev})\left(\frac{1\ \text{min}}{60\ \text{s}}\right)}{(1.8)(\text{degrees})}$$

$$= 166.7\ \text{s}^{-1}$$

Similarly, the stress is given by:

$$S = \frac{3T}{2\pi R^3}$$

$$= \frac{(3)(37.6\ \text{units})(1105\ \text{dyne cm/unit})}{(2)(\pi)(5.00)^3}$$

$$= 158.7\ \text{dyne cm}^{-2}$$

Geometrically, either the cone or the plate can be the rotating element. The torque may be measured either on the other member or through the coupling drive system exactly as in the Couette viscometers.

The classic instrument of this category has been the *Ferranti-Shirley viscometer* (Ferranti Electric, Corrmack, NY), whose use in a wide range of pharmaceutical and cosmetic literature testifies to its general versatility. In addition, cone-plate systems attachments are available for the Brookfield, Contraves, and Rotovisko systems.

## Penetrometers

These instruments were developed to measure the consistency or hardness of relatively rigid semisolids. The cone and the needle forms are the most commonly used in pharmacy and a few of them are shown in Fig. 6.23. The usual cone is that specified by the ASTM, which has a 30° cone point protruding from a 90° cone. However, other solid-angled cones have been used (Fig. 6.24).

In use, the cone is mounted on an instrument that measures its movement with time.

**Fig. 6.23:** Various types of cone and the needle penetrometers

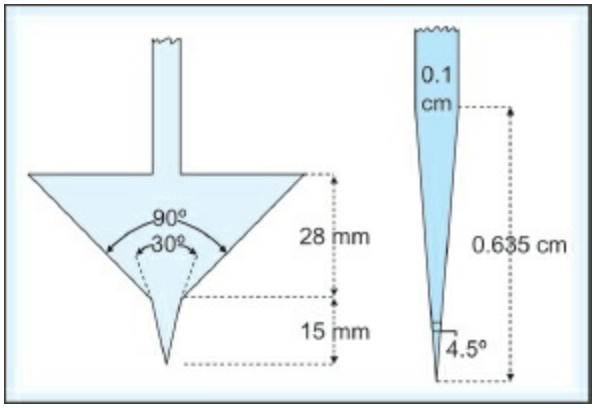



**Fig. 6.24:** Basic configuration of the double cone and needle penetrometers

The tip is set at the surface, and a spring tension is attached to the top of the cone. At time zero, the cone is released. Usually, the penetration occurring in a fixed time is determined. The travel distance is usually reported in decimillimeters, $10^{-4}$ m.

It has been shown that these measurements can provide a self-consistent

measure of yield value (*YV*) through the following equation:

$$YV = \frac{KW}{P^n} \qquad \text{... (30)}$$

where, *K* is an appropriate proportionality constant, *W* is the load in grams on the penetrometer, *P* is the penetration in mm (multiples of 0.1 mm), and *n* is an index. The value of *K* depends on the cone angle, 9670 for 30°, 2815 for 60°, and 1040 for 90°. The index *n* must be determined from a plot of the classically-measured yield value against penetration for a variety of samples. The value of $n = 1.6$ applies to margarine and fat. No equivalent numbers have been reported for products of pharmaceutical interest. However, the highest possible value of *n* is 2, the theoretic value utilized automatically by some workers. For relative values, either number, or an average would not lead to a serious error. Pharmaceutical studies seem to be confined to empiric reporting of penetration values for comparison purposes or for following graphically the penetration with time as a function of penetrometer load. This is not surprising since rarely the necessary time taken to calibrate full rheogram yield values against penetration parameters. This does not detract from the utility of the penetrometer as an effective tool in evaluating changes in yield value for ointments and semisolid creams. Its particular value is the relative rapidity and ease with which such penetrometer measurements may be made.

## Miscellaneous Measuring Systems

A wide variety of geometries are conceivable in which a test medium either moderates motion of a driven object or serves to transmit force across a gradient. Thus, rising spheres, plates, rods, etc. have found utility in various forms of rheometers. Some have practical elements for utilization but often need empiric rather than theoretic calibration. Typical devices are the various plates, rods, and T-bars characteristic of the normal Brookfield viscometer. These measuring systems have provided most of the practical quality control of emulsions and semisolids for the cosmetic and pharmaceutical industries.

Usually, some systems meet a need of special convenience. They are usually calibrated against Newtonian oil and hence readings obtained are apparent viscosities rather than actual values. For routine evaluation and comparison, these empiric instruments are invaluable even though the results cannot be readily translated to absolute rheological parameters.

## Non-Newtonian Corrections

All of the previous equations related to viscometer have been derived considering Newtonian behaviour. It means that shear rate is constant throughout the viscometer. Comparison of non-Newtonian systems requires correction to a shear rate term in reference to a fixed point in viscometer. In general, the correction takes the form:

$$\gamma_{corrected} = \gamma\, F(n) \quad \ldots (31)$$

where, $F$ is the correction factor and $n$ is constant determined from the slope of a log-log plot of shear stress vs. shear rate. Correction factors for the common viscometers are tabulated in Table 6.2.

**Table 6.2:** Non-Newtonian correction factors for viscometers

| **Viscometer** | **Correction factor, F(n)** |
|---|---|
| Capillary | (3n + 1)/4n |
| Falling-sphere | $1 - 2.104\, d/D + 2.09\, d^3/D^3$ where, d is sphere diameter, D is tube diameter |
| Cup and bob (infinite gap) | 1/n |
| Cone and plate | 1 |

## MEASUREMENT OF THIXOTROPY

Quantitative estimation of thixotrophy can be made from estimating the area of hysteresis, which is a measure of thixotrophic breakdown. It can be obtained with the use of a planimeter. Several coefficients of thixotropic breakdown can be used to quantify thixotrophic behaviour in plastic systems.

**Measurement of structural breakdown with increasing rates of shear:** It is used to calculate the thixotropic coefficient, M, which gives the loss in shearing stress per unit increase in rate of shear. A plastic material (Fig. 6.25) is subjected to increasing rates of shear till a ceiling of shear value of $\upsilon_2$. On decreasing the rate of shear, a down-curve is obtained, from the slope of which plastic viscosity. $U_1$, can be calculated. Without disturbing, the rate of shear is increased to another higher value, $\upsilon_2$, a down-curve is obtained for it as well, from the slope of which plastic viscosity, $U_2$ can be calculated.

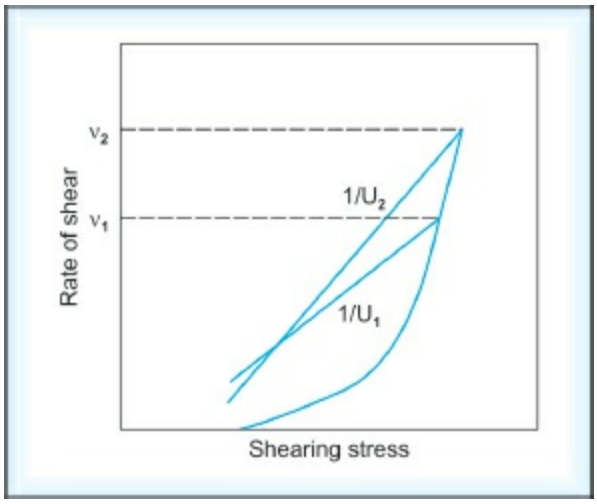



**Fig. 6.25:** Structural breakdown of a plastic system with increasing rates of shear

The value of thixotropic coefficient can be calculated from the following formula:

$$M = (U_1 - U_2)/\ln (\upsilon_2/\upsilon_1) \quad \text{... (32)}$$

where, $U_1$ and $U_2$ are the plastic viscosities of the two down-curves having

maximum rates of shear of $\upsilon_1$ and $\upsilon_2$, respectively. The units of M are dynes.s.cm$^{-2}$.

**Measurement of structural breakdown with time at constant rate of shear:** When the shear rate of a thixotropic material (Fig. 6.26) is increased (up-curve *AB*) and then decreased (down-curve *BE*) at the same rate, a typical hysteresis loop *ABE* is obtained. However, if the sample is taken up to point *B* and then the rate of shear is maintained constant for time '$t_1$' seconds, the shearing stress and so the consistency of the material would decrease to an extent depending on the time for shear, rate of shear and degree of structure in the material. Then a hysteresis loop, *ABCE* is obtained. If the same sample is held for time '$t_2$' seconds at the same rate of shear, a hysteresis loop *ABCDE* is obtained. Based on these rheograms, thixotropic coefficient *B*, i.e. the rate of breakdown of system at a constant rate of shear can be calculated using the following formula:

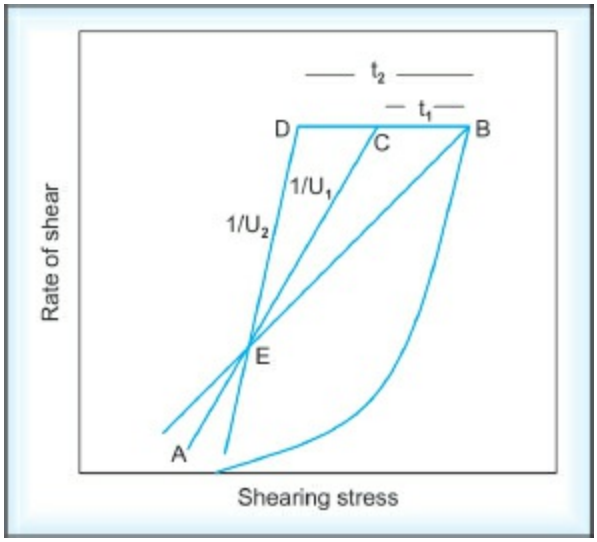



**Fig. 6.26:** Structural breakdown of a plastic system at constant rate of shear with time

$$B = (U_1 - U_2)/\ln (t_2/t_1) \ldots (33)$$

where, U1 and U2 are the plastic viscosities of the two down-curves, after shearing at a constant rate of shear for $t_1$ and $t_2$ seconds, respectively.

## VISCOELASTICITY

Between the limits of elastic deformation and pure viscous flow, there exists a continuum of combinations of these limits. Such behaviour is called *viscoelastic flow*. Viscoelastic materials exhibit both viscous fluidity and elastic solidity when undergoing deformation. Viscous materials resist shear flow and strain linearly with time when a stress is applied while elastic materials strain instantaneously when stress is applied and quickly return to their original state once the stress is removed. Viscoelastic materials have elements of both of these properties and, as such, exhibit time-dependent strain. The relative proportion of elastic deformation and viscous flow depends upon the duration of stress.

When a stress is applied to a viscoelastic material such as a polymer, parts of the long polymer chain change their position and this rearrangement is called Creep. During rearrangement, the polymers remain as a solid material. The rearrangement occurs in order to accompany the stress, and as this occurs, it creates a back stress in the material. When the magnitude of the back stress becomes equal to the applied stress, the material no longer creeps. If the applied stress is removed, the accumulated back stresses will cause the polymer to return to its original form. The material rearranges or creeps, which gives the prefix visco-, and the material recovers, which gives the suffix-elasticity. Thus, viscoelasticity is a molecular rearrangement. A viscoelastic material possesses the following three properties: (i) hysteresis in the stress-strain curve, (ii) time-dependent strain at a constant stress (creep) and, (iii) time-dependent stress at a constant strain (stress relaxation).

Creep and stress relaxation are both manifestations of the same molecular mechanisms, and one should expect that both are related. In particular, the relaxation response moves toward its equilibrium value more quickly than does the creep response.

Viscoelastic property is exhibited by most of the pharmaceutical semisolids such as creams, lotions, ointments, colloidal dispersions, and suppositories. Amorphous and semicrystalline polymers, carbopol gel and aqueous solution of high-molecular weight poly(ethylene oxide) also possess viscoelasticity. Biological fluids like blood, sputum and cervical fluid also exhibit viscoelasticity. Viscoelasticity is widespread even among plastic systems if tested at small shear rates since higher shear rates rupture the

elastic network in these materials. For instance, emulsions are slightly viscoelastic at very low shear due to flocculation of the dispersed globules and interlinking of the flocs.

## Viscoelastic Models

The viscoelastic behaviour of the material may be represented by the use of mechanical models which includes the Maxwell model, the Kelvin-Voigt model, the Standard Linear Solid model and the Weichert model.

Viscoelastic behaviour comprises of elastic and viscous components modeled as linear combinations of springs and dashpots, respectively. Each viscoelastic model differs in the arrangement of these elements.

The elastic property can be represented by Hookean spring represented by the formula:

$$\sigma = E\varepsilon \quad \text{... (34)}$$

where, σ is the stress, $E$ is the elastic modulus of the material, and $\varepsilon$ is the strain that occurs under the given stress. The spring models the instantaneous bond deformation of the material, and its magnitude is related to the fraction of mechanical energy stored reversibly as strain energy.

The entropic uncoiling process is fluid-like in nature, and can be modeled by a “Newtonian dashpot” (represented by the movement of a piston inside cylinder filled with a fluid) such that the stress-strain rate relationship can be given as:

$$\sigma = \eta \frac{d\varepsilon}{dt} \quad \text{... (35)}$$

where, $\sigma$ is the stress, $\eta$ is the viscosity of the material, and $d\varepsilon/dt$ is the time derivative of strain. As stress is applied, piston moves through the fluid and produces shear proportional to viscosity of fluid and as stress is removed the piston does not return to its original position indicating viscous nature.

## Maxwell Model

The Maxwell is a mechanical model in which a Hookean spring and Newtonian dashpot are connected in series, as shown in the Fig. 6.27.

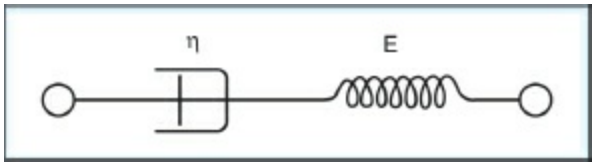



**Fig. 6.27:** Diagrammatic representation of the Maxwell model

When a displacing force is applied, there is an immediate elastic displacement followed by a continuous tinuous viscous flow. The spring stretches immediately, and the dashpot slowly moves independently. When the force is removed, there is an immediate rebound of the elastic displacement, but no viscous flow occurs after that elastic recovery. The spring returns to its original conformation, but the dashpot remains in its new location because there is no force for restoration. Removal of stress leads to complete elastic recovery of the spring but the viscous flow shows no recovery.

The spring should be visualized as representing the elastic component, while the dashpot represents the conformational or entropic component. In a series connection such as the Maxwell model, the stress on each element is the same and equal to the imposed stress, while the total strain is the sum of the strain in each element:

$$\sigma = \sigma_s = \sigma_d \ldots (36)$$

$$\varepsilon = \varepsilon_s + \varepsilon_d \ldots (37)$$

Here, the subscripts *s* and *d* represent the spring and dashpot, respectively. In seeking a single equation relating the stress to the strain, it is convenient to differentiate the strain equation and then write the spring and dashpot strain rates in terms of stress:

$$\frac{d_{\varepsilon \text{Total}}}{dt} = \frac{d_{\varepsilon D}}{dt} + \frac{d_{\varepsilon S}}{dt} = \frac{\sigma}{\eta} + \frac{1}{E}\frac{d\sigma}{dt} \quad \ldots (38)$$

The Maxwell model predicts that stress decays exponentially with time (stress relaxation property), which is accurate for most polymers. One limitation of this model is that it does not predict creep accurately.

## Kelvin-Voigt Model

The Kelvin-Voigt model, also known as the Voigt model, consists of a Hookean spring and Newtonian dashpot connected in parallel, as shown in Fig. 6.28.

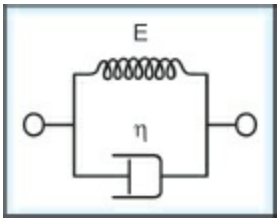



**Fig. 6.28:** Diagrammatic representation of the Kelvin-Voigt model

It is used to explain the creep behaviour of polymers. The constitutive relation is expressed as a linear first-order differential equation:

$$\sigma(t) = E\varepsilon(t) + \eta\frac{d\varepsilon(t)}{dt} \qquad \text{... (39)}$$

When a displacing force is applied to a Kelvin element and held constant, there occurs a viscous flow that decreases with time. In the analog, the viscous flow occurs because of the displacing force, and the net magnitude of that force decreases as the spring stretches to balance the force. When the displacing force is removed, there is a slow return to the original position as the internal force of the spring drives the dashpot back to its original position. In summary, the elastic element has a time component for its attainment of steady state. This model represents a solid undergoing reversible, viscoelastic strain.

Upon the application of a constant stress, the material deforms at a decreasing rate, asymptotically approaching the steady-state strain. When the stress is removed, the material gradually relaxes to its un-deformed state. The Kelvin-Voigt model is extremely good with modeling creep in materials, but with regards to relaxation, the model is much less accurate.

## Standard Linear Solid Model (Maxwell Form)

For polymers whose conformational change is eventually limited by the network of entanglements or other types of junction points, more elaborate spring-dashpot models can be used effectively. Placing a Hookean spring in parallel with the Maxwell unit gives a very useful model known as the Standard Linear Solid model shown in Fig. 6.29.

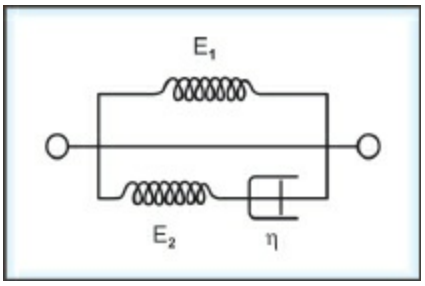



**Fig. 6.29:** Diagrammatic representation of the Standard Linear Solid Model

In this arrangement, the Maxwell arm and the parallel spring $k_e$ experience the same strain, and the total stress ais the sum of the stress in each arm: $\sigma = \sigma_e + \sigma_m$. For this model, the governing constitutive relation is:

$$\frac{d\varepsilon}{dt} = \frac{\frac{E_2}{\eta}\left(\frac{\eta}{E_2}\frac{d\sigma}{dt} + \sigma - E_1\varepsilon\right)}{E_1 + E_2} \quad \ldots (40)$$

When a constant stress is applied, the material will instantaneously deform to some strain (elastic portion), and after that it will continue to deform and asymptotically approach a steady-state strain (viscous portion). The Standard Linear Solid model is more accurate than the Maxwell and Kelvin-Voigt models in predicting material responses, however, mathematically it gives inaccurate results.

## Weichert Model

A real polymer does not relax with a single relaxation time as predicted by the previous models. Molecular segments of varying length contribute to the relaxation, with the simpler and shorter segments relaxing much more quickly than the long ones. This leads to a distribution of relaxation times, which in turn produces a relaxation spread over a much longer time than can be modeled accurately with a single relaxation time.

When it is necessary to incorporate this effect, the Wiechert model illustrated in Fig. 6.30 can have as many spring-dashpot Maxwell elements as are needed to approximate the distribution satisfactorily.

It takes into account that relaxation does not occur at a single time, but at a distribution of times.

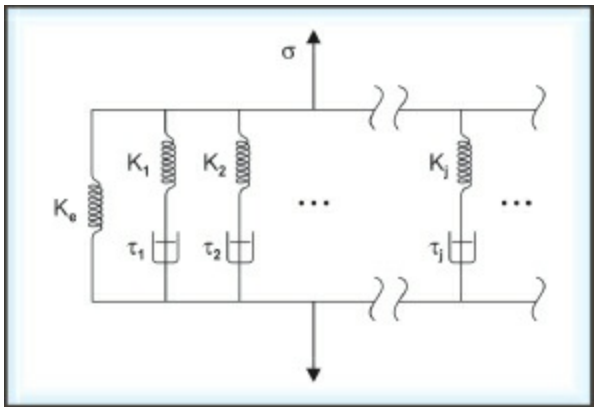



**Fig. 6.30:** Diagrammatic representation of the Weichert model

Schematically, many of these elements may be combined in various permutations. The resulting arrays can be treated mathematically to develop the appropriate characterizing parameters.

Experimental instrumentation to provide data for these mathematical treatments is available. They are of many types depending on the physical nature of the substance to be examined.

## VISCOELASTIC CREEP

When subjected to a constant stress, viscoelastic materials experience a time-dependent increase in strain known as viscoelastic creep. Viscoelastic creep curve can be presented by plotting the creep modulus (constant applied stress divided by total strain at a particular time) as a function of time as shown in Fig. 6.31.

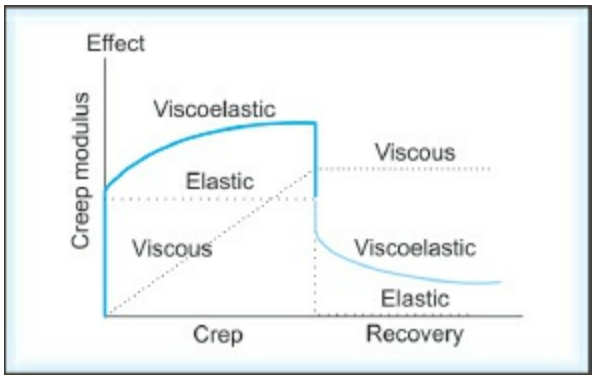



**Fig. 6.31:** Creep curve for a viscoelastic material

The creep curve illustrates viscoelastic properties of the material. The portion AB of the curve represents the elastic behaviour followed by curve BC representing viscoelastic behaviour. The linear portion CD corresponds to the time when bonds in the material rupture under the influence of continuous stress leading to viscous flow. When the stress is removed, the recovery curve DEF is obtained. The portion DE represents instantaneous elastic recovery equivalent to AB followed by slow elastic recovery region EF equivalent to BC. In the recovery curve, no portion corresponding to CD is observed since viscous flow occurs by irreversible destruction of the structure of the material.

# SPECIALIZED PHARMACEUTICAL APPLICATIONS OF RHEOLOGY

## Shear Rates in Pharmaceutical Applications and Resultant Desired Viscosities

Henderson, *et al*. made the first major attempt to establish shear rates corresponding to pharmaceutical use as depicted in Table 6.3.

**Table 6.3:** Shear rate values for different applications

| Parameter | Shear rate value |
|---|---|
| Topical application | 120 |
| Rubbing on ointment tile | 150 |
| Roller mill | 1,000–12,000 |
| Colloid mill | 100,000 |
| Nasal spray (plastic squeeze bottle) | 1000 |
| Pouring from bottle | 100 |
| Rubbing onto skin | 100 to 10,000 |
| Squeezing from a tube or plastic bottle | 10 to 1,000 |
| Pumping of products | 1,000 to 100,000 |
| High-speed filling | 5,000 to 100,000 |
| Yield value | 0.1 |

An example of this extrusion range is that 1 $cm^3/s$ delivery of a liquid cream through an orifice 10 mm in diameter imparts a shear rate of 10 $s^{-1}$, while for the same volume through a 5 mm orifice, such as might occur for a toothpaste, it is 100 $s^{-1}$. Only 0.1 ml/s through a 1-mm orifice imparts a shear rate of 1000 $s^{-1}$. This pharmaceutical characterization, depending upon the application, must bridge an extremely wide range of values of shear rate. The shear rate for individual use by a process may vary and depends on the force applied by the user.

*Example 10:* An interesting example of how to proceed to determine suitable product characterizations is a study by Langenbucher and Lange to determine suitable product squeezing characteristics from a tube and then for ease of application of a cosmetic cream or ointment. They used a series of Newtonian oils of varying viscosity to assess acceptance with a panel.

Study protocol:

- Tubes were of standard aluminum with a diameter of 18 mm, a length of 100 mm, and a 5.3 mm orifice.
- Perception scale of 1 to 5 was used with 1 too thin and 5 too thick.
- A score of 3 was defined as agreeable.
- Perception scale was found to be proportionate to the logarithm of the test medium's Newtonian viscosity.
- A viscosity of 50 to 1000 poises, optimally 200 poises, was considered acceptable for extrusion.
- A viscosity of 0.2 to 5 poises, optimally 1 poise, was preferred for topical use.

Several products for rheological properties were then evaluated as a function of shear rate. One cream with an apparent viscosity of 280 poises at a shear rate of 10 $s^{-1}$ and of 1.8 poises at 10,000 $s^{-1}$ had a panel preference score for extrusion of 3.4 separately for both extrusion and ease of application. From the study of scores as a function of the logarithm of viscosity, these values of 3.4 should have corresponded to 350 and 1.8 poises respectively, an excellent confirmation of the suitability of the screening evaluation. Similar procedures can be used to evaluate any potential product dispenser or use condition with any desired segments of the user population.

*Example 11:* A classic example in pharmaceutical rheology of the need for both structural breakdown under flow and recovery of body (viscosity) is the study of depot procaine penicillin G products by Ober and associates. These workers set two parameters for the evaluation or comparison of products viz. ease of extrusion from a hyperdermic syringe with needle attached, and the degree to which the extrudate would retain itself in a compact volume. An arbitrary force from a 5 g weight on one pan of a two-pan balance was used to press a pad against the tip of the hypodermic needle. If this force was sufficient to plug the needle for a fixed air pressure of 200

psi on the syringe barrel, the formulation was judged too thick. This was found to correlate with the yield value found by rheological measurements. Unfortunately, their values were reported in torque units of dyne centimeters, which require the instrumental parameters to develop shear stress. The suitability of the depot geometry was judged by the relative shape of the product when injected into a gelatin gel, used as a substitute for muscle. The symmetry of the depot was related to the degree of thixotropic recovery of the mass. Using these criteria, these workers then determined how such formulation factors as the particle size and size distribution affected the resulting product.

## Yield Value and Suspensions

Meyer and Cohen demonstrated how the rheological yield value obtainable with their company's Carbopol permitted theoretically the preparation of permanent suspensions. The theoretic yield value (*YV*) for suspension must balance or exceed the force of gravitational settlement. Hence, for spherical objects:

$$YV = \frac{gV(\rho_p - \rho_m)}{A} \qquad \ldots (41)$$

where, $g$ is the acceleration due to gravity, $V$ is the particle volume, $\rho_p$ is the particle density, $\rho_m$ is the suspending medium density, and $A$ is the cross-sectional area of the particle ($\pi R^2$ for round particles of radius $R$).

Meyer and Cohen found theoretic yield values of 65 and 1620 dyne $cm^{-2}$ to suspend sand of density 2.60 and radius 0.030 cm, and marbles of density 2.55 and radius 0.8 cm, respectively.

Suspending agent Carbopol 934 at concentration of 0.18% and 0.40% (w/v) was able to suspend sand and marbles, respectively.

Chong extended this concept by presenting a graphic nomogram (Fig. 6.32 for the necessary yield value to suspend particles of varying density and radius). This graph relates to the difference in density between particle and medium, and is equally valid for particles lighter than the medium, particularly in considering the aeration of a product that has yield value characteristics.

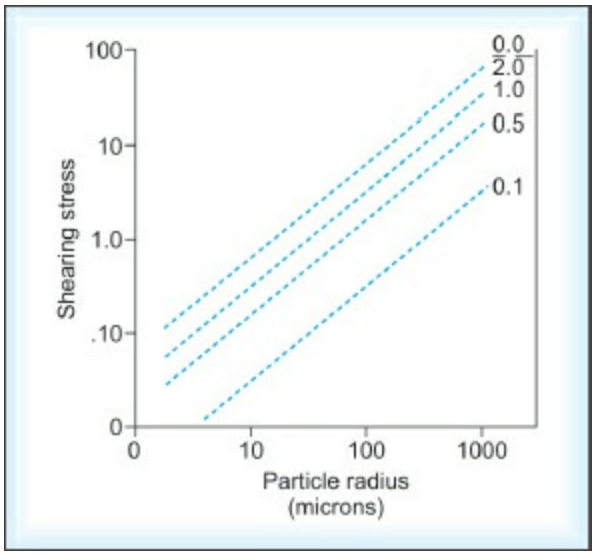



**Fig. 6.32:** Necessary shearing stress to be balanced by yield value for various density differences between particle and medium as a function of particle radius

Any attempt to use this yield value approach obviously requires that real yield values exist at low shear rates as discussed earlier. If the very low shear rate measurements break from the Power law towards Newtonian behaviour, no permanent formulation can result. Unfortunately, no published literature other than the sand study with Carbopol 934 appears to exist relating to stability with aging for this yield value concept.

## Artifactual Observations due to Plug Flow or Slippage Planes

Breakdown of structure under shear may be general, or may occur locally, creating planes of slippage. Frequently, a body of material may remain motionless against the motionless wall of a viscosity-measuring device, while the rest of the material moves as a unit down a tube or with a rotating bob. This behaviour is characteristic of *plug flow*. Such behaviour negates meaningful rheological measurements, except that the stress at which structural rupture occurs may be characterized for that material. Typical rheograms resulting in this behaviour are shown in Fig. 6.33. The Haake Rotovisko has special star-shaped bobs and ribbed bobs available to minimize this effect. It is practical to machine grooves into regular viscometer bobs. These grooves tend to minimize the squeezing out of fluid by syneresis and thus do appreciably raise the shear rate range attainable before plug flow is initiated. This ribbing technique may also be applied in the cone-plate configuration.

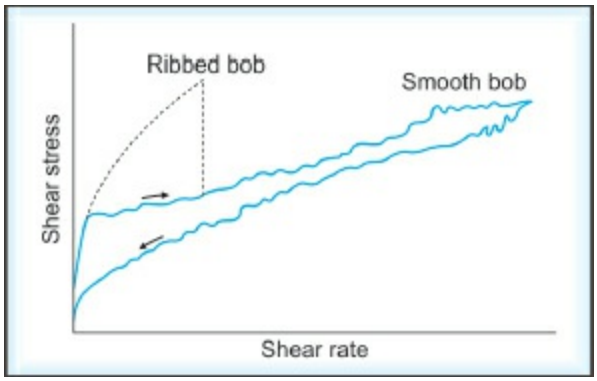



**Fig. 6.33:** Rheogram for a dentrifice obtained with a Hercules Hi-Shear rheometer with smooth and ribbed bobs

Many workers have mistakenly interpreted the point at which plug flow begins as the yield value. Obviously, it represents a significant characterizing parameter, but the shear at which plug flow begins does appear to depend on geometry rather than being an intrinsic term.

In the case of the procaine penicillin study discussed in a previous section, this plug parameter is probably critical to syringe delivery. Indeed, the study evaluates this material fracture point rather than a true rheological yield value.

This loss of cross-linked structure is clearly evident in many low shear rate measurements in solutions of suspending agents, whether montmorillonite clays or xanthan gums. Sometimes, structure may be broken merely by pouring rapidly from a storage container into the rheometer. If the structure is retained in the viscometer, initial torque readings can be quite high before cross-linking structure is lost, which results in reduced torques as the shear rate increases. Similar loss in structure is possible when suspending agents exhibit a loss in suspending power after shaking.

## Measurements of Stickiness or Tackiness

Two types of any tendency to self-adhesion have practical pharmaceutical application. The first is in dermal application, where any sticking of fingers together to a hand or any two skin surfaces can be quite objectionable. The application of rheologic principles to this problem was made by De Martine and Cussler. They found a clear relationship between perceived stickiness and a complex computable function dependent on definable parameters.

An alternative approach, the force to pull two surfaces apart, has been used for measuring stickiness as dermal preparations evaporate. Recently, this type of measure has been developed in detail to evaluate tackiness of coating solutions for tablets as a function of coating composition. Using an Instron Tester, the full time-course of the adhesive force has been followed during extension between the adhesive surfaces. It is of interest to note that tack appears to parallel the orientation effects attributable to elastic parameters of the coating agent.

## Rheological use of Mixing Equipment

A major engineering use of rheological measurements are power requirements in mixers and blenders. In the past, laboratorysized equipment has been instrumented with strain gauges to permit torque measurements to be continuously monitored. This technology is now becoming available in manufacturing equipment.

An interesting example of such a research study was the use of a Brabender Plasti-Corder rotational torque mixer to follow the changes occurring during a wet granulation process. Application of such studies permits stopping further mixing of a granulation at a specific desired point in the granule agglomeration. This can be particularly important when further mixing may create an intractable solid mass or an undesirable fluidizing. Because of the proprietary nature of such studies these are not likely to attain high visibility in the literature.

## Tabletting and Compression Granulation

Dwell time differences during compression, whether slugging or roller compression, can create flow structure differences in granulations. The full viscoelastic implications of compression loading and unloading have recently been developed by Rippie and associates. In particular, they have demonstrated how the fundamental parameters clearly depend on the composition of the tablet.

Instrumented production tablet presses are now available and widely used in the industry for the study of optimum processing conditions. Unfortunately, few of these studies have been released for publication, so that we must rely on such studies as those described in this chapter to realize the potential of such tablet investigations. In particular, the development of formulations to minimize the delay in strain release that contributes to capping is an area in which trade secrets prevent free publication.

## Biorheology

Biorheology is the study of flow and deformation in biological systems. Biological fluids are rheologically complex due to their multicomponent nature but are generally viscoelastic. Rheological properties of biological fluid are important for understanding of their implication in both the healthy and diseased state, and also for successful drug delivery to the body. The rheological nature of various biological fluids is tabulated in Table 6.4.

**Table 6.4:** Rheologic characteristics of biological fluids

| Organ/Fluid | Rheological Behaviour | Utility |
|---|---|---|
| Blood (haemorheology) | Non-Newtonian: Viscoelastic (Fluidized suspension of elastic cells in plasma); at low shear erythrocytes aggregate – viscosity is high; at high shear erythrocytes deform – viscosity is low | Elevated blood viscosity levels have been associated with cardiovascular disorders such as arteriosclerosis, coronary heart disease, angina and myocardial infarction. |
| Bronchial mucus | Non Newtonian: Pseudoplastic (Weak gel containing lipids or glycoconjugates) | Viscosity increases in cystic fibrosis; drugs such as bromhexine and N-acetylcysteine decrease the viscosity of mucus and used to treat chronic obstructive pulmonary disorder. |
| Tear fluid | Non-Newtonian: Pseudoplastic | Viscosity must be high enough to allow tear film to form continuous layer over the eye |
| Vitreous humor | Non-Newtonian (clear | With age, the vitreous |

| | | |
|---|---|---|
| | gel) | humor changes from gel to liquid resulting in pathological conditions |
| Nasal mucus | Non-Newtonian – viscoelastic | In chronic inflammation, there is increase in mucus volume; in sinusitis the dynamic viscosity of mucus is 1.6 Pa/s |
| GIT mucus | Non-Newtonian – viscoelastic (gel) | Ulcerative processes are a result of the weakening of normal GI mucus barrier |
| Cervical mucus | Non-Newtonian – viscoelastic | Transfer of spermatozoa to the uterus; used to predict ovulation |
| Semen | Non-Newtonian – viscoelastic (Semi-gelatinous cellular suspension containing spermatozoa) | Transfer of spermatozoa |
| Synovial fluid | Non Newtonian-strongly pseudoplastic (Clear viscous liquid containing hyaluronic acid) | High viscosity at low shear enables to maintain clearance between articular surfaces; lubricity is aided by pseudoplasticity; when joint moves, i.e. at high shear, viscosity is lowered and protects the cartilage from wear |
| Cerebrospinal fluid | Newtonian | Protects spinal cord and |

| | |
|---|---|
| | brain from physical impact by providing a water cushion |

## SUMMARY

In stability measurements, rheological parameters can provide a method of documenting time-dependent changes.

Equipments suitable for most practical pharmaceutical and cosmetic rheology is commercially available from a variety of sources.

Rheological measurements can provide criteria of product acceptability because they can be correlated to relevant “in use” factors of shear. Recent developments in viscoelastic rheological methodology are now moving from the academic to the industrial applications in pharmaceutical technology. These will lead to further applications in product monitoring and evaluation.

# 7:
# Pharmaceutical Utility

There are some general purpose operations which are though not directly related with the pharmaceutical operations but have great affect on them. They require proper monitoring as well as control to avoid unnecessary breakdown in the ongoing production processes.

Some of these general utilities in a pharmaceutical industry are discussed here to have an insight into some minor points which are many-a-times overlooked initially, though, realized for their importance at the later stages.

To have a successful operation for the purpose of industrial pharmacy, there are certain gaps which need to be filled to have a frictionless platform for a particular pharmaceutical process. For instance, in this chapter some general principles, of controlling the environmental conditions (HVAC) to have proper humidity as well as temperature control as required during the manufacturing of tablets, capsules, parenterals etc., are discussed. Light is also thrown on to the quality of the materials to be used during plant construction and the quality of the major excipients used in almost all pharmaceutical operations i.e. water. Last, but not the least some other applications of gases in a pharmaceutical industry are also looked upon.

## HVAC SYSTEM

Humidity and temperature are the two expediencies which need proper tools for measurement as well as control in a multiple number of pharmaceutical operations. The concept may be realized in case of the processes of drying, storage of chemical substances, control as well as prevention of corrosion, rust or scuffing of highly polished metallic surfaces, and management of explosives and materials having poor electrical conductivity. Apart from these, in the routine processes, for instance, the flow of powders is a susceptible function of moisture content which is determined with the help of humidity measurements. If not considered then the whole tableting processes might break down due to too high humid conditions. Here comes the role of air conditioning that can be defined as a process of treating air so as to control both its temperature and humidity simultaneously to meet the requirements. In addition to human comfort, applications of air conditioning embrace maintenance of proper environmental conditions during production, processing and preservation of desired materials and equipments. It may be obligatory to supply the whole room with air of a certain quality. Otherwise, conditioning may be limited to a small area surrounding a particular piece of equipment.

HVAC (pronounced either “H-V-A-C” or “*H-vak*”) is an acronym that stands for “heating, ventilating, and air conditioning”. HVAC is sometimes referred to as climate control and is particularly important in the design of medium to large scale pharmaceutical industrial setup, where humidity and temperature must all be closely regulated and maintained for a variety of pharmaceutical operations like tablets’ manufacturing, filing and sealing of capsules, sterile product manufacturing and packaging along with safe and healthy working conditions within. Refrigeration is sometimes added to the field’s abbreviation as HVACR. The three functions heating, ventilating, and air conditioning are closely interrelated and are based upon the principles of thermodynamics, fluid mechanics, and heat transfer. HVAC systems can provide ventilation, reduce air infiltration, and maintain pressure relationships between spaces. The delivery of air into room and its removal, is known as room air distribution. In modern pharmaceutical industries, the design, installation, and control systems of these functions are integrated into one or more HVAC systems. Air-conditioning involves both humidification and

dehumidification operations. Before getting into the details of these operations it is important to understand the concept of humidity and principles of psychrometry.

## PSYCHROMETRY

A critical factor in drying operations is the vapour-carrying capacity of the air, nitrogen, or any other gas stream passing over the drying material. This carrying capacity determines not only the rate of drying but also the extent of drying, i.e. the lowest moisture content to which a given material can be dried. The determination of the vapour concentration and carrying capacity of a gas is termed as *psychrometry*. The air-water vapour system is the system most commonly employed in pharmaceutical drying operations. The concentration of water vapour in a gas is called the *humidity* of the gas. Humidity may be expressed in various ways, depending on the information required.

The humidity characteristics of air are best shown graphically in a *psychrometric* or *humidity chart*. Such charts can be found in various handbooks. A psychrometric chart has a formidable look because of the wealth of information presented in a small area. If the different curves in the chart are separated and analyzed individually, however, their utility and ease of use become apparent.

The basic curves of the psychrometric chart are shown in a simplified version in Fig. 7.1. These curves are graphical representation of the relationship between the temperature and humidity of the air-water vapour system at constant pressure. The temperature is shown on the horizontal axis and the vertical axis represents *absolute humidity* (weight of water vapour per unit weight of dry air). The most important curve shown is the curve for *saturation humidity*, i.e. curve CDE. Saturation humidity is the absolute humidity at which the partial pressure of water vapour in the air is equal to the vapour pressure of free water at the same temperature. Under these conditions, the air is completely saturated with moisture, and the humidity does not change when the air is in contact with liquid water at the same temperature.

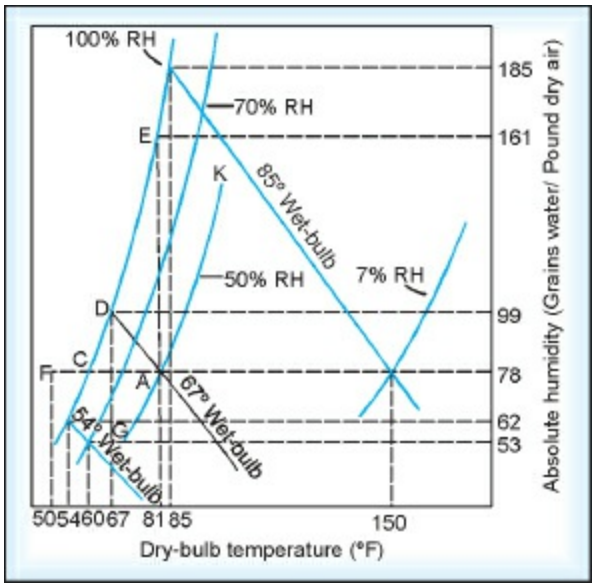



**Fig. 7.1:** Diagram of psychrometric chart showing the relationship of air temperature with humidity

The saturation humidity curve is actually the boundary of a phase diagram. Any point on the curve is uniquely determined by one measurement, either the temperature or the absolute humidity. The relationship between the two variables can be shown more clearly by considering the dotted line, FCA, which corresponds to the absolute humidity, 78 grains water/pound dry air.

At point C, the air is saturated with water vapour, and its temperature, 60°F, is referred to as the *dew point*. The dew point is defined as the temperature to which a given mixture of air and water vapour must be cooled to become saturated (i.e., to hold the maximum amount of moisture without condensation taking place). When the mixture is cooled to temperatures below the dew point, such as 50°F (point F), the water vapour condenses to produce a two-phase system of saturated air (condition C) and droplets of free water.

To make the air usable for drying purposes (without changing the absolute humidity), its temperature must be raised. If the temperature is increased to 81°F (point A), the air is not completely saturated and can accept more water vapour. The relative saturation is usually measured in terms of *percent relative humidity*, the ratio of the partial pressure of water vapour in the air to the vapour pressure of free water at the same temperature. The saturation humidity curve, CDE, is thus the curve for 100% relative humidity (100% RH). Curves of temperature versus absolute humidity at a constant

relative humidity are plotted on the same axes at specific intervals of relative humidity. One of these curves is shown as curve GK with 50% relative humidity. The relative saturation can also be expressed as *percent humidity* or *percent absolute humidity*, the ratio of the absolute humidity to saturation humidity at the same temperature. For air at condition A, the percent absolute humidity is represented by the ratio of the absolute humidity to the absolute humidity of saturated air at that temperature (78/161 = 48%).

If air, under the conditions represented by point A, is used to dry a wet material, the difference in vapour pressure between the surface water and the air causes some of the liquid to evaporate. The latent heat of vaporization of the water cools the evaporating surface below the air temperature. The resultant difference in temperature causes a transfer of heat from the air to the liquid at a rate that increases as the temperature difference becomes larger. Eventually, the heat transferred becomes equal to the heat of vaporization, and temperature stabilizes. The temperature that is reached is called the *wet-bulb temperature*. It is defined as the equilibrium temperature reached by an evaporating surface when the rate of heat transferred to the surface by convection is equal to the rate of heat lost by evaporation. It is called the wet-bulb temperature because it can be measured by means of a thermometer whose bulb is covered by a wick saturated with water. The actual temperature of the air as measured by an ordinary thermometer is called the *dry-bulb temperature*.

The wet-bulb temperature is a function of the temperature and humidity of the air used for the evaporation, and thus can be employed as a means of measuring humidity. For this purpose, a second type of curve is superimposed on the temperature-humidity curves of the psychrometric chart. This is the *constant wet-bulb temperature* line. The constant wet-bulb temperature line is AD for air at condition A, and the temperature corresponding to saturation at point D is the wet-bulb temperature, 67°F.

The humidity of the air is determined by measuring two temperatures, the *wet-bulb* and the *dry-bulb* temperatures. The psychrometric chart is entered at the wet-bulb temperature, and the co-ordinate is followed vertically upwards until it intersects the saturation or 100% relative humidity curve. Then the constant wet-bulb temperature line is followed until it intersects the dry-bulb temperature co-ordinate. Now the absolute humidity can be read directly, and the relative humidity can be found by interpolation between the curves for

constant relative humidity. For example, let us assume a wet-bulb temperature of 54°F and a dry-bulb temperature of 60°F. The 54°F line is followed until it intersects the saturation humidity curve at an absolute humidity of 62 grains water/pound dry air. Then, the 54°F wet-bulb temperature line is followed until it intersects the 60°F dry-bulb temperature line at an absolute humidity of 53 grains water/pound dry air. The relative humidity is found to be 70%.

The constant wet-bulb temperature lines are also valuable for determining the temperature of drying surfaces. The wet-bulb temperature is approximately the same as the *adiabatic saturation temperature*, i.e., the temperature that would be attained if the air was saturated with water vapour in an adiabatic process (a process in which there is no heat gained or lost to the surroundings). This is the case when drying is carried out with the heat supplied by the drying air only. Thus, for the example given above, a wet material being dried with air having a wet-bulb temperature of 70°F would remain at that temperature as long as there is a film of free moisture at the surface. The assumption that the wet-bulb temperature is equal to the adiabatic satura-tion temperature does not hold true, however, for non-aqueous systems.

## Humidity Measurement

The most accurate means of measuring humidity is by the *gravimetric method*. In this procedure, a known amount of air is passed over a previously weighed moisture-absorbing chemical such as phosphorus pentoxide, and the resultant increase in weight of the chemical is measured. Although accurate, the gravimetric method is cumbersome and slow. For rapid determination of humidity, temperature-measurement methods are most often used.

As indicated in the discussion of the psychrometric chart, humidity can be determined by taking two temperature measurements. The simplest instrument for this purpose is the *sling psychrometer* (Fig. 7.2). It consists of two bulb thermometers set in a frame that is attached to a swivel handle. One thermometer, the *dry-bulb thermometer*, has a bare bulb and the bulb of the other thermometer, the *wet-bulb thermometer*, is covered by a wick saturated with water. The psychrometer is whirled through the air, and the two thermometer readings are taken at successive intervals until these temperatures no longer change. In psychrometers used as permanent installations, the movement of air across the thermometer bulbs is induced by a motor-driven fan, and a reservoir is provided to keep the wet-bulb wick saturated with water. The thermometers used in such an installation may be of almost any type—mercury bulb, bimetallic strip, vapour pressure, or electrical.

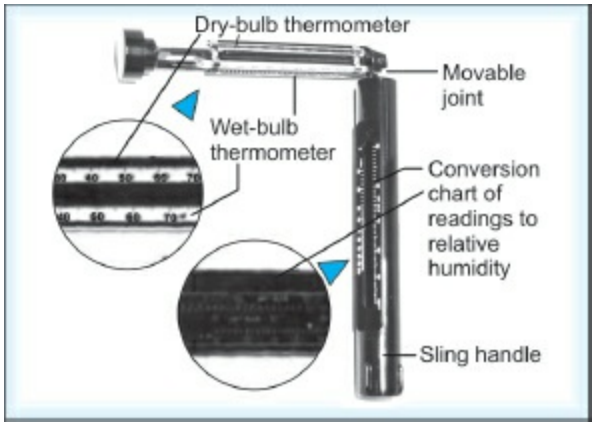



**Fig. 7.2:** Sling psychrometer

Another temperature method for determining humidity is based on the dew point instead of the wet-bulb temperature measurement. The dew point

is determined directly by observing the temperature at which moisture begins to form on a polished surface in contact with the air. The surface is cooled by refrigeration until the first fog of moisture appears.

An entirely different method for measuring humidity employs the *hygrometer*. This instrument utilizes certain materials whose properties change on contact with air of different relative humidities. The *mechanical hygrometer* uses such materials as hair, wood fiber, or plastics, which expand or shrink with changes in humidity. The moisture-sensitive element is connected to a pointer in such a fashion that a change in length causes the pointer to move across a dial calibrated in humidity units. *Electric hygrometers* measure the change in electrical resistance of moistureabsorbing materials with humidity.

## HUMIDIFICATION

Humidification is a process of increasing the moisture in the air. In pharmaceutical industry, if special conditions are not prescribed the humidity required for various operations is 20–30% RH and temperature is within 20–27°C. The purpose of humidification is to prepare air of known temperature and humidity. This is accomplished by bringing the air into contact with water under such conditions that a desired humidity is reached. If the conditions in humidifier are such that the air reaches complete saturation, the humidity is fixed. If, however, the equipment is such (and this holds for most commercial equipment) that the exit air is not quite saturated, then conditions are somewhat indeterminate. The humidity of the exit air can be fixed by varying the water temperature according to the characteristics of the specific piece of equipment on hand. By re-heating to the desired temperature, air of any desired percentage humidity and temperature may thus be obtained. The heating devices are usually coils or banks of finned tubes. The air may be brought into contact with water in a variety of apparatus such as packed towers which shower over the packing or infact any apparatus employed for gas absorption may be used. The usual method, however, is to spray water from spray nozzle into the air. This method requires less space than the tower type of apparatus. The interaction of gas, liquid and conditions of humidification are shown in Fig. 7.3. The ordinate represents temperature and humidity, whereas abscissa indicates the perpendicular distance to the interface. Broken arrow represents diffusion of vapour through the gas phase. Continuous arrow represents the flow of heat (both latent and sensible) through the gas and liquid phases. In adiabatic humidification, the air is in contact with the liquid at constant temperature. The humidity at interface ($H_i$) must be greater than the humidity ($H_y$) of the gas, so that air gets, humidified. Since water is vapourized, the latent heat is transferred from liquid to gas. The gas tempetature ($T_y$) must be higher than the interface temperature ($T_i$) that sensible heat may flow to the interface (heating or cooling without the addition or subtraction of moisture content is called sensible heat or cooling). $H_i$ and $T_i$ represent equilibrium conditions of interface (saturation condition). Latent heat flow from, liquid to gas and sensible heat flow from gas to liquid get balanced. Therefore, there is no temperature gradient in the liquid.

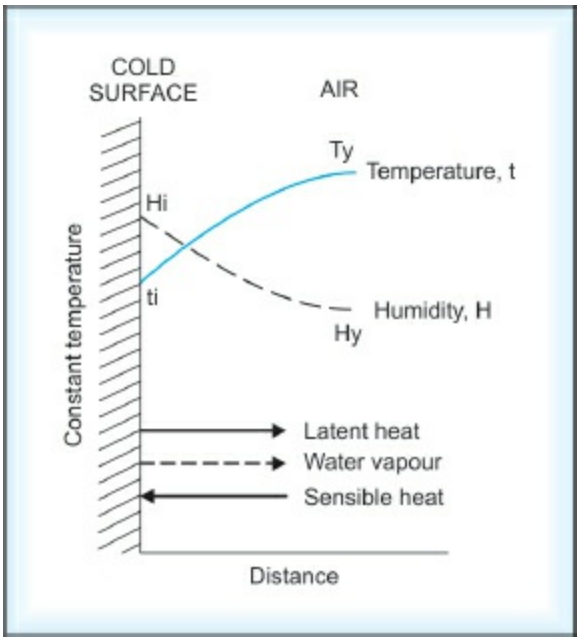



**Fig. 7.3:** Conditions of humidification

## Humidifiers

A humidifier is an equipment that increases humidity (moisture) in the air of the area where it is placed. This is accomplished by bringing the air into contact with water under such conditions that a desired humidity is reached. Various types of humidifiers include:

### *Evaporative Humidifiers*

The most common humidifier, an “evaporative” or “wick humidifier”, consists of just a few basic parts: a reservoir, wick and fan. The wick is a filter that absorbs water from the reservoir. Evaporation of water from the wick is dependent upon the relative humidity. A room with low humidity will have a higher evaporation rate compared to a room with high humidity. Therefore, this type of humidifier is self-regulating: as the humidity of the room increases, the water vapour output naturally decreases. The fan is adjacent to the wick and blows air onto the wick, thus aiding in the evaporation of the water within. Wick humidifiers trap the mineral deposits in the wick.

### *Impeller Humidifier (Cool Mist Humidifier)*

A rotating disc flings water at a diffuser, which breaks the water into fine droplets that float in the air.

### *Vapourizer (Steam Humidifier) (Warm Mist Humidifier)*

Vapourised boils water, releasing steam and moisture into the air. A medicated inhalant can also be added to the steam vapour to help reduce coughs. Vapourizers are more healthful than cool mist types of humidifiers because steam is sterile and free from minerals. Vapourizers use more electricity to boil the water. These types tend to collect minerals on or around the heating element and require regular cleaning to control build-up.

### *Ultrasonic Humidifier*

A metal diaphragm vibrating at an ultrasonic frequency creates water droplets that silently exit the humidifier in the form of a cool fog. Ultrasonic humidifiers should be cleaned regularly to avoid bacterial contamination which may be projected into the air. Impeller and ultrasonic humidifiers do

not selectively add moisture to the air, they also add any suspended material in the moisture to the air such as microorganisms and minerals. The amount of minerals and other materials can be greatly reduced by using distilled water.

## Dehumidification

Dehumidification means decreasing of humidity (removal of moisture) from the air. Dehumidification is accomplished by bringing the moist air in contact with the cold surface (liquid or solid). If moist air is to be dehumidified it might be brought into contact with a spray of water whose temperature is lower than the dew point of the entering air. This may be done in an apparatus similar to that in Fig. 7.3 except that the heater $HC_1$ is not necessary. Alternatively moist air may be passed over a bank of finned metal tubes through the inside of which a cold fluid is passed such that the outside of the tubes is below the dew point of the air so that water will condense out of air (indirect cooling).

Under such conditions, there is a layer of condensed liquid on the metal surface and the process becomes equivalent to the interaction between humid air and cold water. Figure 7.4 represents the principle of the dehumidification process.

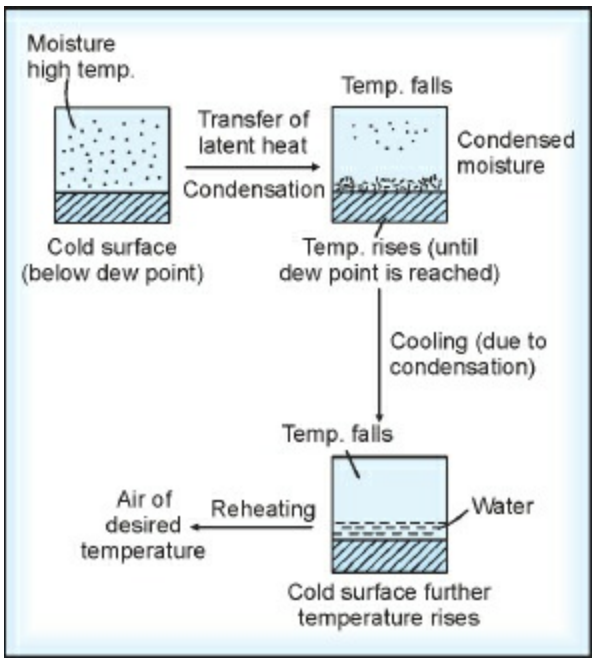



**Fig. 7.4:** Principle of dehumidification process

## Dehumidifiers

Dehumidifiers are used in special heat transfer devices to liquefy vapour by removing their latent heat. Such systems are known as condensers. In pharmaceutical industry many operations are carried out at stated humidity to get optimum results as per the regulation. In areas where air is very humid, it become difficult to carry on operation with hygroscopic and deliquescent substances even under air conditioning, hence dehumidifier are installed there. Dehumidifiers are mainly of three types.

### *Mechanical/Refrigerative Dehumidifiers*

Mechanical/refrigerative dehumidifiers, the most common type, usually work by drawing moist air over a refrigerated coil with a small fan. Since the saturation vapour pressure of water decreases with decreasing temperature, the water in the air condenses, and drips into a collecting bucket. The air is then reheated by the warmer side of the refrigeration coil. This process works most effectively with higher ambient temperatures with a high dew point temperature. In cold climates, the process is less effective. They are most effective at over 45% relative humidity, higher if the air is cold.

### *Desiccant Dehumidifiers*

A desiccant dehumidifier is a device that employs a desiccant material to produce a dehumidification effect. As they are more effective for low-temperature and low (relative) humidity levels, they are generally used for these conditions instead of mechanical/refrigerative dehumidifiers – or are used in tandem with them. Desiccant materials have a high affinity for water vapour. An example of dessicant material is silica gel. Typically, their moisture content is a function of the relative humidity of the surrounding air. Exposed to low relative humidities desiccant materials come to equilibrium at low moisture contents and exposure to high relative humidities results in equilibrium at high moisture contents. The process involves exposing the desiccant material to a high relative humidity air stream, allowing it to attract and retain some of the water vapour and then exposing the same desiccants to a lower relative humidity air stream which has the effect of drawing the retained moisture from the desiccant. The first air stream is the air that is being dehumidified while the second air stream is used only to regenerate the desiccant material so that it is ready to begin another cycle. Note that the first

air stream's water vapour content is reduced while the second air stream's water vapour content is increased. Typically, the low relative humidity air stream is air taken from any available source and heated to reduce its relative humidity. Hence, desiccant dehumidifiers consume heat energy to produce a dehumidifying effect.

In general, a desiccant dehumidifier comprises four major components:

- The component that holds the desiccant, of which there are several types;
- A fan to move the air to be dehumidified (process air) through
- A desiccant holder
- A heater to heat the air that will be used to dry the desiccant (regeneration air)
- A fan to move the low humidity air for drying the desiccant through the desiccant holder.

## *Cooling Towers*

Cooling towers are *heat removal devices* employed to transfer process waste heat to the atmosphere. They may either make use of evaporation to remove the heat and bring the working fluid at lower temperature near to the wet-bulb temperature or rely exclusively on air to cool the working fluid to near the dry-bulb temperature. The cooling that occurs in the former type is termed as "evaporative cooling". It relies on the principle that when flowing water comes in contact with moving air, evaporates. In a cooling tower, only a small portion of the water that flows through the cooling tower evaporates. The heat required for evaporation is given up by the main body of water as it flows through the cooling tower. Simply, the vapourization of this small portion of water provides cooling for the remaining un-evaporated water that flows through the cooling tower. Heat is continuously transferred from the water stream to the air stream and thus the air's temperature is raised the relative humidity to 100% and this air is discharged into the atmosphere.

As represented by Fig. 7.5, water pumped from the tower basin is the cooling water routed through the process coolers and condensers in an industrial facility. The cool water absorbs heat from the hot process streams which need to be cooled or condensed, and the absorbed heat warms the circulating water (C). The warm water returns to the top of the cooling tower and trickles downwards over the fill material inside the tower. As it trickles

down, it contacts ambient air rising up through the tower either by natural draft or forced draft using large fans in the tower. That contact causes a small amount of water to be lost as windage (W) and some of the water (E) to evaporate. The heat required to evaporate the water is derived from the water itself, which cools the water back to the original basin water temperature and the water is then ready to re-circulate. The evaporated water leaves its dissolved salts behind in the bulk of water which has not been evaporated, thus raising the salt concentration in the circulating cooling water. To prevent the salt concentration of the water from becoming too high, a portion of the water is drawn off (D) for disposal. Fresh water makeup (M) is supplied to the tower basin to compensate for the loss of evaporated water, the windage loss water and the drawoff water. A water balance around the entire system is: M = E + D + W.

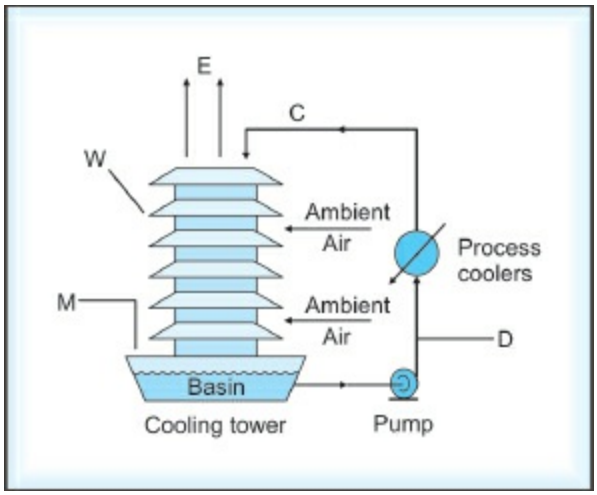



**Fig. 7.5:** Schematic representation of cooling tower

## AIR CONDITIONING

Air conditioning refers to the cooling and dehumidification of indoor air for thermal comfort. In a broader sense, it refers to the process of treating air so as to control its humidity, temperature, cleanliness and distribution simultaneously to meet the requirements of the conditioned space.

Air conditioning applications are broadly divided into *comfort* and *process. Comfort applications* aim to provide building an indoor environment that remains relatively constant in a range preferred by humans despite changes in external weather conditions or in internal heat loads. Process application aim to provide proper conditions during manufacturing, processing, and preserving the material and equipment. In process application of air conditioning, different areas were identified, for example, in granulation and tableting section air conditioning is necessary and the conditions are 45% RH and 22°C and 20% RH and 22°C, respectively. In the production of effervescent products and dry syrups, the condition should be not more than 10–15% RH and 22°C temperature. In the manufacturing of soft gelatin capsule, the temperature should be in the range of 20°C–22°C and 40% RH in the operating area and between 20–30% RH in the drying area. In case of production of all biological products (schedule C and $C_1$) as per USFDA, air conditioning is essential.

Normally, the air is allowed to reach complete saturation by providing appropriate conditions, then humidity is fixed. In commercial equipments, the conditions are somewhat intermediate and the exit humidity is not quite saturated, the desired humidity maintenance can be achieved by reheating to the air to desired temperature. Consider the following example: when air is entering the equipment, let the initial dry bulb temperature of air be p and humidity be Hj (Point A). Let the desired dry bulb temperature of air be t2 and humidity be $H_2$ at the exit (Point B). The air is first allowed to attain the desired humidity by treating with water by using the humidity chart. It is represented by Point C (wet bulb temperature $t_3$). Then it is heated to attain the desired temperature ($t_2$) at the same humidity (represented in Fig. 7.6 by ACB). In the second approach (sequence of steps shown in Fig. 7.7), air is preheated to a temperature, $t_4$. Subsequently, the air is cooled along the adiabatic cooling line, until it reaches the desired humidity. It is then reheated

so as to reach the desired temperature of $t_2$ (represented in Fig. 7.7 by path ADCB).

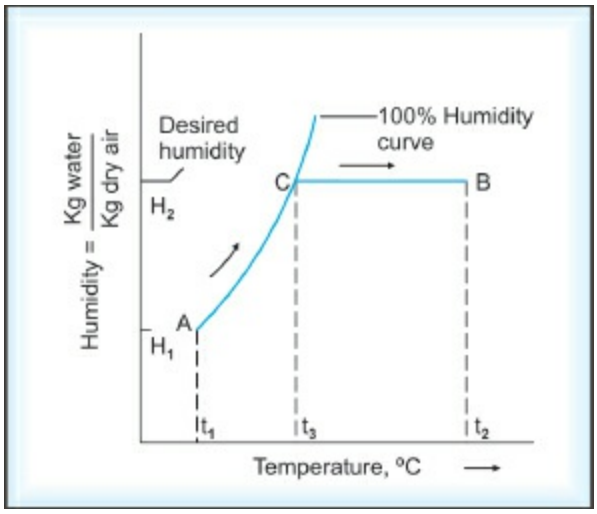



**Fig. 7.6:** Changes in humidity and air temperature for air conditioning (first approach)

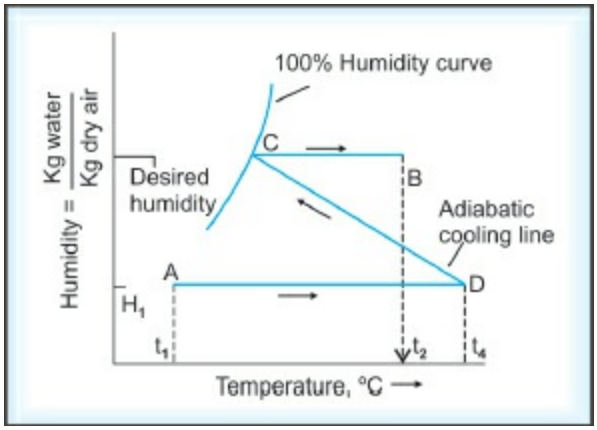



**Fig. 7.7:** Changes in humidity and air temperature for air conditioning (second approach)

Air conditioning cycle comprises of pinching the air from the space, mixing with outer ventilation air and then conditioning before being blown back into the space. The components are: *Outside air damper*, shuts down the outside air intake when the system is turned off. The damper is with a motor to open it and power failure automatically closes it. On the upper end metal mesh is located to avoid any undesired substances from entering into. *Mixing chamber*, mixes the return air from space with the outside ventilation air

*Filter*, cleans the air by removing solid dirt particles. Its position is such that it cleans the return air and ventilation air. The filter is also positioned upstream of any heating or cooling coils, to keep the coils clean. This is particularly important for the cooling coil, because the coil is wet with condensate when it is cooling; *Heating coil* raises the air temperature to the required temperature. *Cooling coil* provides cooling and dehumidification. A thermostat mounted in the space normally controls this coil. A single thermostat and controller are often used to control both the heating and cooling of coils. *Humidifier* adds moisture which is usually maintained by a humidistat in the space. In addition, a high humidity supersede humidistat is often mounted just downstream of the fan, to switch the humidification "off" if it is too humid in the duct. There is minimum possibility of condensate forming in the duct. *Fan* draws the air through the resistance of the system and blows it into the space. These components are controlled to achieve six of the seven airconditioning processes. Heating, cooling, dehumidifying, humidifying, ventilating, cleaning, and lastly the air movement which is not actually controlled by the air conditioning plant but by the way the air is delivered to the space.

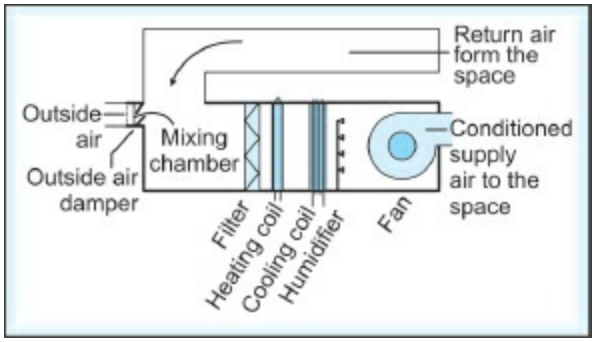



**Fig. 7.8:** Schematic representation of air conditioning cycle

## Air Conditioner

An air conditioner is an appliance, system, or mechanism designed to stabilize the air temperature and humidity within an area (used for cooling as well as heating depending on the air properties at a given time), typically using a refrigeration cycle but sometimes using evaporation, most commonly for comfort cooling in manufacturing and working areas. Modern air conditioning equipments generally fall in two classes (Fig. 7.9).

- Self-contained air conditioner (unitary or packaged)
- Central air conditioner (field erected)

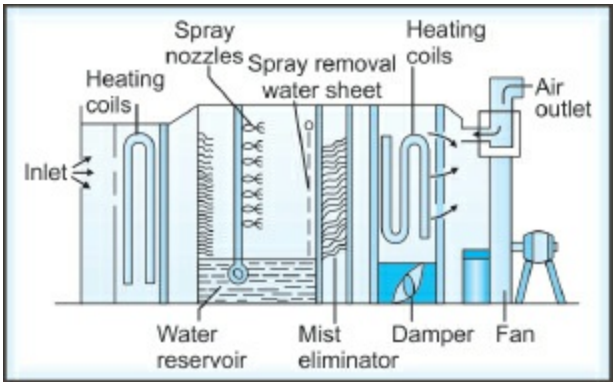



**Fig. 7.9:** Air conditioning unit and control system

***Self-contained air conditioner*** includes window mounted or wall bracket conditioner. Most of these units are air cooled; though water cooled types have also been made. Room air enters the casin at front panel. It is mixed with a part of the outdoor air. The mixture is then forced over the cooling coils by centrifugal fan. Some of the moisture in the air is condensed and deposited by means of a single ring of propeller type condenser. The units are hermetically sealed.

In ***central air conditioner***, air cleaning is usually provided using filters, which can be cleaned for reuse and disposable. Cooling is achieved using either water or by direct expansion in refrigerated coil or air washers. Steam or hot water coil are used for heating purpose. Humidification system may be provided by surface-type water nozzles, steam humidifiers or sprayed coil.

## PHARMACEUTICAL WATERS

Water acts as a solvent for a wide range of substances. Environment Protection Agency (EPA) has given water quality criteria for copious pollutants. Drinking water, which is of major issue under EPA regulation and which is delivered by local municipal system or haggard by private reservoirs, is the preliminary material for all forms of water envisaged by Pharmacopeial monographs. It may be used in the preparation of USP drug substances but not in the preparation of dosage forms or reagents or test solution. There are varieties of grades of water to be used for pharmaceutical purposes. Principally because of its solvent properties and physiological inertness, water is a significant pharmaceutical agent. The level of purity required varies widely with the application. For instance, the one used in the steam boilers only needs to be salt free or in other words softened to avoid scaling and deaerated to avoid corrosion in the system. On the other hand, water used in injectables has to meet much more rigorous quality requirements. The US Pharmacopeia (USP) encloses specifications for several grades of water used in the preparation of medicinal products. These types of water are known as “compendial” because their quality is specified in a nationally recognized standard such as USP. In addition, many companies utilize various non-compendial water systems designed for specific needs. While designing the system for water, it is important to decide on the proper grades to use in various applications. Compendial waters are characteristically very expensive, not only because of the required treatment steps, but also because of extensive validation and testing requirements. Therefore, it is cautious to limit their use to those processes where water becomes the component of the pharmaceutical product, comes in direct contact with the product and last but not the least, is used for the final rinsing of the decisive process equipment. For other applications, various non-compendial grades of water which include potable, softened, deionized, etc., are effectively used without any regulatory conflicts. Other than compendial and non compendial, water can be classified on the basis of the site of production. These includes two types: *Bulk waters*, which are typically produced on the site where they are used; and *packaged waters*, which are produced, sterilized and finally packaged to preserve their microbial quality throughout their packaged shelf life. Exceptionally, there are several specialized types of packaged waters, differing in their designated

applications, packaging limitations, and other quality attributes. These are characteristically produced in large volumes by a multiple unit operation water system and are distributed by a piping system for use at the same site. These pharmaceutical waters meet the quality attributes as specified in the related monographs. Bulk water may further include *potable water*, when, on the whole, process involves the elimination of insoluble matter through appropriate settling, congulating and filtering processes; obliteration of pathogenic microorganisms by chlorination, or other methods; and enhancement of palatability through aeration and filtration through charcoal. The water so obtained is referred to as potable. It is used for drinking, washing, cooling and manufacturing of APIs, except for final stage of production. Precisely, “potable water” means water that meets the standards for drinking purposes of the state or local authority having jurisdiction or water that meets the quality standards prescribed by the US Environmental Protection Agency’s National Interim Primary Drinking Water Regulations, published in 40 CFR Part 141.

Compendial water including different monographs: purified water, water for injection, water for hemodialysis, packaged monographed water, bacteriostatic water for injection, sterile water for irrigation and sterile water for inhalation are described here in detail.

## *Purified Water*

Water subjected to simple purification processes like distillation, ion-exchange (deionized, demineralised), reverse osmosis, or other methods and moreover, contain no added substances are referred to as simply “purified”. It is prepared through potable water which meets EPA standards. The main objective is the exclusion of dissolved solids. Ion-exchange and reverse osmosis are successful to attain this objective by removing electrolytes, whereas, distillation is not equally effective in the removal of weak electrolytes and non electrolytes if they are volatile. But it can render the water sterile and pyrogen-free if repeated number of times. The requirements for Purified Water can be as follows:

- *Low conductivity* (high resistivity): This test summarized as, water contains a minimal amount of ions such as calcium, magnesium, sodium, iron, chloride, sulfate, etc. The intrinsic presence of hydrogen and hydroxide ions establish the theoretical limit of Purified Water

conductivity: approximately 0.05 *m*S/cm (resistivity 18 MO cm) at a pH of 7.0. The practical limits specified in the USP are in the range of 1–5 *m*S/cm, depending on the pH and temperature.

- *Low TOC* (less than 500 ppb): Water shall contain a minimum amount of organic compounds. Such compounds are disagreeable for two main reasons—they may be toxic, and/or they may serve as sources of nutrition for microorganisms.
- *Low microbial count:* Water shall contain a minimal amount of viable microorganisms, including spores.

The current edition of USP has set up the following requirements for USP purified water:

- Purified water is prepared from water complying with the federal regulations for drinking water.
- Purified water contains no added substance.
- Purified water is obtained by a suitable process.
- Its conductivity does not exceed a set level (1.3 *m*S/cm at 25°C).
- Its total organic carbon (TOC) does not exceed set level (500 ppb).

Purified water is used as an excipient in nonparenteral formulations such as tablets, capsules, creams, lotions, etc. and in other pharmaceutical processes, such as cleaning of certain equipment and non-parenteral product-contact components.

## *Water for Injection (WFI)*

For parenteral formulations, water used is the one which meets stringent quality criterias and is known as WFI in official compendias. It is used as an excipient in the preparations where product endotoxin content must be controlled, and in other pharmaceutical processes, such as cleaning of certain equipments and parenteral product-contact components. WFI is usually prepared by distillation in a still purposely designed to produce the desired high-quality water. The specification of still includes that it should (1) carry out prepurification of feed water by chemical softening, deionization, or filtration to improve the quality of the distillate and reduce the frequency of required cleaning due to insoluble scale in the boiler, (2) remove entrained contaminants from the vapour before these are condensed by way through an efficient baffle system, (3) eject volatile constituents from the top of the

system before the vapours are cooled so that they do not redissolve and appear in the condensate, (4) have all surfaces constructed in such a way that when it comes in contact with the vapour and condensate of a material, it will not dissolve in even trace amounts, preferably pure tin, 304 stainless steel, or borosilicate glass. Other than *conventional stills*, two other types of stills frequently used for the production of large volumes of water are the *vapour compression stills* and the *multiple effect stills*. Other than distillation, USP has now approved reverse osmosis and the process is gaining increased attention and use. Reverse osmosis as the name indicates involves movement "up" the concentration gradient and system functions by applying pressure (usually, 200 to 400 psi) to raw water sufficient to force the permeation of water through a semipermeable membrane in the opposite direction from natural osmosis. The most commonly used membranes are made up of cellulose esters or polyamides (nylon) and are effectual in retaining all macromolecules and 85% or more of small ions such as $Na^+$ and $Cl^-$. Pyrogens being the macromolecules are retained along with the viable particles as microorganisms. Greater efficiency and consistency are achieved by allowing the water to pass through two membranes in series. The acceptance of this system for the preparation of WFI is increasing with time as its characteristics are understood explicitly. Apart from the processes, USP establishes the following requirements for WFI USP. It should

- meets all the requirements for Purified Water.
- is obtained by distillation or reverse osmosis and contains not more than 0.25 USP endotoxin units per milliliter.
- is prepared using suitable means to minimize microbial growth.

Allowable concentrations of micro organisms in the official monograph are not specified, but it is recommended that purified water should have not more than 100 colonyforming units per ml (CFU/ml) and WFI should have not more than 0.1 CFU/ml.

*Storage and distribution of WFI:* It is as important as its production. A closed system is enviable, with air exchange through a filter that removes microorganisms, vapours and dirt from the air as the tank is filled and emptied. Moreover, if microorganisms are able to gain entry into the tank, they can be prevented from multiplying by maintaining the temperature of the water at 80°C using steam coils at the bottom of the tank. Normally, WFI

should not be held for more than 24 h at room temperature before it is used, but if held at 80°C, continuous addition of fresh WFI as usage occurs is a common practice. The persistent threat by microbial contamination, in spite of precautions, and subsequent development of pyrogenic substances in the water demand vigilant storage requirements. The distribution of WFI from the storage tank to the place of use may be by direct withdrawal from the tank, or in large plants, through a duct system. When a duct system is used, caution must be taken to prevent contamination counting, construction with welded stainless steel pipe, a closed system with continuous circulation to avoid stoppage, upholding at elevated temperature, complete isolation from all other ducting systems, exclusion of elbows and other pockets in which water can stop for longer time, and a means of scrupulous cleaning and sanitation at recurrent intervals with clean steam or hot alkali.

## *Water for Hemodialysis*

The water used for hemodialysis applications also requires equivalent stringent steps. It is produced at the site of use and is made from EPA drinking water which has been further purified to reduce chemical and microbiological components. It is packaged and stored in unreactive containers that prevent bacterial entry.

## *Packaged Monograph Waters*

They are packaged forms of either purified water or WFI that have been sterilized to preserve their microbe free properties. These waters may have specific projected uses as designated by their names and may also have precincts on packaging configurations depending upon the use. Many-a-times, these packaged waters are used in lieu of the bulk form of water from which they were derived. Packaged monographed waters include *sterile purified water* and *sterile WFI*, packaged and sterilized. Sterile purified water is used mainly in the preparation of non-parenteral compendial dosage forms or in analytical applications demanding purified water where use of a validated purified water system is not practical and only a relatively small quantity is required to be sterile or where bulk-packaged purified water is not suitably microbiologically controlled. Sterile WFI is used for extemporaneous prescription compounding and as sterile diluent for parenteral products. It may also be used for applications where bulk water for injection or purified water is indicated but where access to a validated water

system is either not practical or where only a relatively small quantity is needed. Sterile WFI is packaged in single dose containers not larger than 1L in capacity.

## *Bacteriostatic WFI*

Bacteriostatic WFI is sterile water for injection in which one or more suitable antimicrobial preservatives has been added. It is intended to be used as diluent in the preparation of parenteral products, most typically for multidose products that require repeated content withdrawals. It may be packaged in single-dose or multiple-dose containers not larger than 30 ml.

## *Sterile Water for Irrigation*

It is WFI packaged and sterilized in singledose containers of larger than 1 L in size that allows rapid delivery of its contents.

## *Sterile Water for Inhalation*

It is WFI that is packaged and made sterile, and is intended to be used in inhalators or in the preparation of inhalation solutions. It carries a less rigorous specification for bacterial endotoxins than sterile WFI, and therefore, is not suitable to be used for parenteral applications. It is thus categorized separately here.

## PHARMACEUTICAL GASES

Inert gases have been used to dislodge oxygen from a solution and reduce the possibility of oxidation of the formulation. For instance, sodium bicarbonate injection decomposes, particularly during autoclaving, to produce sodium carbonate, carbon dioxide, and water. If the solution is saturated with carbon dioxide the same can be prevented. The gases have many other applications as in sterilization. “Gaseous sterilization procedures refer to the exposure of items to a chemical sterilant in a gaseous phase in a manner to ensure that a predetermined lethal effect is conveyed to the items being processed with the objective of eliminating the pretreatment bioburden and providing a desired margin of safety”. The details of gaseous sterilization are explained in Chapter 22 of this book. Here, some general principles of validation and procedure to be followed for gaseous sterilization are discussed. As with any sterilization procedure, the use of cGMP and control of the manufacturing environment are prerequisites in the maintenance of barriers against microbial contamination. To get utmost advantage, gaseous sterilizations validation should be carried out early in the development of a new or modified procedure, or new drug or device. It is accentuated that sample-based, end product testing does not guarantee a high quality product but can only efficiently identify and statistically quantify the incidence of substandard product. The gaseous sterilization procedure in conjunction with the manufacturing process must be designed to maximize the probability of the product being suitable for its intended use. Before a gaseous sterilization procedure’s program is initiated, the following items should be in hand:

1. Product specifications, in terms of physical, chemical, microbial and pharmacological properties, where appropriate.
2. Specifications for raw materials and components.
3. Determination of the required Sterility Assurance Level (SAL) based on the use of the items being treated.
4. Compatibility of the process with the items to be treated.
5. Determination of acceptable limits of the major residues after gaseous sterilization procedures.
6. Validation of analytical methods used with adequate calibration and qualification of the measuring equipment.

In some circumstances, a satisfactory demonstration that the gaseous sterilization procedure consistently produces the desired SAL may eliminate the need to test every batch of the items being treated for sterility.

Validation of gaseous sterilization procedures involves the following elements:

- Bioburden
- Manufacturing area environment
- Determination of time and humidity in the preconditioning area
- Determination of temperature, pressure, time and humidity in the chamber
- Load after sterilization
- Loading patterns
- Biological indicator survival
- Vendor certification (if the gaseous sterilization treatment is carried out by an external contractor).

Before the commencement of studies, a written procedure shall be established to prevent unconstitutional changes in the protocol and process, and control change during any phase of the studies until all relevant data is evaluated. The protocol should be written in accordance with the validation. The protocol should give details of the process objectives in terms of the type of material, including any packaging; chamber content and probability of survival (or microbiological safety index, MSI) desired from the process; pre-established details and range for the process parameters which may include pre-humidification time and relative humidity, temperature, loading pattern, relative humidity, partial pressure of gas, etc. a description of all the equipment in terms of type, model, operational range and characteristics; the performance features of each system, sub-system or equipment described which should include pressure gauge accuracy and sensitivity, valve operations, alarm system functions, timer responses and accuracy, gas and water vapour flow rates and pressures, jacket cooling; the methods and procedures for monitoring the performance of equipment systems and sub-systems and the process; the training and qualification of personnel responsible for the evaluation and certification of each stage of the protocol and for the final approval of the entire validation procedure.

Apart from sterilization, gases are also used directly in dosage forms as in aerosols. Although this subject, including the use of so called liquid propellants, is covered elsewhere, note that pressure packs often use nitrogen, nitrous oxide, or carbon dioxide to expel the contents from their containers. The latter two gases are much more soluble in water, so some aeration of the material discharged will take place. Carbon dioxide is six times as soluble in water as nitrogen, and nitrous oxide is about four times as soluble as nitrogen. Thus, if it is intended to have some of the gas dissolved in the product, either nitrous oxide or carbon dioxide may be used. In organic solvent nitrous oxide is somewhat more soluble than carbon dioxide. There is not a great deal of difference in solubility properties however, the possibility exists that the pH-lowering effect of carbon dioxide as it forms carbonic acid may be just as objectionable, as it precipitates carbonate in alkaline product.

Gases are also used for artificial atmosphere and five of them including nitrogen, oxygen, helium, carbon dioxide and nitrogen oxide are official. Nitrogen is used as a diluent for oxygen and may be used as a protective atmosphere for easily oxidizable products. Helium, because of its low density as compared to nitrogen, is used in the preparation of a gaseous mixture composed of 20% oxygen and helium. This mixture is used to ease respiration difficulties. Because of low solubility of helium in blood, the same mixture may be used as an atmosphere for those working under high atmospheric pressures (deep sea divers, caisson workers). When ordinary air is used, rapid decompression causes nitrogen bubbles to be formed in blood, resulting in painful, and sometimes fatal conditions known as bends. Oxygen is used when respiratory problems exist. Ordinarily, it is diluted with nitrogen or helium. Nitrogen (I) oxide usually requires 20–25% oxygen during administration. It is used for surgical operations of short duration.

## MATERIAL FOR PLANT CONSTRUCTION

Materials required in the fabrication of equipment and associated systems dedicated to the production of pharmaceutical products should be capable of withstanding the temperature, pressure and chemical corrosiveness of the product to ensure its purity and integrity. They must also meet a high degree of cleanliness to convey sterile and non-sterile products or solutions, particularly when it applies to their contact surfaces. Therefore, the selection of a material for construction of equipment and plant depends on properties, such as chemical resistance, structural strength, resistance to physical shock (operating pressure), resistance to thermal shock (operating temperature), ease of fabrication, ease of cleansing and its initial and maintenance cost. Various materials used for pharmaceutical plant construction are described in Fig. 7.10.

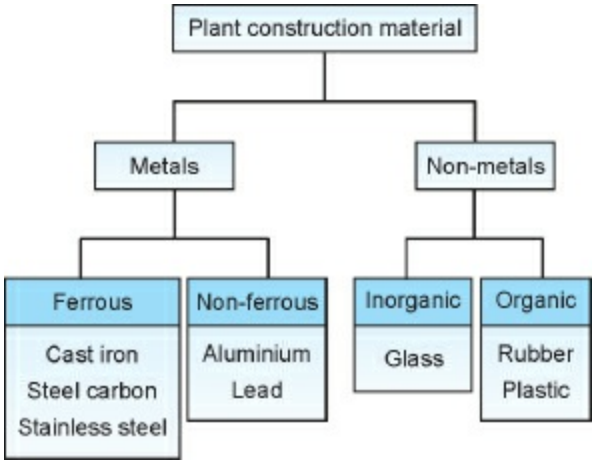



**Fig. 7.10:** Classification of materials used for pharmaceutical plant construction

## Ferrous Metals

High mechanical strength, lower cost and abundant availability makes the ferrous metals one of the most widely used materials for pharmaceutical plant construction. Three different types of ferrous metals are used, i.e. cast iron, steel carbon and stainless steel.

### *Cast Iron*

This consists of iron with a high proportion of carbon (beyond 1.5%). The amount of carbon varies, giving products with different properties. Cast iron is resistant to concentrated sulphuric acid ($H_2SO_4$) and nitric acid ($HNO_3$), and to dilute alkalis. At high temperatures, it is attacked by ammonium salts and reacts with tannins. It is commonly used as support for plant, jackets of steam pan, etc. It is cheap and is therefore, often used in place of more resistant and more expensive materials by coating with plastic, enamel or other protective substances. Its characteristics may also be altered by alloying with other elements such as nickel (Ni), chromium (Cr) and silicon (Si). Different types of cast iron alloys are detailed in Table 7.1.

**Table 7.1:** Properties and applications of ferrous metals used as plant construction material

| Types | Variants | Composition | Properties and applications |
|---|---|---|---|
| Cast iron | Gray cast iron | Carbon, silicon | Low cost, easy to cast, brittle, poor resistance to shock, used in impellers, balls and dies |
| | Malleable iron | White cast iron, carbon-2.5% | Corrosion resistant, type-I easy to machine and type-II less easy to machine |
| | High silicon cast iron | Silicon 13–16%, cast iron | Resistant to corrosion and oxidizing and reducing environment. Not easily machined and welded, used in sulphuric acid services |
| | Nickel resistant cast iron | | Superior toughness, resistant to impact, heat and corrosion, easy to weld and machine, non-resistant to oxidizing and alkaline environment |
| Carbon steel | Low alloy steel | Low conc. of Fe, C, Ni, Cr and Mb. | High mechanical strength, corrosion resistant, used for thin plant walls |
| Steel or mild steel | Alloy with nickel | Ni | High hardness and corrosion resistant |
| | Alloy with silicon, chromium | Ni, Cr | Resistant to abrasion, corrosion and oxidation |
| | Martensitic (type 410) | Cr = 12–20%<br>C = 0.2–0.4%<br>Ni = up to 2.0% | Resistant to mild corrosion, atmospheric and organic exposures, have poor ductility, used in sinks, bench tops, mixing elements, etc. |
| Stainless steel | Ferritic (type 430, α-form) | Cr = 15–30%<br>C = 0.1%<br>Ni = nil | Resistant to moderate corrosion, oxidation and high temperature (up to 800 RC, easy to machine, non-resistant to reducing agents and HCl, used in baffles, heat exchangers, furnace parts, pump shafts, etc. |
| | Austenitic (γ-form) | Cr = 13–20%<br>C = 0.1%<br>Ni = 6–22% | Highly corrosion resistant, sterilizable, non-magnetic, easy to weld and clean, but not easy to machine, used in fermentors, storage vessels, evaporators, etc. |

## *Steel Carbon or Mild Steel*

It is an iron alloy that contains only a small percentage of carbon. It has much less brittleness and much greater mechanical strength than cast iron. Unlike cast iron, it can be worked by welding and machining. It is used for supporting structures like girders and bases for plant vessels and small articles like nuts and bolts. Different types of carbon steel alloys are detailed in Table 7.1.

## *Stainless Steel*

It is an alloy of iron with nickel and chromium which makes it corrosion resistant. Stainless steels are uniquely qualified for bioprocessing applications not only because of their long service life, fabricability and availability, but

also because they are non-contaminant, noncorroding, can be polished to very smooth finishes, are strong and rigid, can withstand heat and chemical sterilization treatments, and are easily welded. Stainless steel alloys are mainly categorized into three groups, i.e., martensitic, ferritic and austenitic, which are detailed in Table 7.1. Out of these, the austenitic 18–8 (18% chromium, 8% nickel) group, also called as “18/8 stainless steel,” is the industry’s workhorse, which includes types 304 and 316 and their L or low carbon content variations. With concerns for higher corrosion resistance, the new and better low-carbon superaustenitic stainless steel AL-6XN (6% molybdenum) and the nickel-based alloys, Hastelloy B and Hastelloy C, are becoming notoriously important in the fabrication of vessels, piping, tubing, and fittings. Finally, the cast stainless alloys such as CF-8 (similar to 304), CF-8M (similar to 316), and CF-3M (similar to 316L), utilized in pumps, various types of valves (particularly ball type), and fittings, also occupy a prominent position in the industry.

## Non-ferrous Metals

### *Aluminium*

Cheap, light in weight, easily fabricable, easy to clean and maintain, highly resistant to industrial fumes, vapours and salt waters, has poor mechanical strength above 150°C, and used in heat transfer applications and meat storage.

### *Lead*

Cheapest, has low melting point, poor structural qualities, high coefficient of expansion, and is used in collapsible tubes with internal linings for fluoride toothpaste, cold water pipes, waste pipes, etc. Alloy with Ag and Cu improves resistance to corrosion, creep and fatigue. Alloy with Sb, Ti and As provides strength.

## Inorganic Non-metals

### *Glass*

The detailed description of glass is described in Chapter 27. Different types of glass are detailed in Table 7.2.

**Table 7.2:** Properties and applications of glass used as plant construction material

| Variants | Composition | Properties and applications |
|---|---|---|
| Soft glass (soda glass) | Sodium and calcium silicate | Used in glass bulbs and window glasses |
| Hard glass (potash glass) | Potassium and calcium silicate | Acid resistant, used in glass apparatus |
| Flint glass (potash lead glass) | Potassium and lead silicate | Used in optical instruments |
| Jena glass | Zinc silicate and barium borosilicate | Acid and alkali-resistant, used in laboratory glassware |
| Pyrex glass | Silicon dioxide, boron oxide, sodium oxide and small amount of K, Ca and Mg | Heat resistant, used in laboratory glassware and reactor vessels |
| Quartz glass | Pure silica | Resistant to heat (withstand temperature up to 1000°C), used in silica crucibles, distilled water stills |

### *Glassed Steel*

Resistant to corrosion, acids (except HF and hot conc. $H_2SO_4$), brittle, gets damaged by thermal shock, used in pipings when transparency is required and for handling of strong acids, alkalis and saline solutions.

## Organic Non-metals

### *Plastics*

ASTM defines a plastic as ‘a material that contains, as an essential ingredient, one or more organic polymeric substances of large molecular weight, is solid in its finished state, and, at some stage in its manufacture into finished articles, can be shaped by flow’. Despite of the many advantages of stainless steel, the wide availability of thermoplastics and fluorinated plastics offering lower initial costs, less weight, complete resistance to corrosion, elimination of the passivation process, and extremely smooth internal surfaces have increased their application, particularly in systems such as purified water distribution loops. The detailed description of plastics is described in Chapter 27. Table 7.3 lists different types of plastics used as material for plant construction.

**Table 7.3:** Properties and applications of plastics used as plant construction material

| Name | Properties | Applications |
|---|---|---|
| PVDF | Strong, resistant to abrasion and permeation of gases, have low surface and joint extractables | Used in purified water distribution systems. Provide wide temperature service range (40–284°F) and cleanability at elevated temperature |
| PTFE (Teflon) | Resistant to chemicals, high impact strength, low surface energy, low coefficient of friction, elastic recovery and self-lubricating properties. Has useful temperature limit (200–500°F). | Used in sanitary seals, O-rings, valve diaphragms, mechanical seals and rotary valves. Used as a liner or coating material for valves and pumps |
| PFA resin | Melt-processible, resistant to heat and chemical, least affected by contaminants and biofilm growth | Suitable for tubing, used in preparation of WFI |

| | | |
|---|---|---|
| Polypropylene | Widely used thermoplastics, most economical, resistant to heat, stiff, resistant to caustics, solvents, acids and organic chemicals. Possess high heat deflection temperature (195–240°F) | Suitable for laboratory and industrial drainage pipes |

## Rubber

They are widely used as lining materials. A number of synthetic rubbers have been developed, and while none has all the properties of natural rubber, they are superior in one or more ways. The isoprene and polybutadiene rubbers are duplicates of natural rubber. Natural rubber is resistant to dilute mineral acids, alkalies and salts, but oxidizing media, oils, benzene and ketones attack it. Hard and soft rubbers are frequently used for handling acids, especially dilute aqueous solutions. They cannot resist oxidizing agents and are swelled by organic solvents. Soft rubber is used as a lining for steel while hard rubber is used alone. Hard rubber is made by adding 25% or more of sulphur to natural or synthetic rubber and as such is both hard and strong. Synthetic rubbers, are becoming of increasing importance due to their superiority to natural rubber in many properties, such as resistance to oxidation, solvents, oils and many chemicals. Different types of synthetic rubbers are detailed in Table 7.4. For further details *refer* Chapter 27.

**Table 7.4:** Properties and applications of rubbers used as plant construction material

| Name | Properties | Applications |
|---|---|---|
| Neoprene (polychloroprene) | Stable at high temperature, resistant to attack by sunlight, oils, aromatic or halogenated solvents, does not burn readily | Used as insulating material in electric cables, rubber stoppers, cap-liners, dropper assemblies for eye drops, etc. |
| Butyl rubber dried | Resistant to mineral acids and alkalis (except $HNO_3$ and $H_2SO_4$), have low water vapour permeability | Used for closures of containers of freeze-products |
| Silicon rubber (polysiloxanes) | Resistant to high and low temperatures and attack to aliphatic solvents, oils and greases | |
| Polyisoprene | Translucent, flexible and stable at high temperatures | |

# 8:
# Dissolution

Dissolution is a physicochemical process by which a solid substance enters the solvent phase to yield a solution. Dissolution (release of a drug from the dosage form) is a key prerequisite for any orally-administered drug to be systemically effective. In the case of oral drug products, the release properties are mainly influenced by disintegration of the solid dosage form, into granules deaggregation of granules to yield fine particles and dissolution of drug from the fine particles into solution. Both dissolution and disintegration are parameters of prime importance in the product development strategy, with disintegration usually reflecting the effect of formulation and manufacturing process variables, and the dissolution reflecting not only the effect of processing and formulation but also influence the solubility and particle size, which are largely properties of the drug raw material. From the dosage form perspective, dissolution of the active pharmaceutical ingredient, rather than disintegration of the dosage form, is often the rate-determining step in presenting the drug in solution to the absorbing membrane.

Prior to the emergence of dissolution tests as being official in the pharmacopoeias, disintegration tests were the only official in vitro tests used by major pharmacopoeias throughout the world as means of in vivo release predictability and product performance.

Like disintegration testing, dissolution tests do not prove conclusively that the dosage form will release the drug in vivo in a specific manner, but dissolution does come one step closer, in that it helps establish whether the drug can become available for absorption in terms of being in solution at the site of absorption.

Initially, dissolution tests were introduced to characterize the release profiles of low-solubility (< 1%) drugs in aqueous media. But now, the emphasis is to adopt dissolution tests in monographs of all oral solid dosage forms with minor exceptions (e.g. non-absorbed drugs). Not surprisingly, any report in literature on formulation and development of any solid dosage form starts with dissolution testing.

The importance of dissolution testing increased steadily and of the 674 monographs for oral tablets and capsules listed in USP 28, 582 include dissolution testing. USP 28 also includes dissolution tests for suppositories, granules and suspensions, and drug-release tests for 47 modified-release tablets and capsules and two transdermal delivery systems.

With evolution and advances in dissolution testing technology, and the understanding of scientific principles and mechanism of test results, a clear trend has emerged, wherein dissolution testing has moved from a traditional quality control test to a more valuable tool to guide formulation development, assess product quality, monitor the manufacturing process, and in some cases predict in vivo performance of solid oral dosage forms and as a surrogate measure for bioequivalence (BE) and to provide biowaivers. Dissolution profile comparison has additionally been used extensively in assessing product sameness, especially when postapproval changes are made.

This chapter focuses first on the mathematical concepts of dissolution along with influence of physicochemical properties on the dissolution of drugs, and then describes all the compendial dissolution apparatuses, their modifications and their use for dissolution testing of many types of conventional and novel dosage forms. The last section of the chapter provides some emphasis on the role of dissolution testing in the regulation of pharmaceuticals.

## DISSOLUTION MECHANISM

Drug dissolution is a multistep process involving heterogenous reactions at solid-liquid phase boundaries. The heterogenous reactions that constitute the overall mass transfer process are considered to take place in two steps: (i) convective transport of the soluble solid through hydrodynamic boundary layers surrounding solid-liquid interphase to the bulk phase (diffusion-limited model), and (ii) a reaction at the solid-liquid interphase, the so called interfacial transport (reaction-limited model).

The slower of these steps exercises a dominating influence upon the rate of dissolution. In dissolution, the interfacial transport is virtually instantaneous and so, the rate of dissolution will more frequently be determined by the rate of the slower step of diffusion of dissolved solid through the hydrodynamic boundary layers.

The flow close to the solid surface can be separated into two main regions. Within the bulk flow region viscosity is negligible, whereas near the surface a small region exists that is called the hydrodynamic boundary layer. In this region, adherence of molecules of the liquid to the surface of the solid body slows them down. The hydrodynamic boundary layer is dominated by pronounced velocity gradients within the fluid that are continuous, and does not, as is sometimes purported, consist of a “stagnant” layer. According to Newton’s law of friction, pronounced velocity gradients lead to high friction forces near the surface of a solid particle. The hydrodynamic boundary layer grows further downstream of the surface since more and more fluid molecules are slowed down. The hydrodynamic boundary layer governs the overall mass transfer.

## Diffusion Limited Model or Film Theory

The rate of diffusion is described by *Fick's law of diffusion* expressed as:

$$dC/dt = k \Delta C \ldots (1)$$

where, $k$ is rate constant ($s^{-1}$) and $C$ is concentration of solid in solution at any point and at time $t$.

The first dissolution experiments were conducted by *Noyes and Whitney* (1897) and found that the dissolution rate ($dC/dt$), is a linear function of the difference between the bulk concentration ($C_b$) at time $t$, and the saturation solubility ($C_s$). This statement can be formulated mathematically as follows:

$$\frac{dC_b}{dt} = k_d(C_s - C_b) \qquad \ldots (2)$$

where, $k_d$ is the dissolution rate constant.

Later on, *Nernst and Brunner* (1904) showed that $k_d$ is a composite constant being proportional to the diffusion coefficient $D$ and the surface area of the dissolving body, $S$. Thus equation (2) was modified to what is known as Nernst and Brunner equation:

$$dC/dt = D S (C_s\text{-}C_b)/V h \ldots (3)$$

where, $h$ designates the thickness of hydrodynamic boundary layer, and $V$ is the volume of the dissolution medium.

In 1931, *Hixon and Crowell* makes the equation (2) applicable to dissolving compact objects by expressing the surface (S) of equation (3) with respect to the weight (w). By this consideration, Eq (2) when integrated yields *Hixon-Crowell* or *Cubic root law* which relates time to the cubic root of weight:

$$w_0^{1/3} - w^{1/3} = k_2 t \qquad \ldots (4)$$

where, $\omega_0$ is the initial weight and $k_2$, a constant.

According to the equations of Noyes and Whitney, and Nernst and Brunner, the dissolution rate depends on a small fluid 'layer' called the hydrodynamic boundary layer, adhering closely to the surface of a solid particle that is to be dissolved.

The above approaches can be categorized as various expressions of the diffusion layer model as a physical explanation for dissolution process, where the limiting step has been considered to be the diffusion of molecules through a stagnant film of liquid around the solid surface. The diffusion limited model has been exclusively based on experiments in rotating or stationary disk apparatus or flow through cells under well defined hydrodynamic conditions.

Departures from the diffusion layer model have been reported in powder dissolution experiments in the USP paddle apparatus and in gastrointestinal absorption studies. In these cases, the dissolution rate does not seem to be proportional to the difference between the concentration of dissolved drug and the saturation solubility, as suggested by the diffusion layer model.

## Reaction-limited Model

Dissolution mechanism based on interfacial transport is so called reaction-limited model. The important difference between the reaction-limited model and the diffusion layer model is that while in the latter the solubility is a thermodynamic parameter and drives the dissolution rate, in the former solubility (Cs) is the result of chemical equilibrium. In a reaction-limited model, dissolution is considered to be a reaction between the undissolved species and the dissolution medium molecules. The rate of dissolution is therefore driven by the concentration of the undissolved species and solubility is considered to be the concentration when the reaction equilibrium is reached.

The model exhibits rich behavior mimicking the profiles obtained by the classic equations but also shows additional flexibility, which may prove useful in dissolution curve fitting. The assumptions utilized by the model may be particularly applicable to in vivo conditions, potentially explaining the observed variability and deviations from the diffusion layer model principles. The approach could also be applicable for surfactant facilitated dissolution of drugs.

Interfacial barrier and Danckwert models, which describe the heterogeneous features of drug dissolution, are the two approaches which can be considered as reaction-limited dissolution models since both do not rely on the premises of the diffusion principles.

### *Interfacial Barrier Model (Limited Solvation Theory)*

The Interfacial Barrier Model (limited solvation theory) proposed by Wilderman (1909) considered that interfacial transport, rather than diffusion through the film, is the limiting step due to a high activation energy level for the former. According to this model, an intermediate concentration can exist at the interface as a result of solvation mechanism and is a function of solubility rather than diffusion. This model has not been studied in detail and an explicit mathematical description for the dissolution kinetics is not available.

However, the rate of drug dissolution ($G$) that is controlled by the interfacial reaction was reluctantly expressed as:

$$G = K\,(C_s - C_b) \ldots (5)$$

Where, $K$ is the effective interfacial transport constant.

## *Danckwert Model*

The next model for reaction-limited dissolution is Danckwert Model which appeared in 1951. According to this model, constantly renewed macroscopic packets of solvent (eddies) reach the solid surface and absorb the molecule of solute delivering them to the bulk solution. The concept of Danckwert model for drug dissolution is shown in Fig. 8.1 and is expressed by the equation:

$$dC/dt = A\,(C_s\text{-}C_b)\,\sqrt{\gamma D/V} \ldots (6)$$

where, γ is the rate of surface renewal.

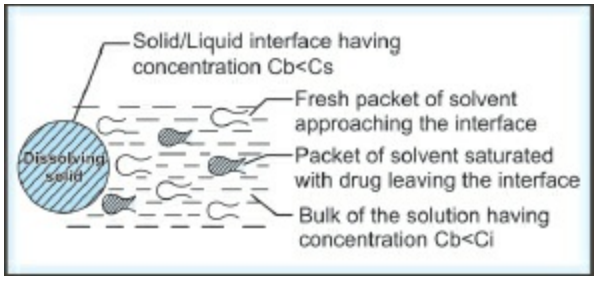



**Fig. 8.1:** Danckwert model for drug dissolution

## FACTORS AFFECTING DISSOLUTION

Once the dosage form reaches the absorption site, it must disintegrate, deaggregate and release its therapeutic agent. Figure 8.2 indicates solubility problem that may be encountered after the administration of a drug in an oral dosage form. The lack of ability of a drug to go into solution is sometimes a more important limitation to its overall rate of absorption than its ability to permeate the intestinal mucosa. For many drugs that cross the intestinal mucosa easily, the onset of drug levels will be dictated by the time required for the dosage form to release its contents, and for the drug to dissolve.

The dissolution process is primarily dependent on pharmaceutical variables with a possible exception of pH dependency, a patient—related variable. The relative importance of the various processes illustrated in Fig. 8.2 may be explained in terms of Nernst and Brunner equation, whereby dissolution is described by a diffusion-limited model:

$$dC/dt = D\,S\,(C_s\text{-}C_b)/V\,h$$

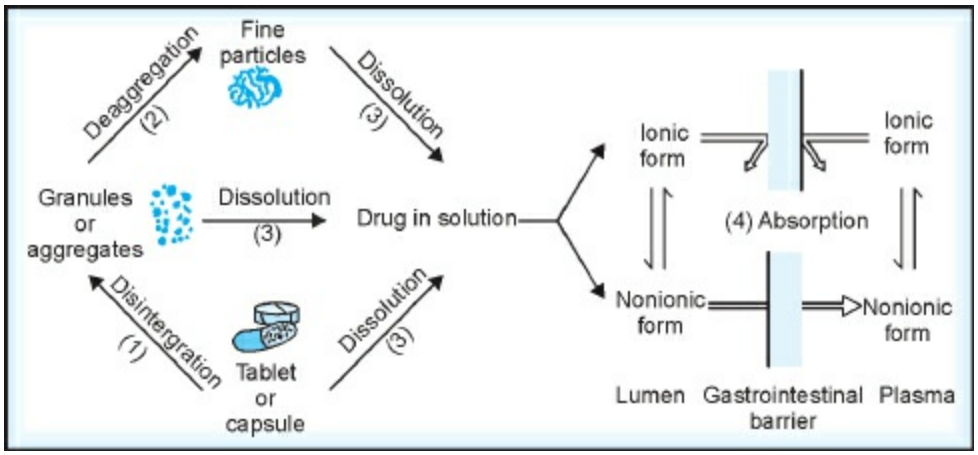



**Fig. 8.2:** Process involved in absorption of drug from dosage form

There are many physicochemical and physiological factors which can have a great influence on the factors in Eq. (3) and therefore, on the dissolution rate.

# DRUG-ASSOCIATED FACTORS

## Factors Influencing the Effective Surface Area of Drug Particles, S

### *Particle Size*

An important factor determining the dissolution rate is the particle size of the drug. The dissolution rate is directly proportional to the effective surface area of the drug, which may be increased by physically reducing the particle size. Faster initial dissolution rates obtained by grinding or milling the drug can often be attributed to both an increase in the surface area as well as changes in surface morphology that lead to a higher surface free energy. However, an increase in edges, corners defects and irregularities on the surfaces of coarse-grade drug particles can also influence the effective hydrodynamic boundary layer, and hence the dissolution rate. Micronization to particle sizes of about 3–5 μm is often a successful strategy for enhancing the dissolution rate of poorly water-soluble drugs, as exemplified for griseofulvin and chloramphenicol. It has been shown that the microcrystalline from of nitrofurantoin (<10 μm) is more rapidly and completely absorbed as compared to macrocrystalline form (74–177 μm). The effective surface area also depends on the ability of the dissolution medium to wet the particle surface. When the dissolution medium has poor wetting properties or when the drug is hydrophobic (aspirin, phenobarbital, and phenacetin), micronization might result in a decreased dissolution rate. It is probable that decreasing the particle size of a hydrophobic drug actually decreases its effective surface area due to agglomeration. In fact, smaller particles of hydrophobic drug have more air adsorbed on their surface resulting in floating of drug on the dissolution medium.

### *Wetting*

A rough estimate of the wettability of a hydrophobic drug by a given medium can be obtained from the contact angle at the liquid/solid interface and the structure of the drug. A compound, not very well wetted by water, possesses a large contact angle. It is worth noting that media such as SGF, SIF and water used during in vitro dissolution testing, do not contain surfactants and therefore fail to address the possibility of wetting and the associated

enhancement in dissolution rate.

The bile salts in the small intestine play a very important role in decreasing the contact angle of poorly-soluble drugs. Addition of bile salts to the medium containing a drug with large contact angle results in a enormous increase in the dissolution rate. However, bile salts have little influence on the dissolution rate of drugs with a small contact angle.

## *Deaggregation*

The rate of disintegration of the dosage form and the size of the resulting aggregates can be the rate-limiting factors step in the dissolution process. As illustrated in Fig. 8.2, after a dosage form disintegrates into larger particles, these large particles must deaggregate to yield fine particles. Hence, deaggregation is often a prerequisite for dissolution. The formulation that deaggregates rapidly results a larger surface area forms has particles of exposed to dissolution medium aggregation decreases the effective surface area, inhibits the exposure of drug to the dissolution medium and thus decreases the dissolution rate.

## *Manufacturing Processes*

Various manufacturing processes can affect dissolution by altering the effective surface area of the drug particles either by addition of hydrophilic diluents or surface-active agents.

## *Granulation*

In terms of achieving higher dissolution rates, wet granulation is considered superior to dry granulation and direct compression procedures. Wet granulation imparts hydrophilic properties to the surface of the granules, and thus improves the dissolution rates of poorly-soluble drugs. Additionally, the use of hydrophilic fillers and diluents such as starch, lactose and microcrystalline cellulose tends to improve the dissolution of the active ingredients.

## *Compression Force*

The high compression force results in an increase in the inter particle bonding, density and hardness, decrease in solvent permeability and inhibition of wettability of the tablet due to the formation of a firmer and

more effective sealing layer by the lubricant under the high pressure and temperature generated during compression. All these effects, due to high compression force, decrease the dissolution rate.

## *Formulation Ingredients*

It has been shown that the dissolution rate of a pure drug can be altered significantly when mixed with various excipients that are added to satisfy certain pharmaceutical functions, such as diluents (fillers), dyes, binders, granulating agents, disintegrants and lubricants. Generally, identical tablet and capsule products manufactured by different pharmaceutical manufacturers, are found to exhibit significant differences in dissolution rates of their active ingredients. Several studies have shown that poor tablet and capsule formulations cause a marked decrease in bioavailability and impairment of the clinical response. Such findings, especially for life-saving drugs like digoxin and tolbutamide tablets, as well as for chloramphenicol and tetracycline were the triggering factors that compelled the drug-regulatory and compendial agencies to institute the dissolution test as a legal requirement for most solid dosage forms.

## *Granulating Agents and Binders*

Tablets granulated with gelatin solution provide a faster dissolution rate in gastric fluid than those prepared using sodium carboxymethylcellulose or polyethylene glycol (PEG) 6,000 as a binder. This observation is attributed to the fact that gelatin imparts hydrophilic characteristics to the hydrophobic drug surface and decreases contact, whereas sodium carboxymethylcellulose is converted to its less soluble acid form at low pH of the gastric fluid, and PEG 6000 forms a poorly-soluble complex. Even gelatin obtained from various processes and from different origins, has been shown to affect the dissolution rate of dosage forms.

## *Disintegrating Agents*

The type and amount of disintegrating agents employed in a formulation significantly control the overall dissolution profile of dosage form. Increasing the starch content from 2% to 5% results in a three-fold increase in the dissolution rate. The effect of starch depends on its swelling characteristics when it comes in contact with water, which results in better and quick disintegration. Another disintegrant, Primojel is found to be ineffective,

particularly if added after granulation.

## *Lubricating Agents*

Lubricating agents are often added so that the powder mass does not stick to the processing machinery. The nature, quality and quantity of lubricants added can affect the dissolution rate. The effect of lubricants on the dissolution rate of salicylic acid tablet has been studied and it has been concluded that magnesium stearate, a hydrophobic lubricant, tends to retard the dissolution rate, whereas sodium lauryl sulfate enhances it, due to the hydrophobic character and surface activity, of the latter which increases wetting and better solvent penetration into the tablets. Effect of lubricants on the dissolution rate of drugs from dosage form would also depend on properties of granules, the lubricant itself and the amount of lubricant used. If granules are fast disintegrating and hydrophilic, a water-soluble surface-active lubricant will have an insignificant effect on dissolution. Most lubricants, such as magnesium stearate, aluminum stearate, stearic acid and talc, are hydrophobic. They decrease the effective drug-solvent interfacial area by changing the surface characteristics of the tablets, which results in reduction of its wettability, prolongation of its disintegration time, and decrease the rate of dissolution.

## Factors Influencing Saturation Solubility of Drug, $C_S$

The saturation solubility is a key factor in the Nernst and Brunner equation, as, together with the concentration of drug already dissolved and the thickness of the hydrodynamic boundary layer, it determines the concentration gradient across the boundary layer, which is the driving force for dissolution. Although the saturation solubility, (*Cs*) influences the apparent dissolution rate constant, it is an intrinsic property of a drug compound and can therefore affect the hydrodynamic boundary layer indirectly. High aqueous solubility, for example, leads to concentration-driven convection at the surface of the drug particles. Various physicochemical and physiological factors influence the saturation solubility of a drug in the gastrointestinal tract. These include its crystalline form, lipophilicity and the ability of the drug to be solubilized by native surfactants and co-ingested foodstuffs, its aqueous solubility and $pK_a$, and the GI pH profile.

### *Polymorphism*

Many drugs are able to crystallize in several forms, each having a different energy, and thereby differing physicochemical properties such as melting point, solubility, heat of fusion, density and refractive index, etc. Depending on the relative stability, one of the polymorphic forms will be more stable than the others. Stable polymorph represents the lowest energy state, highest melting point and least aqueous solubility of the drug. Other forms, termed as metastable forms, represent higher energy states, low melting points and higher aqueous solubilities. Polymorphic form has been shown to influence the solubility, and therefore, the dissolution rate of numerous drugs including chloramphenicol, novobiocin and chlorpropamide. The polymorphic form III of riboflavin is 20 times more soluble than form I. The bioavailability of a drug may also be increased by selection of an appropriate polymorph, but for reasons other than increase in drug solubility. In the case of chloramphenicol palmitate, the rate of enzymatic hydrolysis differs among polymorphs. Since only the free base form of chloramphenicol is efficiently absorbed, the conversion rate of the ester to the free base will be the controlling factor in the overall rate of absorption. Enhanced dissolution rate through appropriate polymorph selection does not, however, always lead to improved

biovailability as can be seen for metastable form of diflunisal. Commercial exploitation of metastable forms with higher solubility is limited by the possibility of interconversion of polymorphs both during manufacture, and subsequently, during storage of the dosage form.

Amorphous forms of drugs, having no internal crystal structure, represent the highest energy state and have solubilites higher than their crystalline counterparts. Among pseudopolymorphs, organic solvates have higher aqueous solubilites than anhydrous form of drug, whereas hydrates have lower aqueous solubilites than anhydrous form.

## *Solubilization*

Solubilization can be defined as 'the preparation of a thermodynamically-stable solution of an insoluble or very slightly-soluble substance by the use of amphiphilic components. A relationship between the extent of solubilization and the lipophilicity of the compound has been observed by Collett and Koo. They found a linear relationship between the π-values of functional groups on substituted benzoic acids and their aqueous/micellar partition coefficients in polysorbate 20.

In the small intestine, drug solubility can be enhanced by amphiphilic bile components such as bile salts, monooleins and lecithin. Increase in solubility of up to 100-fold upon the addition of physiological concentrations of bile salts to aqueous media has been observed. When amphiphilic bile components are present in concentrations higher than their critical micelle concentration (CMC), micellar solubilization of the drug can occur. Solubilization via simple bile salt micelles has been reported for many poorly soluble drugs, including glutethimide, griseofulvin, digoxin and gemfibrozil. Addition of lecithin causes an increase in the molecular weight of bile salt micelles and increases the solubility of drug compounds. The ratio of bile salts to lecithin may also influence the extent of solubilization.

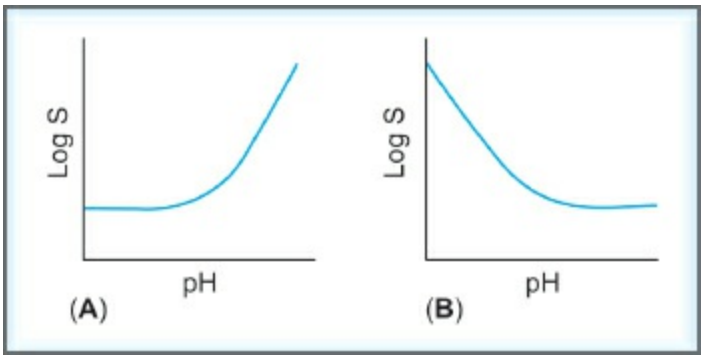
Mithani et al. developed a model to predict solubility enhancement by bile salts based on two physicochemical properties of the drug, namely its log *P* and aqueous solubility. A linear correlation was observed between log *P* and the logarithm of the solubilization ratio *SR*, which is described in Eq. 7 as:

$$[SR] = 0.64 \times \log [P] + 2.09 \quad \ldots (7)$$

In addition to partitioning behavior, other factors such as molecular weight of the drug and specific interactions may be important in determining the extent of solubilization.

## *pKa – pH Profile*

The solubility of weak acids and weak bases is dependent on their ionization constants, $K_a$, and the pH of the dissolution medium. The intrinsic solubility can be defined as the solubility of a compound in its free acid or base form. For weak acids, this is approximated by the solubility at pH values more than one unit below the p*K*a. As the pH value increases, the solubility of the acid increases due to contribution from the ionized form. At pH values exceeding pH = *pKa*+1, a linear relationship between the logarithm of solubility and the pH is observed, until the limiting solubility of the ionized form is reached. An inverse relationship exist for weak bases (Figs 8.3A and B).




**Figs 8.3A and B:** Graphical representation of relation between pH and solubility for (A) Weak acids; (B) Weak bases

a. With increasing pH, the solubility of a weak acids is enhanced, according to the following equation:

$$S = S_0 * \left(1 + \frac{K_a}{[H^+]}\right) \qquad \ldots (8)$$

where, $S$ is the pH-dependent solubility, $S_0$ is the pH-independent solubility and $K_a$ is the acidity constant.

b. With decreasing pH, the solubility of a weak bases is enhanced according to the following equation:

$$S = S_0 * \left(1 + \frac{[H^+]}{K_a}\right) \quad \ldots (9)$$

The pH of the GI fluids, therefore, influences the saturation solubility of ionizable drugs. The pH varies widely with location in the GIT, as shown in Table 8.1.

**Table 8.1:** Average pH values in healthy humans in the fasted and fed states at various sites in the upper GI tract

| Location | Average pH-fasted state | Average pH-fed state |
|---|---|---|
| Stomach | 1.3 | 4.9 |
| Duodenum | 6.5 | 5.4 |
| Jejunum | 6.6 | 5.2–6.0 |
| Ileum | 7.4 | 7.5 |

## *Gastric pH*

There are complex variations in pH of stomach between the fed and fasted states. Upon ingestion of a meal, the gastric pH, at first increases because of buffering effect of food and then after 3–4 h of the meal intake, fasted-state pH has been reestablished due to secretion of gastric acid. Weak bases like dipyridamole (*pKa* = 6.4), ketoconazole (*pKa* = 6.5 and 2.9) and itraconazole (*pKa* = 3.7) are less soluble in the stomach if given immediately after food intake because the gastric fluids are less acidic. This effect will however, be partly offset by the longer gastric emptying time in the fed state, which will afford more time for the drug to go into solution. In contrast, the solubility of a weak base, fluconazole (*pKa* = 1.5), is sufficiently high and its administration to patients with elevated gastric pH does not lead to dissolution rate-limited absorption. Apart from gastric acid secretions, the pH of lumenal fluids is also dependent on other factors like age, pathophysiological conditions such as achlorhydria and AIDS, and concurrent drug therapy such as $H_2$-receptor antagonists and proton pump inhibitors.

## *Small Intestinal pH*

The small intestinal pH, at first, decreases in response to a meal with the arrival of acidic chyme from the stomach but later the fasted-state pH is reestablished as a result of pancreatic bicarbonate output. Poorly-soluble weak acids with *pKa* values less than 6, e.g. indomethacin (*pKa* = 4.5) and furosemide (*pKa* = 3.9), they are relatively insoluble in the gastric juice and dissolution occurs primarily in the upper small intestine. However, in the case of very weak acids like paracetamol (*pKa* = 9.5) or hydrochlorthiazide (*pKa* = 8.8), the variations of GIT pH are irrelevant to the solubility because these compounds are always in the ionized form throughout the physiological pH range.

## *Salt Form*

The dissolution rate of the salt form of a drug is greater than the unionized form of the former due to a higher solubility. However, the solubility of a salt depends on the counterion with a smaller counterion having higher solubility. The solubility of p-amino salicylic acid (ASA) increases in the manner: unionized ASA (1 g in 600 ml) < $K^+$ ASA (1 g in 10 ml) < $Ca^{+2}$ ASA (1 g in 7 ml) < $Na^+$ ASA (1 g in 2 ml). The studies on the beneficial effects of changing unionized drugs into salt form have also been reported for tetracycline and tolbutamide.

## *Buffer Capacity*

For the dissolution of weak acids and bases, the pH of the boundary layer is especially significant. The pH at the surface of the dissolving solid influences the dissolution rate of an ionizable compound in a manner that exactly parallels with its influence on the surface pH. Co-ingestion of liquid and food, with high buffer capacity, may alter the dissolution rate of ionizable drugs in the stomach. The buffer capacity in the small intestine is, however, governed mainly by pancreatic juice. Any increase in the buffer concentration in the small intestine would increase the dissolution rate of weakly-acidic drugs because the pH at the drug surface would be nearer to that of the bulk lumenal pH. It is also worth mentioning here that, since the intestinal juice contains significant levels of bicarbonate buffer, in vitro dissolution tests' results that are obtained in unbuffered or weakly buffered media are unlikely to reflect the in vivo dissolution behaviour of ionizable drugs.

## *Complexation*

## *π-Donor/π-Acceptor Complexes*

Nicotinamide enhances the solubility of poorly-soluble drugs such as progesterone, diazepam and some anti-cancer nucleosides by forming complexes via a π-donor/π-acceptor mechanism. In the case of progesterone, the solubility is enhanced almost 600fold, in comparison to its aqueous solubility, using nicotinamide.

## *Cyclodextrins*

Cyclodextrins are torus-shaped oligosaccharides composed of glucose molecules, which can form inclusion complexes by taking up a guest molecule into the central cavity. A feature of cyclodextrin complexes is that they are stable in aqueous solution. α-, β-and γ-cyclodextrins contain six, seven and eight glucose units, respectively. The α-cyclo-dextrin, with an internal diameter of 6 Å, is too small to complex with most of active pharmaceutical ingredients. The β-cyclo-dextrin possesses low aqueous solubility (1.8% in water at 25°C) because of the formation of stable intramolecular hydrogen bonds, and this limits the degree to which it can be used to improve the solubility of drugs.

It has been shown that inclusion of drugs such as benzodiazepines and digoxin in γ-cyclodextrins, and piroxicam, fenbufen, ibuprofen and cinnarizine in β-cyclodextrins, can result in a very significant increase in their dissolution rates. However, whether the increase in dissolution rate translates to an increase in bioavailability depends on the dissolution of drug is a rate-limiting step.

Recently, hydroxyalkylated derivatives of β-cyclodextrins have been developed. Not only do they have a better aqueous solubility than the parent cyclodextrins, but they also possess a lower toxicity. In the case of carbamazepine, a 2500-fold increase in the aqueous solubility was reached by complexation with 2-hydroxypropyl-β-cyclodextrin (HP-β-CD). The bioavailability of cinnarizine was significantly enhanced, when administered in a capsule, with HP-β-CD, an effect attributed to the more rapid dissolution rate. Further solubility enhancement could be achieved by substituting the β-cyclodextrins with maltosyl- or glucosyl-groups, as demonstrated for lipophilic vitamins and the phytosterol viz. fucosterol.

## *Co-ingested Food*

The solubility of drug in GI tract can also be enhanced when the drug is highly soluble in a co-administered food component. When dicumarol is administered with defatted milk, the resultant plasma levels are five times higher. The main milk protein, casein, is responsible for the significant increase in dicumarol solubility and bioavailability. Griseofulvin absorption is also enhanced by concomitant food intake, especially after ingestion of heavy fat meals. However, in the case of griseofulvin, the prolonged gastric-emptying time and the amplified bile output may also play a role in the enhanced bioavailability.

## *Temperature*

The temperature influences the drug's saturation solubility and also affects the kinematic viscosity as well as diffusion coefficient. Therefore, when performing dissolution experiments, temperature should be monitored carefully or preferably kept constant.

## Factors Influencing the Volume Available for Dissolution, V

One of the ways in which food intake influences the dissolution rate is through the increase in volume of the GI contents. A meal can increase the available gastric volume by as much as 1.5 L. Not only do the ingested food and fluids directly influence the volume in the upper GI tract, they also stimulate secretion of gastric acid, bile and pancreatic juice.

## Factors Influencing the Diffusivity of Drug Molecules, D

The Stokes-Einstein equation describes the relationship between diffusivity and viscosity, showing that diffusivity, *D*, is inversely to proportional to the viscosity, $\eta$

$$D = \frac{k \times T}{6\pi\eta r} \qquad \ldots (10)$$

where, *D* is the diffusivity, *T* is the temperature, $\eta$ is the viscosity of the medium, *r* is the radius of the drug molecule and *k* is the Boltzmann constant.

Dissolution studies, carried out in polysorbate 80 solution of increasing concentration, have shown that at high polysorbate 80 concentrations, the dissolution rate decreases, even though the solubility is enhanced. A dramatic increase in the viscosity or decrease in diffusivity is identified as the reason for this surprising result.

The diffusivity of the dissolving drug can be increased by food intake, however, the extent of the effect will depend on the food components, i.e. the type of meal ingested and the volume of co-administered fluids. Water-soluble fibers such as guar, pectin and some hemicelluloses swell, increase the viscosity of aqueous solutions and decrease the dissolution rate.

## Factors Influencing the Boundary Layer Thickness, h

In the fasted state, the proximal GI tract is quiescent with the cyclic appearance of short bursts of intense propagated motor activity. During the quiescent state (Phase 1), owing to the lack of contractions, the fluid is stagnant and hence the boundary layer is thick. However, the transit time of the drug and therefore, the time available for dissolution is prolonged. The intense motor activity (Phase 3) serves to clear the proximal GIT of the undigested residues and to prevent the overgrowth of bacteria. However, this is an essentially propagative pattern and owing to the short residence time, does not favour dissolution. Phase 2 behavior is intermediate between that of Phase 1 and Phase 3 behaviour.

Co-administered food causes an enhancement in segmental contractions associated with an increase in mixing efficiency. In other words, the boundary layer becomes thin. Therefore, efficiency of dissolution and absorption are probably greater in the fed than in the fasted state. This is illustrated by studies on the rate of glucose uptake from solutions in an intestinal loop model, which showed that absorption is most efficient in the fed state and least efficient during Phase 3 motor activity in the fasted state (Fig. 8.4).

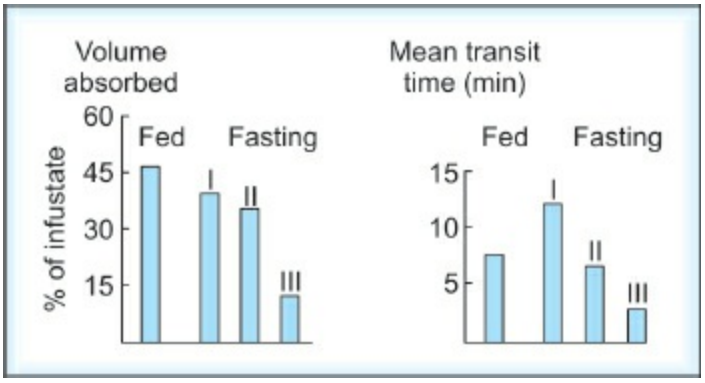



**Fig. 8.4:** Effect of motor pattern in the fasted and fed states on absorption and transit time in the case of a glucose-electrolyte solution introduced into an intestinal loop

Anticholinergics, such as trihexiphenidyl, increase the transit time and therefore, the time available for dissolution, but probably increase the boundary layer thickness, thereby reducing the driving force for dissolution.

Motility-inducing agents, such as cisapride, increase propagative contractions and because of the resultant decreased transit time, do not favour drug dissolution.

## Miscellaneous: Dissolution Test Parameters

### *Eccentricity of the Stirring Device*

USP specifies that the stirring shaft must rotate smoothly without significant wobble. Eccentricity can be measured with a machinist's indicator in terms of total indicator reading (TIR), which determines the sum of the distance on both sides (180°C) of the axis of rotation.

### *Vibration*

Vibration has the effect of changing the flow patterns of the liquid and of introducing unwanted energy to the dynamic system. Both the effects may result in significant changes in the dissolution rates. For most of the drugs, the speeds of the rotational device selected by official compendium are 50 or 100 rpm. Precise speed control is best obtained with a synchronous motor that locks into the line frequency. Such motors are not only more rugged but are far more reliable. Periodic variations in rpm, commonly referred to as torsional vibration might result in possible disturbance in rotational devices.

### *Alignment of the Stirring Element*

USP states that the axis of the stirring element shall not deviate more than 0.2 cm from the axis of the dissolution vessel, which defines centering of the stirring shaft to within ± 2 mm. It also constrains tilt. A series of tests using paddle apparatus suggest that tilt in excess of 1.5° may increase dissolution rates from 2% to 25%, which is a significant variation. The user should be able to adjust his equipment to obtain alignment of the vertical spindles to within 1° perpendicularly with the base of the drive to which the flasks are mounted.

### *Agitation Intensity*

The thickness of the diffusion layer is inversely proportional to the agitation speed, and therefore, agitation conditions can markedly affect diffusion-controlled dissolution. Agitation intensity within and between the various in vitro dissolution testing devices can be varied by the dimensions of the dissolution vessel, volume of the dissolution medium, and the degree of agitation or shaking.

## *Dissolution Media*

Acidic solution tends to disintegrate the tablets slightly faster than neutral solution (water) or basic buffer, and thereby may enhance the dissolution rate by increasing the effective surface area. However, because of corroding action of the acid on dissolution apparatus, currently it is a general practice to use distilled water unless investigative studies show a specific need for the acidic solution to generate meaningful dissolution data. Another approach for avoiding deleterious effects of hydrochloric acid is to replace it with sodium acid phosphate, in order to maintain the required low pH. Addition of surfactants and wetting agents lowers the contact angle and consequently improves penetration by the dissolution medium. Low level of surfactants are recommended to be included in the dissolution medium as this seems to give a better in vitro-in vivo correlation.

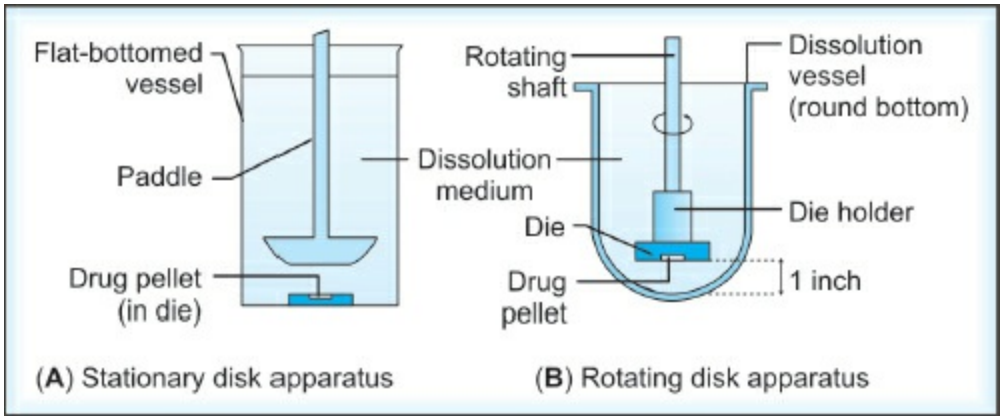
The foregoing discussion shows that the physico-chemical properties of a compound have a strong influence on its dissolution in the gastrointestinal tract, and hence on whether or not dissolution will be the rate-limiting step in its absorption. The poor match between physiological conditions and those used in in vitro dissolution test systems is the primary reason for the inability to predict in vivo dissolution from in vitro data.

## INTRINSIC DISSOLUTION

The intrinsic dissolution (ID) of a pure substance is the rate at which it dissolves from a constant surface area whilst the temperature, pH, agitation and ionic strength of the dissolution medium are constant. Thus, for a drug substance, the ID is independent of formulation factors and measures the inherent solubility of the drug in the dissolution medium. ID determinations can be used to characterize bulk drug substances and excipients, and to test the chemical equivalence of drugs synthesized by different processes. They can also provide an important insight into the dissolution behavior of a drug in physiological conditions.

The ID is a key indicator of the potential bioavailability of a candidate drug. ID e" 1.0 mg/min/cm$^2$ suggests that drug dissolution will not be the rate-limiting step to absorption while an ID d" 0.1mg/min/cm$^2$ suggests that drug dissolution will be the rate limiting step to absorption. An intermediate value suggests that drug dissolution may be the rate-limiting step to absorption.

For determination of ID, two variations of the ID apparatus exist-st*ationary disk apparatus* (Fig. 8.5A) and the *rotating disk apparatus* (Wood Apparatus, Fig. 8.5B). Both types of apparatuses employ a non-disintegrating compacted pellet of pure drug produced by compressing an aliquot of powder in a stainless steel punch and die. This is mounted in a holder so that only one face of the pellet is exposed to dissolution medium.




**Figs 8.5A and B:** Schematic representation of intrinsic dissolution apparatus: (A) Stationary disk apparatus; (B) Rotating disk apparatus

In the stationary disk apparatus, the holder containing pellet is placed face up inside a flat-bottomed dissolution vessel prefilled with the appropriate volume of dissolution medium. The medium is stirred by means of a rotating paddle positioned 6 mm above the pellet surface. In rotating disk apparatus, the die is inverted and screwed onto a shaft, which is then lowered into the dissolution medium until the face pellet is 3.8 cm from the bottom of the vessel. The shaft is then rotated in the same way as USP apparatus 1 and 2.

The ID is calculated by plotting the cumulative amount of substance dissolved per unit area of the exposed pellet surface against time until 10% of the drug pellet has dissolved. However, this compendial calculation may prove difficult to apply to poorly-soluble compounds where 10% dissolution may not be achieved.

## COMPENDIAL METHODS

While selecting apparatus for dissolution testing the following ideal features of dissolution apparatus must be considered:

1. The apparatus must be simple and easy to operate under a variety of conditions.
2. The dimensions and positioning of all the components must be precisely specified.
3. The apparatus must yield repeatable results.
4. The apparatus must be sensitive enough to reveal process variations.
5. The apparatus should permit a uniform and non-turbulent liquid agitation.
6. The apparatus should provide minimum mechanical abrasion to the dosage form.
7. Evaporation of the solvent medium must be eliminated.
8. Medium must be maintained at a fixed temperature within a specified narrow range.
9. Nearly perfect sink conditions should be maintained.
10. Samples should be easily withdrawn without interrupting the flow of the liquid.
11. The apparatus should allow good interlaboratory agreement.
12. The apparatus should be versatile and capable of evaluating disintegrating, nondisintegrating, dense or floating tablets or capsules, and finely powdered drugs.

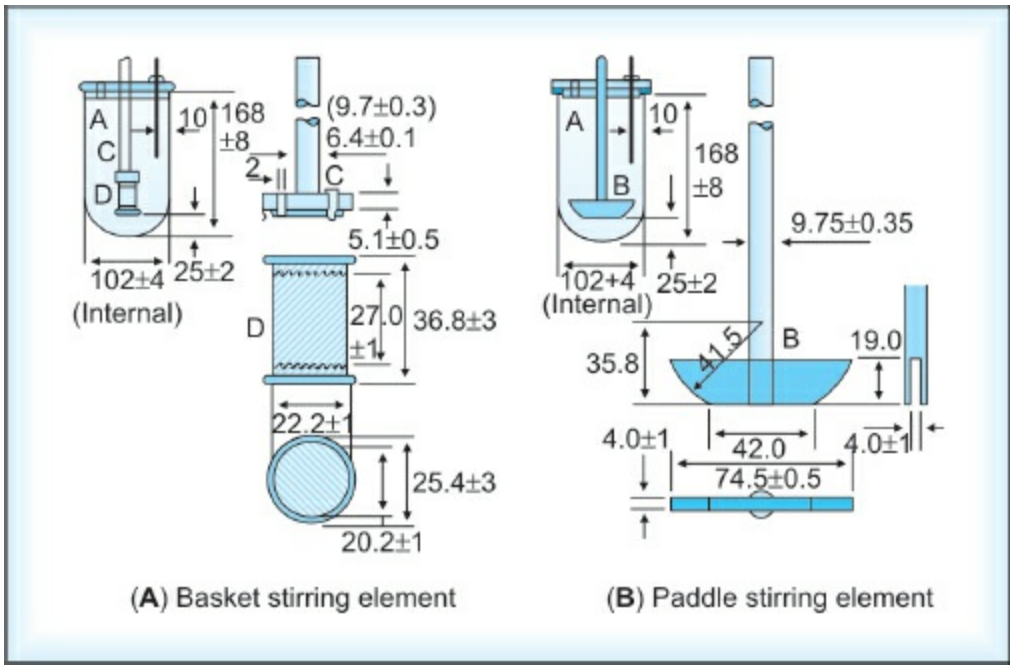
According to USP 30, the official dissolution apparatus are classified into 7 types: USP Apparatus 1-basket, USP Apparatus 2-paddle, USP Apparatus 3-reciprocating cylinder, USP Apparatus 4-flow-through cell, USP Apparatus 5-paddle-over-disk, USP Apparatus 6-rotating cylinder, and USP Apparatus 7-reciprocating holder. In Indian Pharmacopoeia, only two apparatuses are official: IP Apparatus 1-paddle, and IP Apparatus 2-basket. In British and Japanese Pharmacopoeia apparatuses 1, 2 and 4 are official, whereas in European Pharmacopoeia apparatus 1, 2, 3, 4, 5 and 6 are official. Selection of dissolution apparatus and method depends upon the dosage form and

intention of dissolution.

## Basket Apparatus (USP Apparatus 1)

The basket method was first described in 1968 by Pernarowski, Woo, and Searl. The apparatus consists of a motor, a metallic drive shaft, a cylindrical basket and a covered vessel made of glass or other inert transparent material. The contents are held at 37°±0.5°C. The bath liquid is kept in constant and smooth motion during the test and there should be no significant motion, agitation, or vibration caused by anything other than the smoothly rotating stirring element. Ideally, the apparatus should provide clear observation of the stirring element and sample. The vessel is cylindrical with a hemispherical bottom and flanged upper rim. It is 160–175 mm high and has an inside diameter of 98–106 mm, and a nominal capacity of 1000 ml (Fig. 8.6A). A fitted cover may be used to retard evaporation but should provide sufficient openings to allow ready insertion of a shaft and a thermometer, and allow withdrawal of samples.

The shaft is positioned so that its axis is within 2 mm of the axis of the vessels and the lower edge of the blade is 23 to 27 mm from the inside bottom of the vessel, and should rotate smoothly, without significant wobble.




**Figs 8.6A and B:** USP apparatus: (A) Basket stirring element; (B) Paddle stirring element

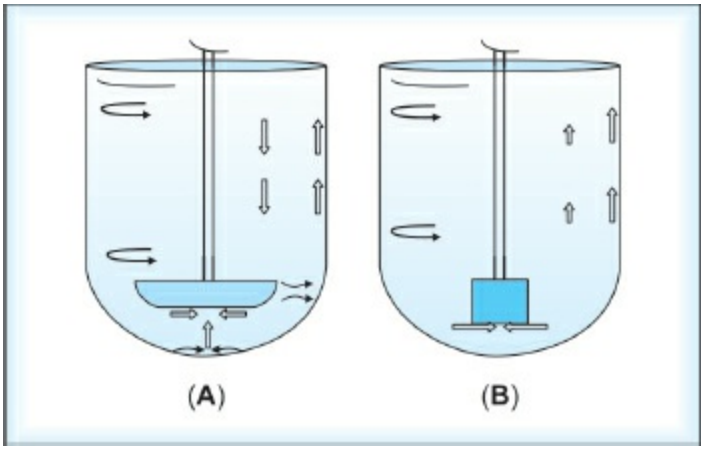
The shaft rotation speed should be maintained within ±4% of the rate specified in the individual monograph using motor with a speed regulator. The shaft has a vent and three spring clips to fit the basket into position. Basket is fabricated of stainless steel, type 316 or equivalent. Welded seam, stainless steel cloth (40 mesh or 425 mm) is used, unless an alternative is specified. For acidic media, 2.5 mm thick gold coating on the basket may be used. For testing, a sample unit is placed in a dry basket at the beginning of each test. Basket method suffers from several drawbacks, few of which have been overcome using cylindrical vessel with a hemispherical bottom having improved hydroynamics (see section: Flow Patterns in Baskets and Paddles). Irrespective of apparatus design, there are still several potential problems. The wire basket corrodes following exposure to acidic media. The basket method gives poor reproducibility due to inhomogeneity of the agitation conditions produced by the rotating basket and clogging of the basket can occur due to adhering substances. Additionally, particles can fall from the rotating basket and sink to the bottom of the flask where they will not be subjected to the same agitation as that inside the basket. Finally, there is possibility of dissolution being accelerated due to abrasion of the surface of the dosage form as it rubs against the basket mesh, the so-called ‘cheese-grater effect.’

## Paddle Apparatus (USP Apparatus 2)

In this apparatus, a paddle replaces the basket as the source of agitation. As with the basket apparatus, the shaft should be positioned no more than 2 mm at any point away from the vertical axis of the vessel and rotate without any significant wobble. The specifications of the shaft are given in Fig. 8.6B. The metallic blade and shaft comprise a single entity that may be coated with a suitable inert coating to prevent corrosion. The dosage form is allowed to sink to the bottom of the flask before rotation of the blade commences.

### *Flow Pattern in Paddle and Basket*

Figures 8.7A and B illustrate the flow patterns for the paddle and the basket apparatuses, respectively. An undertow can be observed visually in the paddle apparatus for stirring rates exceeding 125 rpm. The hydrodynamic region below the paddle, and even more so, below the basket, appears to be somehow 'separated' from the region above the stirring device. Diffusion-driven exchange of the dissolved mass between these two regions is unhampered, but little exchange of particulate material takes place.




**Figs 8.7A and B:** Schematic flow pattern for: (A) the Paddle apparatus; (B) The Basket apparatus

Flow rates given for the basket apparatus, however, are valid for bulk flow only, and do not reflect the influence of hydrodynamics on dissolution inside the basket. Nevertheless, vessel hydrodynamics of regions outside the basket may be relevant for dissolution of solid formulations w.r.t. the fractions of particulate material that have fallen though the basket screen.

Further, hydrodynamics inside the basket may also be influenced by the ‘outside’ bulk hydrodynamics and stirring rate in such a way that, starting with a rotational speed of about 100 rpm or more, contact between the bulk fluid and the formulation inside the basket becomes restricted. At these rates, the basket may be regarded as a ‘closed container’, with limited access to ‘fresh’ dissolution medium and less turbulent flow conditions. For some specific purposes, the basket could even be used to serve as a ‘rotating cylinder,’ with the formulation placed outside the basket at the bottom of the vessel. Such a modified apparatus could be advantageous when mild but reproducible hydrodynamic conditions are desired.

## Reciprocating Cylinder Apparatus (USP Apparatus 3)

Reciprocating Cylinder Apparatus consists of a set of cylindrical, flat-bottomed glass vessels; a set of glass reciprocating cylinders; inert fittings (stainless steel type 316 or other suitable material), and screens that are made up of suitable nonsorbing and non-reactive material and that are designed to fit into the tops and bottoms of the reciprocating cylinders. This has been commercially developed as the Bio-Dis apparatus, which allows tubes containing the sample to be plunged up and down in a small vessel containing the dissolution medium (Fig. 8.8). It has been designed to allow the tubes to be dipped sequentially in up to six different media vessels, using programs that vary the speed and duration of immersion. It allows automated testing for up to six days and the manufacturers advocate its use in the testing of extended-release dosage forms. However, there is some evidence that samples tested using this apparatus tend to yield higher values of the amount released than might be found using alternative procedures. The components conform to the dimensions shown in Fig. 8.8, unless otherwise specified in the individual monograph.

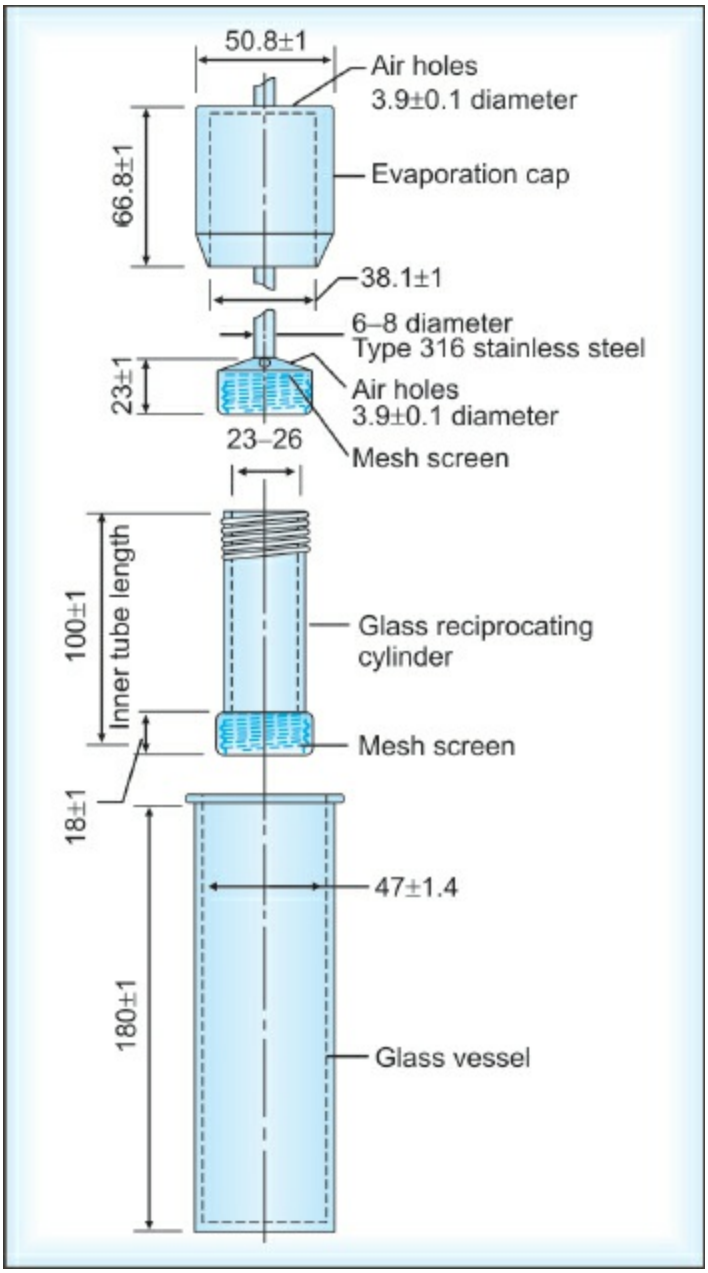



**Fig. 8.8:** USP apparatus 3 (reciprocating cylinder apparatus)

## Flow-through Cell Apparatus (USP Apparatus 4)

Limited-volume apparatuses with a finite volume of dissolution fluid suffer from the problem that they operate under non-sink conditions, which results in limitations when poorly-soluble drugs are being tested. The drawbacks of non flow-through apparatuses include (i) lack of flexibility and homogeneity, (ii) establishment of concentration gradients, (iii) semi-quantitative agitation and (iv) variable shear.

A flow-through system and reservoir may be used to provide sink conditions by continually removing its solution medium and replacing it with fresh medium. Consequently, the flow-through apparatus has been developed, which features a dissolution cell of low volume (often <30 ml) and a reservoir to provide fresh medium (Fig. 8.9). It is prescribed for testing extended-release dosage forms. The basic components are reservoir, pump, heat exchanger, column (cell), tablet support and filter system. The systems enable medium to be taken from a suitable reservoir, and passed straight through the apparatus containing the dosage form, to be either assayed and removed or recirculated.

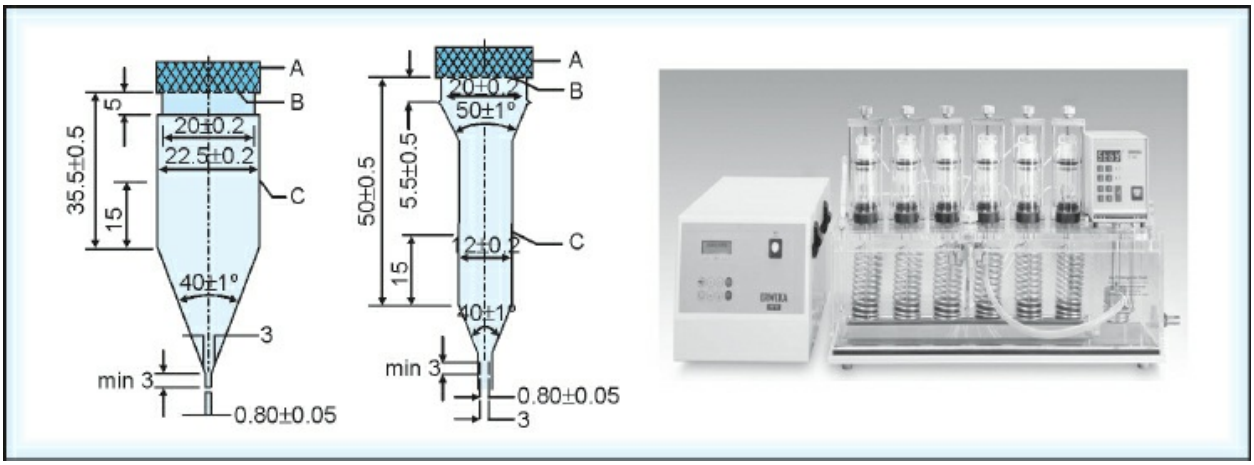



**Fig. 8.9:** USP apparatus 4—flow-through cell

The maintenance of a controlled flow is crucial to column methods and can be influenced by the inlet system. Laminar flow of solvent through the cell is achieved by placing glass beads at the bottom of the cell to facilitate similar disintegration of all surfaces of the sample. It is common to place the tablets on such supports, but attrition (by glass beads) may encourage

breakdown of the dosage form, thereby increasing dissolution rates. Tablet support and consistent positioning in the liquid flow are prerequisites for consistent results. Laminar flow conditions are typically used for tablets, hard gelatin capsules, powders and granules. Suppositories and soft gelatin capsules are placed in the cell without beads for turbulent flow. The columns may be short with tapered inlet and outlet sections but generally are long sections of straightsided tubing to provide hydrodynamic stability. The material under test is placed in the vertically-mounted dissolution cell, which permits solvent to be pumped in from the bottom. The cell type selected is dependent on the dosage form being tested (Fig. 8.10). Standard cells of 12 mm internal diameter used for tablets and capsules show a higher dissolution rate as compared to 26 mm cells due to higher flow velocity. The suppository cell is designed for use with suppositories and soft gelatin capsules. Fats and gels used in these dosage forms are separated in this type of cell so as not to block the filters. For testing powders and granules, a modified cell containing two screen plates is used, where in the sample is placed between the screen plates. The cell for implants has a 1 ml volume of due to the low flow rates (e.g. 5 ml/hr) required for testing this dosage form in a test that may continue over several weeks. The flow rate of the dissolution medium through the cell must be specified for each product. The USP recommends a flow rate between 4 and 16 ml/min with an allowance of ±5%.

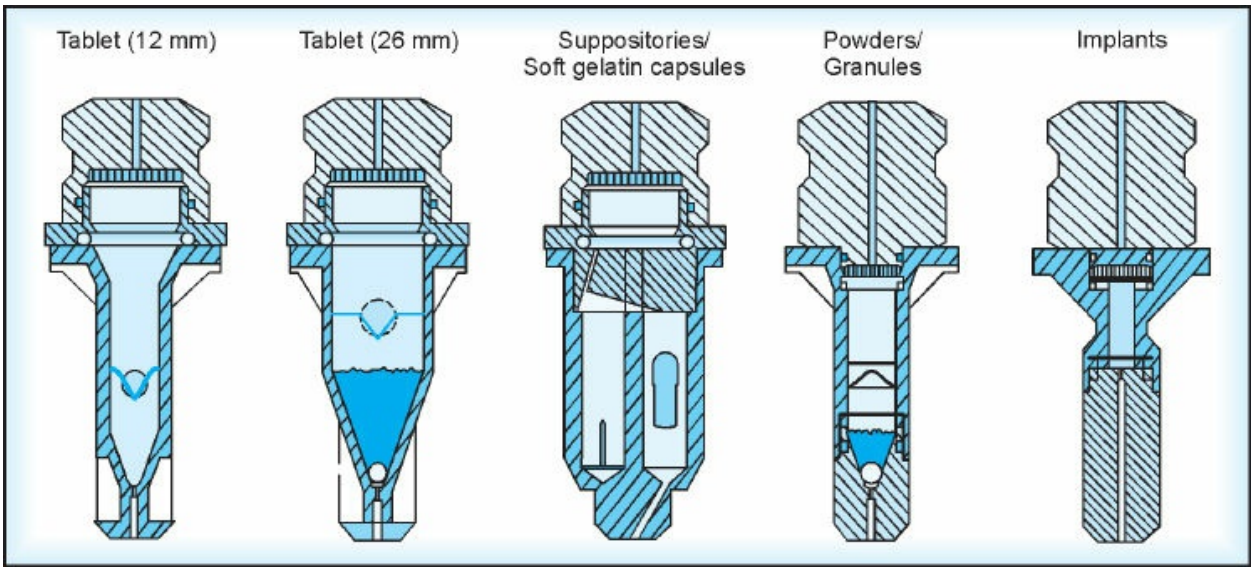



**Fig. 8.10:** Various flow-through cells for dissolution testing of different dosage forms

## Paddle-over-disk Apparatus (USP Apparatus 5)

This uses the paddle apparatus (USP 2) with the sample, usually a transdermal delivery system, being attached to a stainless steel disk, which is then placed at the bottom of the vessel, directly under the paddle.

## Rotating Cylinder Apparatus (USP Apparatus 6)

This is a modification of the basket apparatus with the basket being replaced by a stainless steel cylinder. This apparatus is generally used for transdermal delivery systems by attaching to the outside of the cylinder.

## Reciprocating Holder Apparatus (USP Apparatus 7)

Apparatus consists of a sample holder that oscillates up and down in the medium vessel. The sample holder may take the form of a disk, cylinder, or a spring on the end of a stainless steel or acrylic rod, or it may simply be a rod alone (Fig. 8.11). This apparatus may be used for transdermal products, coated drug delivery systems, or osmotic pump devices. The sample is attached to the outside of the sample holder. It is prescribed for the drug-release testing of psuedoephedrine hydrochloride extended-release tablets USP, where in the tablets are enclosed in a 5 × 5 $cm^2$ of nylon, which is then attached to the rod.

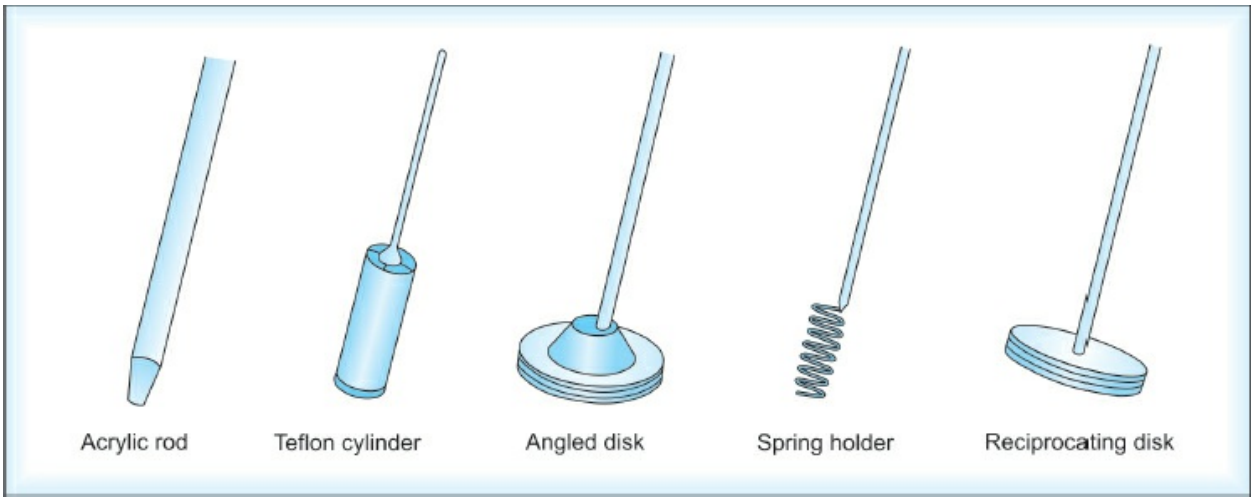



**Fig. 8.11:** Sample holders for dissolution testing of different dosage forms

# DISSOLUTION OF DOSAGE FORMS

With advancement in research in drug delivery, modernization of technology and more emphasis on in vivo predictability of therapeutic effect by means of in vitro tests, dissolution tests have been gaining more and more popularity. The dissolution tests have been successfully implemented on conventional pharmaceutical dosage forms, and generalized monographs described in pharmacopoeias are sufficient to test any such new formulation. However, formal guidelines to evaluate sustained or controlled release products do not exist. The current trend is to evaluate each and every dosage form on individual basis. The regulatory authorities and formulation scientists face an enormous challenge in generalizing dissolution testing because most individual drug candidates for sustained or controlled release dosage forms and their delivery design possess diverse physicochemical and pharmacokinetic properties requiring specific considerations. Difficulties are also face in simulating in vivo conditions in vitro. As a result, modifications in the standard USP test apparatuses as well as the development of novel apparatuses are required in order to mimic in vivo conditions for accurate analysis of these dosage forms.

## Immediate-release Tablets

Immediate-release (IR) dosage forms are intended for rapid delivery of a drug into the blood circulation. Studies of dissolution in immediate release drugs are typically done with USP apparatuses 1–4, those being the rotating basket, paddle, reciprocating cylinder and flow through cell, respectively. In vitro dissolution test for IR solid oral dosage forms are used to assess product quality and performance and batch variation, and to guide development of new formulation. If the tested tablets have low density or they are floating-type, then use of sinkers is advisable.

Sinkers are defined in USP as ‘not more than a few turns of a wire helix’ used for floating dosage forms. Sinkers can act as barriers to dissolution when the wire is wound too tightly around the dosage unit. Sinker recommended in pharmacopoeia is shown in Fig. 8.12.

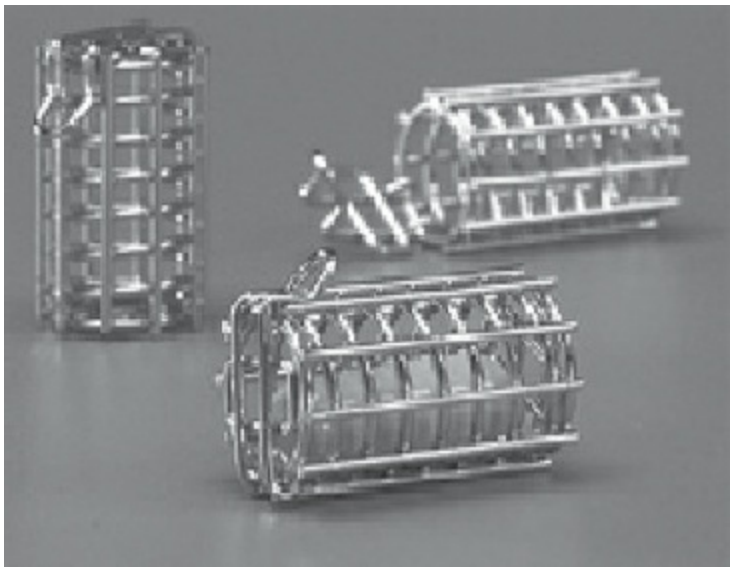


**Fig. 8.12:** Sinker recommended in pharmacopoeia

The requirements for IR dosage forms are met if the quantities of active ingredients dissolved from the dosage units tested conform to Table 8.2. The testing is continued until the results conform at either stage 1 or stage 2.

**Table 8.2:** Acceptance table for immediate release dosage forms

| Stage | Number tested | Acceptance criteria |
|---|---|---|

| | | |
|---|---|---|
| $S_1$ | 6 | Each unit is not less than D* + 5% |
| $S_2$ | 6 | Average of 12 units ($S_1$ + $S_2$) is equal to or greater than D, and no unit is less than D-15%. |
| $S_3$ | 12 | Average of 24 units ($S_1$+ $S_2$ + $S_3$)is equal to or greater than D, not more than 2 units are less than D-15% and no unit is less than D-25% |

* D is the amount of dissolved active ingredient specified in the individual monograph, expressed as percentage of the stated amount.

## Extended-release Tablets

An ideal dissolution apparatus for an extended-release (ER) product should be able to tackle at least some of the challenges that the formulation scientists face in simulating in vivo conditions. The apparatus would be capable of simulating (i) the entire pH range of the GI tract, (ii) food-induced physiological changes that occur in the GI tract and (iii) the motility pattern and other mechanical forces encountered by the dosage form in the GI tract. To overcome some of these difficulties, several modifications have been suggested in the literature (Table 8.3), but none of these modified methods is comprehensive enough to be adopted for testing of various types of ER products.

**Table 8.3:** Modifications proposed to standard dissolution methods for testing of extended release dosage form

| Factor studied | Modification | Drug/dosage form used |
|---|---|---|
| GI motility | Paddle-bead method – polystyrene beads inserted into dissolution medium | Phenyl-propanolamine HCl |
| pH of GIT | Dialysis cell containing dosage form in small volume of fluid, immersed in dissolution medium | Theophylline beads, Theophylline matrix tablets |
| High fat food | Pretreatment of dosage form in peanut oil | Theophylline beads, Theophylline matrix tablets, Propanolol HCl capsules |
| Poor solubility of drug | Addition of solubilizer | Felodipine matrix tablet |

Apparatuses 1, 2 and 7 are mentioned in the USP for the dissolution testing of ER tablets. Modified dissolution apparatus 2 has been stated in

USP for nifedipine, felodipine and metformin hydrochloride ER tablets. The new ‘stationary tablet basket’ apparatus contains a stationary stainless steel tablet basket (in which the tablet is placed) located 1 cm above the paddle (Fig. 8.13). Flow-through cell or reciprocating cylinder is also utilized for the dissolution testing of the ER tablets.

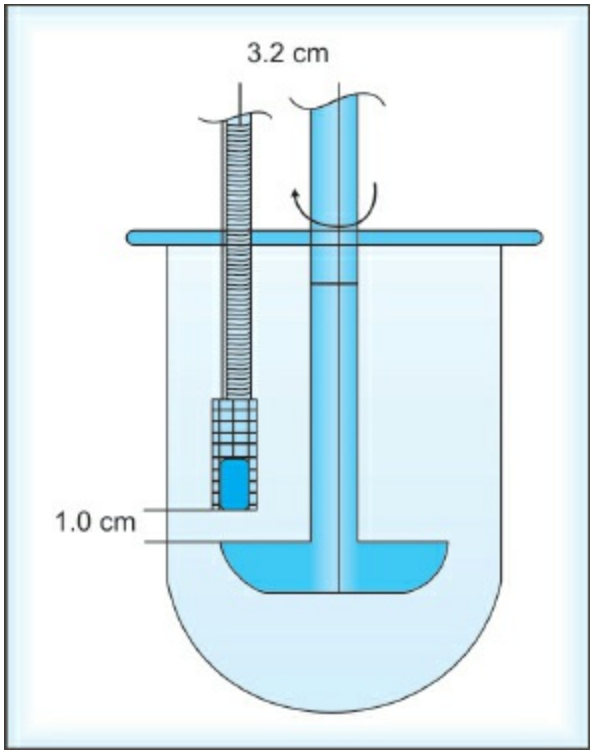



**Fig. 8.13:** Configuration of stationary tablet basket

The requirements of the extended release dosage forms are met if the quantities of active ingredient dissolved from the dosage units tested conform to Table 8.4. The testing is continued unless the results conform at either level 1 or 2.

**Table 8.4:** Acceptance Table for extended release dosage forms

| Level | Number tested | Criteria |
|---|---|---|
| $L_1$ | 6 | No individual value lies outside each of the stated ranges and no individual value is less than the stated amount at final test time |
| $L_2$ | 6 | The average value of the 12 units ($L_1$ + $L_2$) |

| | | |
|---|---|---|
| | | lies within each of the stated ranges and is not less than the stated amount at final test time; none is more than or less than 10% of labeled content at final test time |
| $L_3$ | 12 | The average value of the 24 units ($L_1$+$L_2$+$L_3$) lies within each of the stated ranges and is not less than the stated amount at final test time; not more than 2 of the 24 units are more than or less than 10% of labeled content at final test time and none of the unit is more than or less than 20% of labeled content at final test time |

## Chewable Tablets

Chewable tablets are rapidly disintegrating, used to achieve a fast onset of action and are designed to increase compliance among individuals who are unable to swallow traditional tablets. The most acceptable apparatus for testing of chewable tablet is USP apparatus 2, the same as for traditional tablets. However, ampicillin chewable tablets require the use of apparatus 1, and carbamazepine chewable tablets use apparatuses 2 and 3 as two different tests. Furthermore, the use of a reciprocating cylinder, along with glass beads has been recommended in order to create a large amount of agitation within the dissolution medium. Mechanical breakage of the tablet prior to the dissolution test is also recommended.

## Sublingual/Buccal Tablets

The sublingual or buccal route is suitable for medications that cannot be taken by the oral route due to GIT degradation, or their susceptibility to the hepatic first pass effect. These tablets are also advantageous for patients who are unable to swallow whole tablets. Since these medications are designed to dissolve the drug in a short time period, it is obvious that disintegration and not necessarily dissolution is the true rate-limiting step for drug release of these dosage forms. In lieu of this assumption, USP 29 states the use of disintegration test for ergotamine tartrate sublingual tablets few of the sublingual tablets, such as isosorbide dinitrate sublingual tablet, are tested officially using apparatus 2 with water as the dissolution medium. However, in vivo dissolution is limited for these tablets by the amount of saliva present in the mouth. As a result, dissolution tests using standard USP apparatuses and large volumes of liquids might not produce results that reflect in vivo dissolution. Drug dissolution in smaller volumes or using different apparatuses might mimic the physiological condition of oral cavity. Comparatively better in vitro results are obtained when hydrocortisone hemi-succinate muco-adhesive tablets were tested using USP Apparatus 3 at a rate of 20 strokes/min.

Another modification, comprising a single stirred, continuous flow-though filtration cell with a dip tube to remove finely divided solid particles, has been reported. The volume of liquid in the cell is 10 ml and the fluid is pumped through to give a short residence time with almost complete removal in about 8 min. Figure 8.14 shows a schematic drawing of this apparatus.

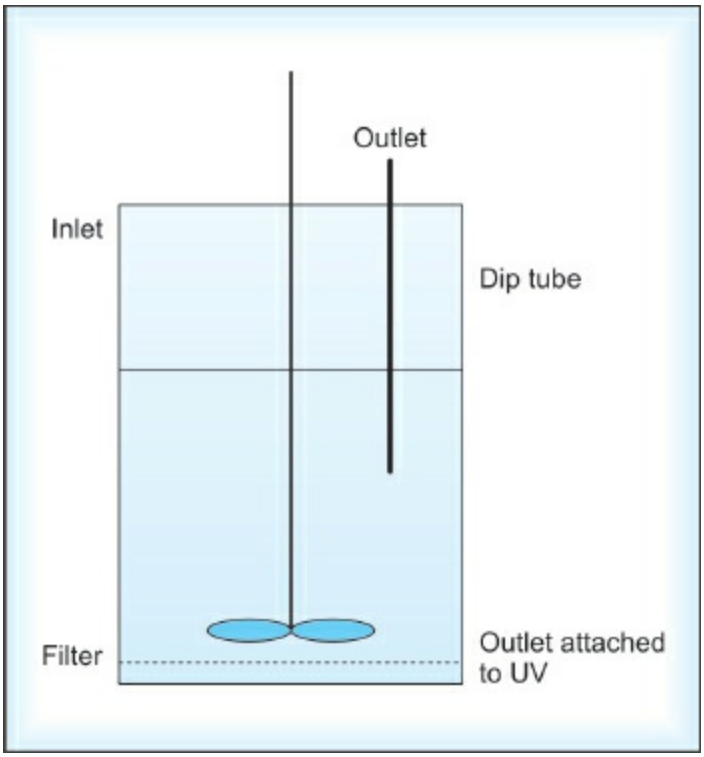



**Fig. 8.14:** Schematic diagram of dissolution apparatus for sublingual/buccal tablets

A special disintegration test using a Texture Analyzer Instrument to accurately determine the rate of drug release from sublingual medications has been developed. In this method, the tablet is attached to a cylindrical probe and placed under a constant force to promote disintegration. The tablet is then submerged in a defined volume of medium, and the time for complete tablet disintegration versus distance travelled is determined. Drug release studies for buccal tablets are normally performed using USP Apparatus 2. However, to mimic the intended drug release in one direction only (buccal mucosa), modified Franz diffusion cells and intrinsic dissolution apparatus are utilized. The use of modified Franz diffusion cell is demonstrated in the testing of nicotine buccal tablets.

## Chewing Gums

For chewing gums, no apparatus till date has been official. A 3-piston apparatus, which in essence “chews” the gum at a rate of 60 cycles/min in a test medium with pH of 6.0 at 37°C, has been developed and its use has been recommended by European Pharmacopoeia. There are several obvious disadvantages using this method, for instance, the chewing gum may adhere to the equipment, thus affecting its ability to imitate in vivo condition. One of the unique methods to evaluate chewing gums is to allow human volunteers to chew the medicated gum for a specific period of time (i.e. 10, 20, 30, or 40 min), followed by analyzing the residual quantity of amount of active ingredient remaining in the gum. This method definitely warrants some scrutiny in methodology but is a prime example, demonstrating the need of developing an appropriate in vitro test apparatus to analyze the release of medication from chewing gums.

## Powders

The USP does not state any official method for dissolution testing of powders. The only application of powder dissolution in the USP is the evaluation of intrinsic dissolution of powders. Dissolution testing of finely-divided particles can be performed using apparatus 2 or the flow-through cell apparatus. However, it has to be noted that in the standard USP apparatuses, the dispersal of powders may have an impact on the dissolution behavior. In an attempt to keep both drug and excipients together, a modified-basket method was developed to better simulate the in vivo environment. The basket used in this setup was dipped into molten wax in order to seal the bottom. In this modified apparatus, excipients were able to interact with the drug for a longer period of time. A modified vessel, 'Peak Vessel' is designed to eliminate "mounding or coning" by having a cone molded into the bottom of the glass vessel (Fig. 8.15). The peak vessel is non-compendial, but may have utility in products that contain dense excipients that can have a tendency to cone rather than disperse freely inside the vessel.

**Fig. 8.15:** Peak vessel

## Capsules

Soft gelatin capsules may contain either hydrophilic or hydrophobic drug components. In the case of hydrophilic capsules, dissolution tests can be performed quite easily using USP Apparatus 2, but this becomes more difficult for hydrophobic medications. For soft gelatin capsules, which are dietary supplements, the USP has added a rupture test based on the time needed for capsule shell to rupture in 500 ml water. The capsule shell must rupture within 15 min but no drug release is measured. In vitro dissolution tests of lipophilic drugs from oil containing soft gelatin capsules have up to be performed in the USP paddle or basket apparatus. Because of the unfavorable oil-water partition coefficient of hydrophobic drugs, surface-active compounds have been added to the aqueous dissolution media in order to avoid long dissolution times. However, it is speculated that exposure of the gelatin shell to such media may induce physical and/or chemical changes in the drug, arising either through complex formation or cross-linking reactions. Cross-linked capsules can be tested using USP *two-tier dissolution test*, which involves the use of enzymes, pepsin and pancreatin in the dissolution medium.

The official methods have the serious disadvantage that the dissolution conditions for hydrophobic floating materials are poorly defined and sample taking might be difficult. One way to solve such problems is to use a flow-through cell in which the site of dissolution is smaller flow conditions are better defined, sample taking is simple because the drug is removed from the excipient by continuous extraction with an aqueous perfusion medium and automatically filtered.

The standard flow-through cell is not suitable for lipid-filled soft gelatin capsules, because after capsule rupture, the oil phase is quickly drawn into the filter on the top of the cell, which can get clogged. A new flowthrough cell for lipid-filled soft gelatin capsules has been designed and a schematic view of this device is shown in Fig. 8.16. In this flow-through cell, the dissolution medium enters through the medium inlet, on the right-hand side of the cell, moves to the left side of the cell, pushes the air out through a capillary, and then the flows through the center channel to the filter. After the capsule ruptures in the right-side of the cell, the lipid content rises up, due to its lower density. When the lipid phase reaches the triangular area top of the

left-side cell, it stays there. Thus, the dissolution medium continuously extracts the drug from the lipid layer as it flows through the cell. A rotating dialysis cell method has also been investigated for dissolution of tocopherol nicotinate from soft gelatin capsules. The use of two-phase dissolution medium (organic and aqueous) has also been utilized for conducting dissolution in lipid-filled soft gelatin capsules.

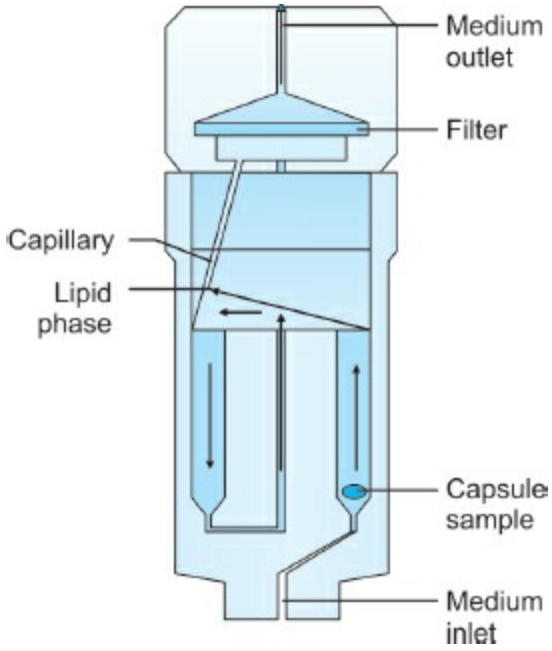



**Fig. 8.16:** Schematic view of flow-through cell for lipid filled soft gelatin capsule

## Suppositories

To find a standard method to test drug release from lipophilic suppositories is challenging as in the case of lipid-filled soft gelatin capsules. This is due to the melting and deformation of the suppository in the dissolution medium. For in vitro release study, one requires knowledge of the melting point range of the suppository base. The testing temperature should be similar to physiological conditions. However, some studies allow higher temperatures, particularly for acetaminophen suppositories used to treat fever. Most common apparatus for dissolution test of suppositories is apparatus 2. The same has been reported in the USP for testing of indomethacin suppository. A modified paddle or basket method with a wired screen, and a sinker or a modified flow-through cell with a specific dual-chamber suppository cell have been recommended for lipophilic suppositories. Hydrophilic suppositories release the drug by dissolving, as opposed to melting, in rectal fluids. Conventional basket, paddle, or flow-through cell seems to be suitable for dissolution testing of hydrophilic suppositories. However, no simulated rectal fluid exists at the moment to simulate the in vivo dissolution of suppositories.

## Transdermal Patches

The USP has published three different in vitro drug release tests for dissolution testing of patches. These include paddle over disk, cylinder method, and reciprocating disk method, i.e. apparatuses 5, 6, 7, respectively. The paddle-over-disk method is the most widely used because it is simple and easy to reproduce. The testing conditions should be ideally adjusted to pH 5–6, reflecting physiological skin conditions. The testing should be done at 32°C at 100 rpm agitation speed. Nicotine transdermal patch is official in USP. The above mentioned three different apparatuses are recommended for drug release testing of this patch. However, there are numerous examples of using Franz diffusion cell for release studies of transdermal systems in the literature.

## Semisolid Dosage Forms

Semisolid dosage forms include creams, ointments and gels. Currently, no monograph exists in the USP which uses dissolution testing of semisolid bases. In research, the drug release test is normally performed using the Franz diffusion cell system. A schematic picture of Franz diffusion cell is shown in Fig. 8.17. Franz diffusion cell system consists of a receiver chamber and a donor chamber. The dissolution medium is contained in receiver chamber where as drug product is introduced into donor chamber. The skin or membrane is placed between the two chambers. The sampling is done from receiving chamber through a sampling port. Common membranes are Tuffryn®, Supor®, Cellulosic, Acetate Plus®, Nylon, Teflon, and polycarbonate. The receiving medium must be similar to physiological conditions of the skin. Critical components of the in vitro release test for semisolid products include selection of an assay method, diffusion cell volume, selection of an appropriate membrane, nature of the receiving medium, equipment related parameters, viz. stirring speed and temperature, and validation of the method.

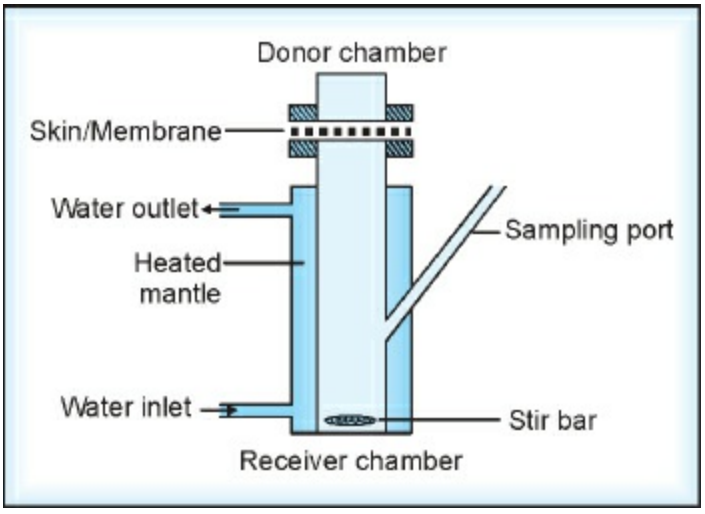



**Fig. 8.17:** Franz diffusion cell system

Enhancer cell, designed by Vankel Technology Group, is another device which is used for dissolution testing of semisolid products. This device is a Teflon cell with adjustable volume and a screw cap to retain the skin or artificial membrane. The semisolid product is put into the cell and a membrane is used to provide a defined surface to determine drug release. The

assembly can be used with any dissolution tester and is available with 4.0, 2.0, or 0.5 $cm^2$ surface area. The Paddle-over-enhancer-cell method provides release rates comparable to Franz cell technology.

## Aerosols

To date no single in vitro test system has yet emerged as the ideal choice for performing dissolution measurements to estimate in vivo solubility in the lung fluids. The custom made flow-through dissolution apparatus is the only method, which has been used to study the dissolution of inhaled glucocorticoid particles. In this method, the aerosol particles, obtained using impaction, are collected onto a glass prefilter for dissolution studies. The dissolution medium is pumped to flow through the aerosol particles, previously collected and immobilized on the glass fiber filter placed between 0.45pm membrane filters. The dissolved fraction of the dose, which passess the upper filter, is collected separately for individual analysis at pre-determined intervals.

## Suspensions and Granules

Release of drug from particles in granules and suspensions is usually intended to occur rapidly in vivo, and testing of such dosage forms is not common. However, some IVIVC studies have confirmed the importance of dissolution rate determinations of suspensions as a discriminative test for rapid screening of new formulations. For testing of suspensions like cefuroxime axetil oral suspension and indomethacin oral suspension, and granules like Flunixin Meglumine Granules, USP Apparatus 2 at a rotation speed between 25 to 50 rpm has been used frequently. However, the rotating filter apparatus has gained wide acceptance for suspensions because it provides mild laminar liquid agitation, and also functions as an in situ non-clogging filter.

## Nutritional/Dietary Supplements

USP 26/NF 21 contains a section on Nutritional Supplements and USP 30/NF 25 contains a section on Dietary Supplements. The dissolution procedures for the nutritional supplements use Apparatus 1 and 2, and require the measurement of one vitamin and folic acid (if applicable), and one mineral (if applicable). Oil-soluble vitamins are exempted from the dissolution requirement. Both pharmacopoeial and non-pharmacopoeial standard methods have been discussed with some non-pharmacopoeial dissolution apparatuses, and method are intended to be used in research and development only and might not be suitable for routine quality control. However, despite the fact that they are not pharmacopoeial standard methods, they have the potential to provide valuable information on the expected in vivo drug release. Therefore, it is necessary to further develop in vitro assays for novel dosage forms and establish standard protocols for their drug release tests.

## DISSOLUTION PROFILE COMPARISON

In recent years, more emphasis has been placed on dissolution testing within the pharmaceutical industry, and correspondingly, by regulatory authorities. Indeed, the comparison of dissolution profiles has extensive application throughout the product development process. Under appropriate test conditions, a dissolution profile can characterize the product more precisely than a single-point dissolution test. A dissolution profile comparison can be used to:

- Develop in vitro-in vivo correlations, which can help to reduce costs, speed-up product development and reduce the need to perform costly bioavailability (BA)/bioequivalence (BE) human volunteer studies;
- Establish final dissolution specifications for the pharmaceutical dosage form;
- Establish the similarity of pharmaceutical dosage forms, for which composition, manufacturing site, scale of manufacture, manufacturing process and/or equipment may have changed within defined limits.

Methods used to compare dissolution profile data can be categorized under the following general headings:

i. Exploratory data analysis methods-graphical and numerical summaries of the data.

ii. Mathematical methods—methods that typically use a single number to describe the difference between dissolution profiles, e.g. $f_1$ and $f_2$ equations, and Rescigno's indices.

iii. Statistical and modeling methods-ANOVA, modeling-based methods, mixed effects model, and Chow and Ki's method.

Among several methods investigated for dissolution profile comparison, the method described by *Moore and Flanner* is the simplest and widely applicable. Moore and Flanner describe a model-independent mathematical approach to compare the dissolution profile using two factors-a '*difference factor*' (*Eq. (11)* $f_1$) and a '*similarity factor*' (*Eq. (12*), $f_2$). Both equations are endorsed by the FDA as acceptable methods for dissolution profile comparison, but the $f_2$ equation is preferred. The equations are:

$$f_1 = \left\{ \frac{\sum_{t=1}^{n} |R_t - T_t|}{\sum_{t=1}^{n} R_t} \right\} \times 100\% \quad \text{... (11)}$$

$$f_2 = 50 \log_{10} \left\{ \left[ 1 + \frac{1}{n} \sum_{t=1}^{n} W_t (R_t - T_t)^2 \right]^{-0.5} \times 100 \right\} \quad \text{...(12)}$$

where, $n$ is the number of dissolution time points, $R_t$ and $T_t$ are the reference and test dissolution values, respectively at time t, and $W_t$ is an optional weighting factor.

The $f_1$ equation is the sum of the absolute values of the vertical distances between the test and reference mean values, i.e. $(R_t - T_t)$ at each dissolution time point, expressed as a percentage of the sum of the mean fractions released from the reference formulation at each time point. The $f_1$ equation is zero when the mean profiles are identical and increases proportionally as the difference between the mean profiles increases.

The $f_2$ equation is a logarithmic transformation of the average of the squared vertical distances between the test and reference mean dissolution values at each dissolution time point, multiplied by an appropriate weighting, i.e. $W_t(R_t - T_t)^2$.

The transformation is such that the $f_2$ equation takes values less than 100. The value of $f_2$ is 100 when the test and reference mean profiles are identical. Values of f between zero and 15, and of $f_2$ between 50 and 100 ensure sameness or equivalence of the two dissolution profiles. This range of $f_2$ values equates to an average difference between mean dissolution profiles of up to 10%. In addition to these criteria, the FDA scale-up and post-approval changes-modified release guidance states that the average difference at any dissolution time point between the test and reference mean profiles should not exceed 15%. The main advantages of the $f_1$ and $f2$ equations are that they are easy to compute and they provide a single number to describe the comparison of dissolution profile data.

For modified-release (MR) dosage forms including extended- or delayed-release dosage forms, dissolution profile data would be generated at a number of time points (typically five), until either 80% of the drug is releases or the

dissolution profile reached an asymptote. In contrast to MR dosage forms, dissolution profile data for immediate-release (IR) may be generated at fewer time points. The typical USP requirements for IR dosage forms are that 75% of the active ingredient from the dosage unit should be dissolved in water or acid at 37°C in 45 min in the USP 1 or USP 2 Apparatus, which is operated at the appropriate speed (typically 100 rpm for USP 1 or 50 rpm for USP 2). Further requirements for dissolution profile comparison are:

- At least 12 units should be used for each profile determination. Mean dissolution values can be used to estimate the similarity factor, $f_2$. To use mean data, the percentage coefficient of variation at the earlier point should not be more than 20% and at other time points should not be more than 10%.
- The dissolution measurements of the two products (test and reference, pre- and postchange, two strengths) should be made under the same test conditions. The dissolution time points for both the profiles should be the same, e.g. for IR products-15, 30, 45, and 60 min, and for extended-release products-1, 2, 3, 5, and 8 h.
- Because $f_2$ values are sensitive to the number of dissolution time points, only one measurement should be considered after 85% dissolution of the product.
- For drug products dissolving 85% or greater in 15 min or less, a profile comparison is not necessary.

## IN VITRO-IN VIVO CORRELATION (IVIVC)

IVIVC is a general term that refers to a relationship between a biological property (in vivo input rate) produced by a dosage form and a physicochemical characteristic of the same dosage form (in vitro dissolution). As is well known, in vitro dissolution testing is a powerful and useful method for determining product quality and sometimes for evaluating the clinical performance of dosage forms. The utility of in vitro dissolution as a surrogate for in vivo bioavailability is very attractive and has been demonstrated for several products. In order to successfully develop an IVIVC, in vitro dissolution has to be the rate-limiting step in the sequence of steps leading to absorption of the drug into the systemic circulation. Furthermore, to utilize dissolution test as a surrogate for bioequivalence (where a relatively simple in vitro test is used in place of human testing), the IVIVC must be predictive of in vivo performance of the product. Four categories of IVIVCs have been described in the FDA guidance.

**Level A:** A level A correlation represents a point-to-point relationship between the in vitro dissolution and the in vivo input rate (e.g. the in vivo dissolution of the drug from the dosage form). Generally these correlations are linear, however, non-linear correlations are also acceptable. Linear relation between absorption and dissolution demonstrates superimposable data with a slope of one, an intercept of zero, and a coefficient of determination of one.

In the case of a successful level A correlation, an in vitro dissolution curve can serve as a surrogate for in vivo performance. Therefore, a change in the manufacturing site, method of manufacture, raw material supplies, minor formulation modification, and even product strength using the same formulation can be justified without the need for additional human studies. A level A correlation is considered most informative and very useful from a regulatory viewpoint.

**Level B:** A level B correlation uses the principles of statistical moment analysis. The mean in vitro dissolution time is compared either with to the mean residence time or with the mean in vivo dissolution time. Although this type of correlation uses all of the in vitro and in vivo data, it is not considered a point-to-point correlation. Further, since it does not uniquely reflect the actual in vivo plasma level curve, this is not very useful from the regulatory

perspective.

**Level C:** A level C correlation establishes a single point relationship between one dissolution parameter (e.g. $t_{50\%}$ or percent dissolved in 4 h) and one pharmacokinetic parameter (e.g. AUC or $C_{max}$). A level C correlation does not reflect the complete shape of the plasma concentration-time curve, and therefore is not the most useful correlation from a regulatory perspective. However, this type of correlation can be useful in early formulation development. Some of the parameters used in single-point correlations are presented in Table 8.5.

**Table 8.5:** Parameters used for correlating in vitro dissolution with in vivo data

| In vitro parameters | In vivo parameters |
|---|---|
| Time for a specific amount dissolved | Area under the concentration vs time curve |
| Amount dissolved at specific a time point | Fraction absorbed, absorption rate constant |
| Mean dissolution time | Mean residence time |
| Parameter estimated after modeling dissolution process | Concentration absorbed at time *t* |

**Multiple level C:** A multiple level C correlation relates one or several pharmacokinetic parameters of interest to the amount of drug dissolved at several time points of the dissolution profile. Multiple level C correlation can be as useful as level A IVIVC from a regulatory perspective. However, if one can develop a multiple level C correlation, it is likely that a level A correlation can be developed as well.

If the drug is highly permeable and in vitro dissolution is the rate-limiting step, it is very highly likely that a successful IVIVC can be developed.

## General Principles of IVIVC

i. IVIVC should be developed using two or more formulations with different release rates:
   - Only one release rate is sufficient if dissolution is condition-independent

ii. In vitro dissolution profiles should be generated using an appropriate dissolution methodology:
   - The dissolution method used should be the same for all the formulations

iii. A bioavailability study should be conducted to determine the in vivo plasma concentration time profiles for each of the formulations:
   - Preferably, this should be a crossover study in adequate number of subjects.

iv. In vivo absorption profile is plotted against the in vitro dissolution profile to obtain a correlation (Fig. 8.18).

A predictive IVIVC can empower in vitro dissolution as a surrogate for in vivo bioavaila-bility/bioequivalence. This can be used to:

- Grant biowaivers
- Set meaningful dissolution specifications that take into account the clinical consequences.

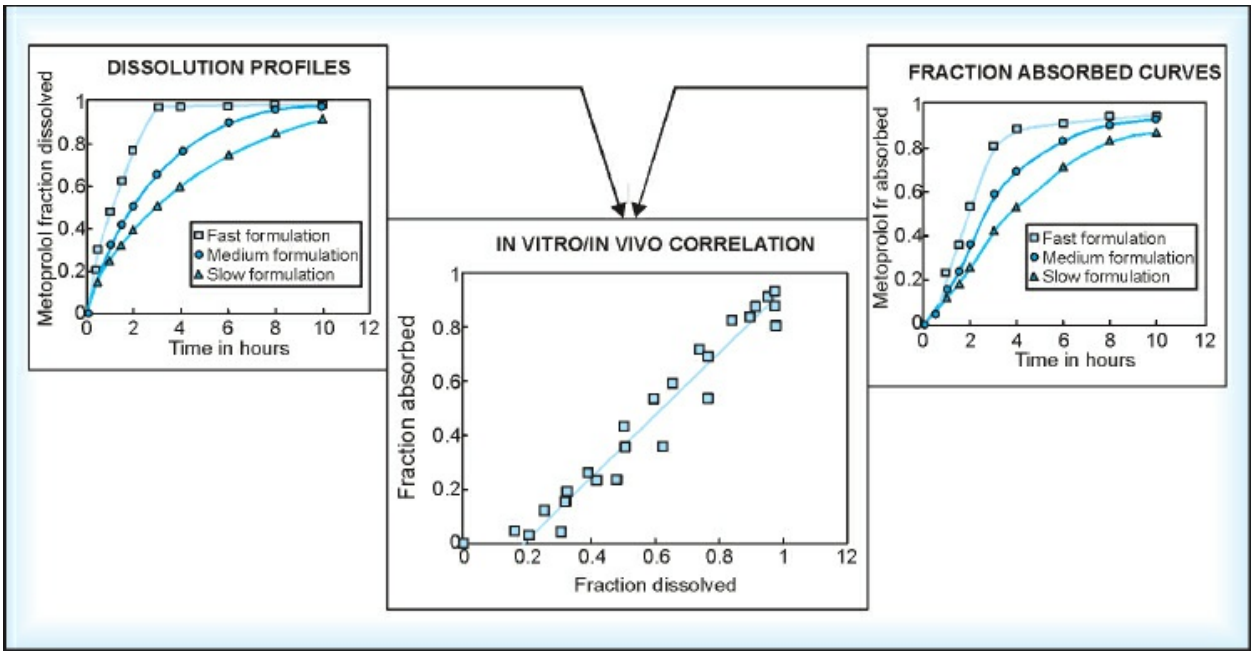



**Fig. 8.18:** Development of in vitro-in vivo correlation

## DISSOLUTION BASED BIOWAIVERS

The introduction of the Biopharmaceutical Classification System (BCS) precipitated a tremendous surge of interest in dissolution and dissolution testing methodologies. BCS is devised to classify drugs based on their aqueous solubility and intestinal permeability. The BCS characteristics (solubility and permeability), together with the dissolution of the drug from the dosage form, take into account the major factors that govern the rate and extent of drug absorption from dosage forms. It classifies the drug substance (and therefore the drug product) into four classes:

**Class 1:** High-solubility/High-permeability (HS/HP)
**Class 2:** Low-solubility/High-permeability (LS/HP)
**Class 3:** High-solubility/Low-permeability (HS/LP)
**Class 4:** Low-solubility/Low-permeability (LS/LP).

BCS takes into consideration GI physiological factors, such as pH, gastric fluid volume, gastric emptying, intestinal transit time, etc, and permeability factors. According to the BCS guidance:

- The drug substance is considered highly soluble when the highest dose strength is soluble in 250 ml or less of aqueous media over the pH range of 1–7.5.
- The drug substance is considered highly permeable when the extent of drug absorption in humans is determined to be 90% or more of an administered dose based on a mass balance determination or in comparison to an intravenous reference dose.
- An IR drug product is considered rapidly dissolving when 85% or greater of the labeled amount of the drug substance dissolves within 30 min, using basket method (USP Apparatus 1) at 100 rpm or paddle method (USP Apparatus 2) at 50 rpm in a volume of 900 ml or less in each of the following media:

  (i) 0.1N HCl or simulated gastric fluid without enzymes (ii) a pH 4.5 buffer, and (iii) a pH 6.8 buffer or Simulated Intestinal Fluid without enzymes.

The BCS provides a new perspective to dissolution testing. It provides scientific rationale to lower regulatory burden and justifies a biowaiver under

certain circumstances. The BCS also predicts the possibility of obtaining an in vitro/in vivo correlation.

Justification of a biowaiver is based on a combination of the BCS classification of the drug substance and a drug product dissolution profile comparison. Prerequisites for dissolution based biowaivers are shown in Fig. 8.19. Specifically, to obtain a biowaiver for an IR generic product:

- The reference product should belong to BCS Class 1, HS/HP.
- The test and reference drug products should dissolve rapidly (85% or greater in 30 min or less) under mild test conditions in pH 1.2, 4.5, and 6.8.
- The test and reference products should meet the profile comparison criteria under all test conditions.

For IR products:

1. A biowaiver is applicable for lower strength(s) when the highest strength is shown to be bioequivalent to the innovator product and the formulation(s) of the generic product is (are) proportional to the highest strength and meets dissolution profile comparison criteria.

For MR products:

1. A biowaiver is applicable for beaded capsules when the lower strength differs only in the number of beads of active drug and the dissolution profile is similar in the recommended dissolution test media and conditions.
2. A biowaiver is applicable for extended-release tablet formulations, where the lower strength(s) are compositionally similar to the highest strength the same release mechanism is followed and the dissolution profile is similar at pH 1.2, 4.5 and 6.8.

The biowaiver criteria described in BCS guidance are regarded as very conservative. Discussions are underway to consider relaxing some of the requirements for biowaiver of drug products. These dissolution-based biowaivers exemplify the role of dissolution in regulating pharmaceutical drug products.

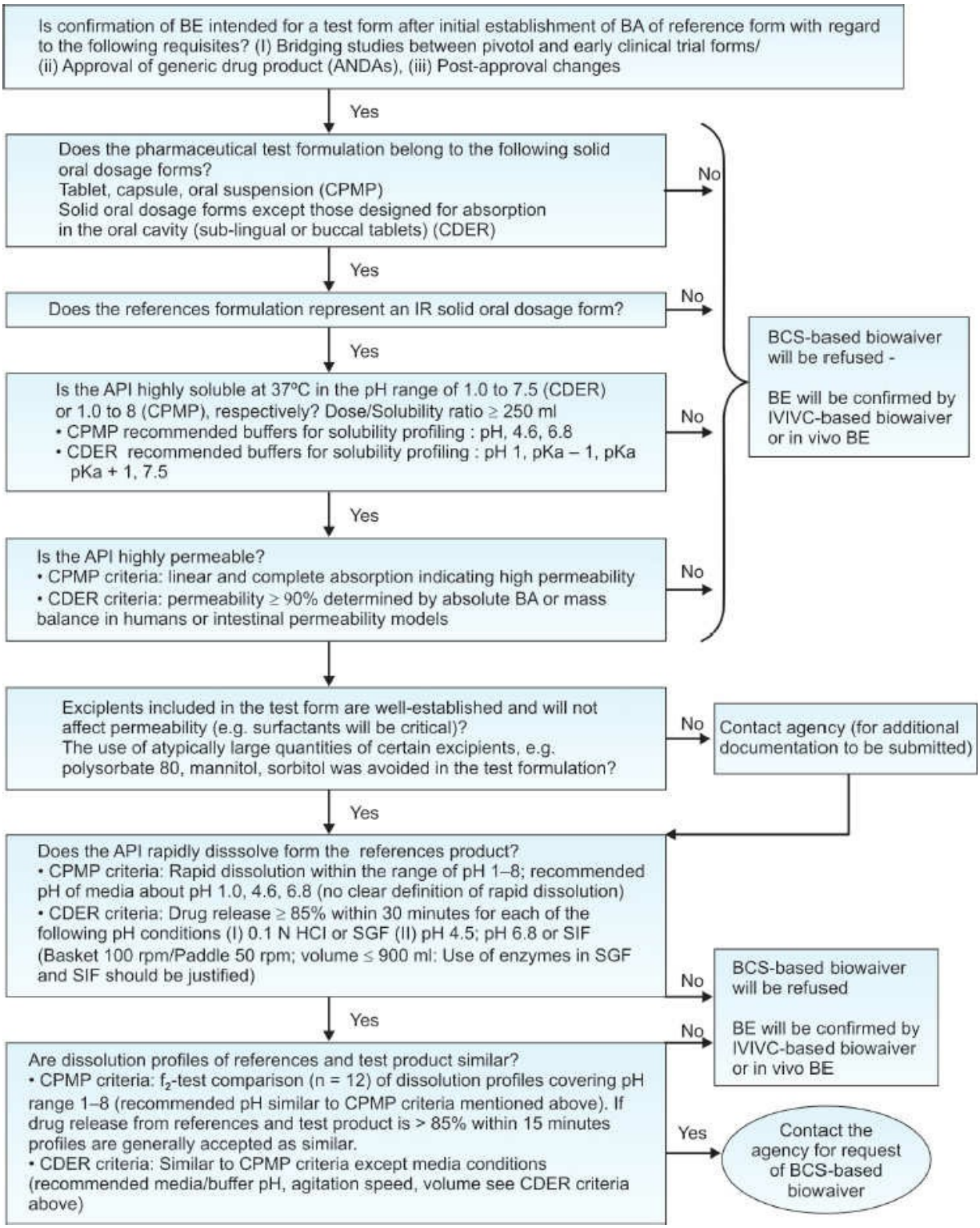



**Fig. 8.19:** Prerequisites for dissolution based biowaivers

## BIORELEVANT DISSOLUTION TEST DESIGN

Dressman and Reppas introduced the concept of using more biorelevant dissolution media, which consist of ingredients that provide physicochemical properties similar to the contents of the human GIT. Their composition is given in Table 8.6 and their use for a particular GI segment is shown in Table 8.7.

**Table 8.6:** Composition of some biorelevant dissolution medium

| **Media constituents** | **Quantity** |
|---|---|
| FaSSGF, pH 1.8 | |
| Sodium chloride | 2 g |
| Hydrochloric acid, conc. | 3 g |
| Triton X 100 | 1 g |
| Deionized water *qs ad* | 1 L |
| Blank FaSSIF, pH 6.5 | |
| $NaH_2PO_4 \times H_2O$ | 3.4.38 g |
| NaCl | 6.186 g |
| NaOH | 0.348 g |
| Deionized water *qs ad* | 1 L |
| Blank FeSSIF, pH 5.0 | |
| Glacial acetic acid | 8.65 g |
| NaCl | 11.874 g |
| NaOH pellets | 4.04 g |
| Deionized water *qs ad* | 1 L |
| SCoF, pH 5.8 | |

| | |
|---|---|
| 1 M Acetic acid | 170 ml |
| 1 M NaOH | 157 ml |
| Deionized water *qs ad* | 1 L |
| SGFsp, pH 1.2 | |
| Sodium chloride | 2 g |
| Hydrochloric acid conc. | 7 g |
| Deionized water *qs ad* | 1 L |
| FaSSIF | |
| Sodium taurocholate | 8.25 g |
| Lecithin | 2.954 g |
| Blank FeSSIF *qs ad* | 1 L |

**Table 8.7:** Biorelevant dissolution mediabased on gastrointestinal site

| **Segment** | **Pre-prandial mediummedium** | **Post-prandial** |
|---|---|---|
| Stomach | FaSSGF | Ensure plus |
| Duodenum | FaSSIF (pH 6) | FaSSIF (pH 5) |
| Upper jejunum | FaSSIF (pH 6.5) | FaSSIF (pH 5) |
| Lower jejunum | FaSSIF (pH 6.8) | FaSSIF (pH 6) |
| Upper ileum | FaSSIF (pH 7.2) | FaSSIF (pH 7.2) |
| Lower ileum | FaSSIF (pH 7.5) | FaSSIF (pH 7.5) |
| Proximal colon | SCoF | SCoF (pH) |
| *pH adjusted by NaOH | | |

Biorelevant medium is a term used to describe a medium that has some relevance to the in vivo dissolution conditions for the compound. Choice of a biorelevant medium is based on a mechanistic approach that considers the

absorption site, if known, and whether the rate-limiting step to absorption is the dissolution or permeability of the compound.

- The composition of Fasted State Simulated Gastric Fluid (FaSSGF) is similar to that of simulated Gastric Fluid without pepsin (SGFsp). However, the pH of FaSSGF is closer to average values of gastric pH observed in the fasted state and a minor amount of a non-ionic surfactant (Triton X 100) has been added to lower the surface tension to that observed in aspirated human gastric juice.
- A composition for the upper small intestine in the fasted state simulated intestinal fluid (FaSSIF), as well as the buffer ($FaSSIF_{blank}$) solution which forms the basis of this mediums. In order to precisely assess the effect of bile salts on solubility and dissolution of the drug substance, results of $FaSSIF_{blank}$ and FaSSIF should be compared. Analogous compositions are also presented for the fed state in the upper small intestine (FeSSIFbl$_{an}$k and FeSSIF).
- To simulate conditions in the lower small intestine, media are buffered at higher pH values and contain progressively lower concentrations of bile salts
- Due to fermentation of undigested carbohydrates by the colonic bacteria, the pH in the proximal colon is usually lower than that of the ileum. This is reflected in the composition of static coefficient of friction (SCoF), which is essentially an acetate buffer. The use of acetate is appropriate as it is known that the products of carbohydrate fermentation include very short chain acids (acetate, propionate, and butyrate are typical).

To challenge the ability of MR dosage forms to resist exposure to high ionic strength, the ionic strength of any of the above-mentioned media can be increased, typically with sodium chloride in the first instance. However, it must be said that the osmolarity in the GI tract rarely falls outside the range 50–600 mOsm/Nm and that if this range is exceeded an artefactual discrimination may result.

A practical feature of these physiologically-based dissolution testing procedures is that they use compendial devices in combination with the biorelevant dissolution media. The procedures thus provide a strong capability for predicting in vivo performance of the drug and/or drug product. More than a mere research project, this technology was proven to be useful as

a surrogate for (BA)/(BE) studies.

In the last 10 years, the use of biorelevant testing conditions has become standard in the characterization of new compounds, development of formulations and development of appropriate quality control tests. Still, there are areas where the biorelevant media can be improved. For example, in the fed state, lipid digestion products may also contribute to the solubilization of lipophilic compounds, thus, inclusion of lipid digestion products in the media would doubt less be of interest for prediction of fed vs fasted state dissolution in vivo. Another continuing area of focus will be the refinement of efforts to predict food effects for MR formulations and to validate the media for various types of MR formulations (hydrogels, osmotic pumps, coated pellets, etc.). All in all, we can be confident that the use of biorelevant media in formulation development will continue to expand and find new applications.

## DEVELOPMENT OF DISSOLUTION TESTS

**Step 1:** Understand the drug's solubility behavior over the usual pH range encountered in the GI tract.

**Step 2:** Determine the solubility in the following pH range.

a. If dose: solubility ratio (D:R) is less than 250 ml at all pH values:

- Drug is considered highly soluble
- Dissolution is very unlikely to limit drug absorption
- A simplified dissolution test then follows the procedures

b. If D:R lies between 250 and 1000 ml at all pH values.

- Drug is still unlikely to exhibit dissolution rate-limited absorption
- This should be confirmed by dissolution of the pure drug in biorelevant media
- To achieve acceptable dissolution in simple buffer solutions, drug is likely to require micronization, use of an appropriate salt form, and/or addition of a small amount of surfactant.
- A development of dissolution tests then follows the procedures

c. If D:R is greater than 1000 ml even in biorelevant media.

- Development of an oral dosage form is going to "require allocation of considerable resources."

## For Highly-soluble Drugs

Drugs can be defined as drugs having good solubility characteristics (i.e. dissolution is unlikely to be rate-limiting to absorption) when D:S<1000 ml across a pH range of approximately 1–7 in simple buffer solutions and D:S<250 ml in biorelevant media. For these compounds, it is often possible to use the same dissolution test procedure throughout the product life cycle. Exceptions to this would include development of a completely different type of dosage form such as an enteric-coated dosage form, MR product or orally disintegrating dosage form, etc. The most appropriate dissolution apparatus for IR products of compounds with good solubility is the USP Apparatus 2.

After establishing that the solubility is appropriately high over a pH range of approximately 1–7 in simple buffer media, the next step is to verify that the dissolution of the pure drug.

## For Very Poorly-soluble Drugs

For very poorly-soluble drugs, dissolution medium may contain a percentage of a surfactant (e.g., sodium lauryl sulfate, Tween 80 or cetyl trimethylammonium bromide (CTAB) that is used to enhance drug solubility. The need for surfactants and the concentrations used should be justified. Surfactants can be used as either a wetting agents or to solubilize the drug substance when the critical micelle concentration (CMC) is reached. If a drug is ionizable, surfactant concentration and pH may be varied simultaneously, and the combined effect can substantially change the solublizing characteristics of the dissolution medium. Table 8.8 lists dissolution medium selection criteria as defined in regulatory, industrial and compendial guidances.

**Table 8.8:** Regulatory, industrial, and compendial guidance for selection of dissolution medium

| Guidance/Compendial reference | Volume | pH | Additives |
|---|---|---|---|
| Federation International Pharmaceutique (FIP) | 500–1000 ml | pH 1–6.8; above pH 6.8 with justification, not to exceed pH 8 | Enzymes, salts, surfactants |
| USP | 500–1000 ml; upto 2000 ml for drugs with limited solubility | Buffers, pH 4–8 or dilute acid (0.001–0.1N HCl) | Salts, surfactants, enzymes for crosslinked capsules |
| WHO, EP, PhEur, JP | Determined per product | Adjust pH within ±0.05 units of prescribed value | Determined per product |
| FDA | 500, 900 or 1000 ml | pH 1–6.8; above pH 6.8 with justification, not to exceed pH 8 | Surfactants, enzymes for crosslinked capsules |

**Step 3. Dissolution of the pure drug:** This test can be simply performed by sprinkling the dose on 500 ml of pre-warmed medium in the paddle apparatus. Test is performed at pH values of about 2 and 6.5, typical of the gastric and small intestinal pH, respectively, in young, healthy subjects (i.e. those with the same GI characteristics as the subjects who will be later enrolled in BA/BE studies). If dissolution of the pure drug powder is

completed in 10–15 min in both media, this is an indication that any well-designed IR formulation (powder, granule, tablet, or capsule) should be able to achieve 85% release of the labeled content within 30 min under similar test conditions. Failure of the pure powder to completely dissolve within 15 min or great variability among samples in the % dissolved at 15 min may indicate that the drug has some wetting problems that should be addressed during formulation. The same test conditions used for the pure drug powder can now be used to compare formulations.

## Section II:

# Pharmaceutical Dosage Form Design

# 9:
# Preformulation

Development of a new drug entity is a concerted effort and it takes US $800 million, and 10 to 15 years to develop a successful drug. In order to initiate safe, efficacious and robust formulation development, important physico-chemical properties of a new drug entity need to be determined early in the discovery process. This is helpful in selecting the optimized molecule for formulation development. Once an optimized molecule is selected for clinical testing, it is important to understand the biopharmaceutical properties of the drug entity with respect to its in vivo absorption and disposition. This information is helpful in making appropriate formulation choices to maximize the drug's bioavailability. Once the formulation is ready to enter clinical trials, it is important to characterize mechanical properties of the formulation to ensure that it can be manufactured at a large scale in a reproducible and effective manner.

Preformulation commences when a newly synthesized drug shows sufficient pharmacologic promise in animal models to warrant evaluation in man. These studies should focus on those physico-chemical properties of the new compound that could affect drug performance and development of an efficacious dosage form. A thorough understanding of these properties may ultimately provide a rationale for formulation design, or support the need for molecular modification. In the simplest case, these preformulation investigations may merely confirm that there are no significant barriers to the compound's development. The objectives of the preformulation study are therefore, (1) to establish the physico-chemical parameters of a new drug entity, (2) to determine its kinetics and stability and (3) to establish its compatibility with common excipients. Preformulation studies may have a significant impact on manufacturing, storage, and performance of drug products. It not only helps to guide dosage form selection, but also provides insights into how drug products should be processed and stored to ensure their quality.

A well designed preformulation study is therefore necessary to fully

characterize molecules during the discovery and development processes so that the new drug entities have the appropriate properties, and there is an understanding of the shortcomings that must be overcome by the formulation process.

## Preliminary Evaluation and Molecular Optimization

Once a pharmacologically active compound has been identified, a project team consisting of representatives from the disciplines indicated in Fig. 9.1 has the responsibility for assuring that the compound enters the development process in its optimum molecular form. Prior to starting preformulation studies, the physical pharmacist should meet with the principal investigators involved in the drug's development to obtain information on the known properties of the compound and the proposed development schedule. Since drug research is usually targeted for a specific therapeutic area, potency relative to competitive products as well as the probable human dosage form(s) may be known. Similarly, the medicinal chemists may have insight regarding the molecule's weaknesses as a result of their efforts to synthesize the compound. In addition, a literature search should be conducted to provide an understanding of the probable decay mechanism(s) and conditions that promote drug decomposition. This information may suggest a means of stabilization, a key stability test, or a stability reference compound (such as aspirin for a compound undergoing ester hydrolysis). Information on the proposed mode of drug administration as well as a literature review on the formulation, bioavailability, and pharmacokinetics of similar drugs often proves useful when deciding in how to optimize the bioavailability of a new drug candidate.

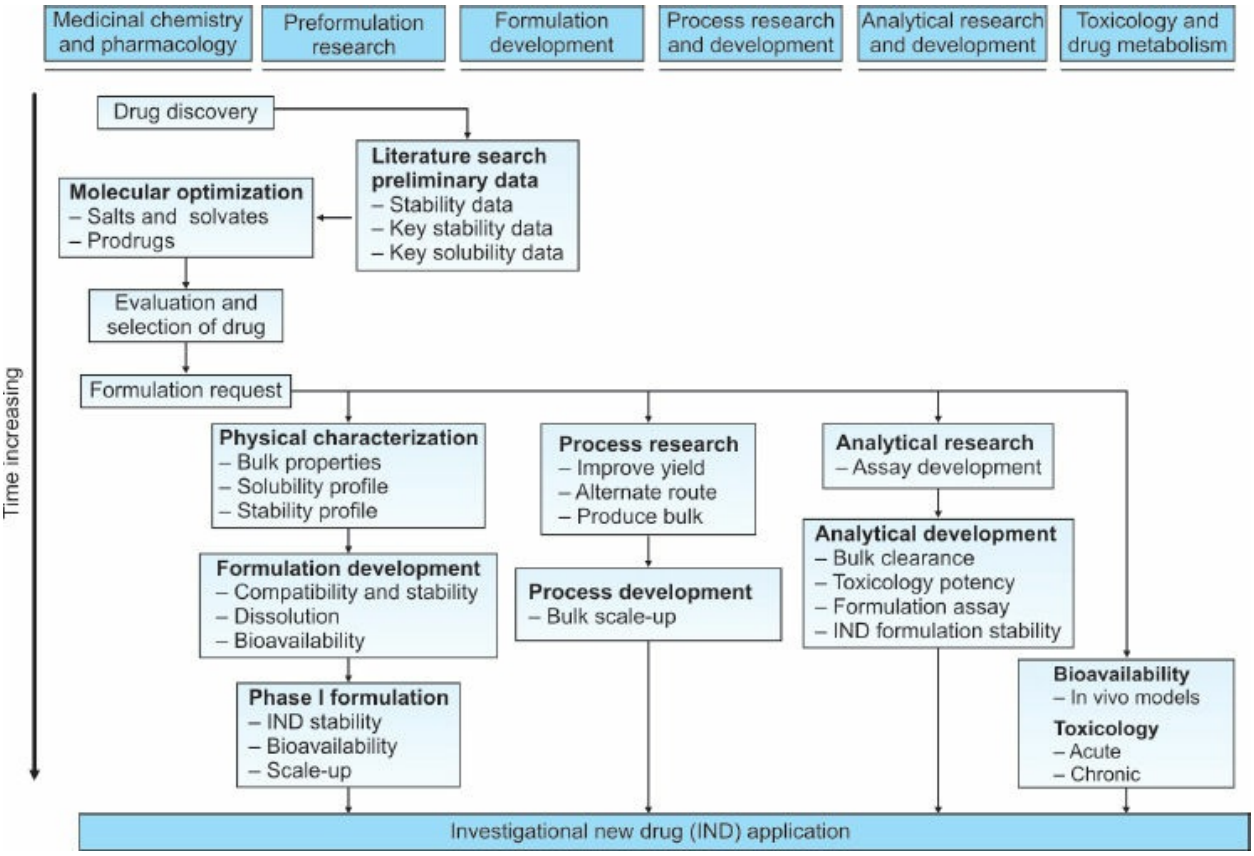



**Fig. 9.1:** Flow diagram illustrating the multidisciplinary development of a drug candidate

The selection of an "optimized molecule" is of utmost importance since various solids differ vastly in physico-chemical properties as listed in Table 9.1. While each discipline may have its own criteria for an "optimized" molecule, the physical pharmacist must focus on how the product will be formulated and administered to the patients. Commonly, stability and/or solubility shortcomings can adversely affect these aspects of drug performance.

**Table 9.1:** Solid state physico-chemical properties of the drug

| | |
|---|---|
| 1. Packing properties | a. Molar volume and density<br>b. Refractive index<br>c. Conductivity electrical and thermal<br>d. Hygroscopicity |

| | |
|---|---|
| 2. Thermodynamic properties | a. Melting and sublimation temperatures<br>b. Internal energy (i.e. structural energy)<br>c. Enthalpy (i.e. heat content)<br>d. Heat capacity<br>e. Entropy<br>f. Free energy and chemical potential<br>g. Thermodynamic activity<br>h. Vapour pressure<br>i. Solubility |
| 3. Spectroscopic properties | a. Electronic transitions (i.e. ultraviolet visible absorption spectra)<br>b. Vibrational transitions (i.e. infrared absorption spectra and Raman spectra)<br>c. Rotational transitions (i.e. far infrared or microwave absorption spectra)<br>d. Nuclear spin transitions (i.e. nuclear magnetic resonance spectra) |
| 4. Kinetic properties | a. Dissolution rate<br>b. Rates of solid state reactions<br>c. Stability |
| 5. Surface properties | a. Surface free energy<br>b. Interfacial tensions<br>c. Habit (i.e. shape) |
| 6. Mechanical properties | a. Hardness<br>b. Tensile strength<br>c. Compactibility tableting<br>d. Handling flow and blending |

When the first quality sample of the new drug becomes available, probing experiments should be conducted to determine the magnitude of each suspected problem area. If a deficiency is detected, then the project team should decide on the molecular modification(s) that would most likely improve the drug's properties. Salts, prodrugs, solvates, polymorphs, or even new analogs may emerge from this modification effort.

## Salt Formation

Although each of these modification approaches has proven beneficial, the salt and pro-drug approaches are the most common. Most salts of organic compounds are formed by the addition or removal of a proton to form an ionized drug molecule, which is then neutralized with a counter ion. Ephedrine hydrochloride, for example, is prepared by the addition of a proton to the basic secondary nitrogen atom on ephedrine, resulting in a protonated drug molecule (ephedrine-$H^{(+)}$), which is neutralized with a chloride anion (ephedrine-HCl). In general, organic salts are more water-soluble than the corresponding unionized molecule, and hence, offer a simple means of increasing dissolution rates, and possibly improving bioavailability. During synthesis, salts are usually formed in organic media to improve yield as well as purity. Some of the problems commonly encountered in evaluating salt forms include poor crystallinity, various degrees of solvation or hydration, hygroscopicity, and instability due to an unfavorable *pH* in the crystalline microenvironment. Hydrochloride, bromide, chloride, iodide, citrate, maleate, pamoate, phosphate, sulfate and tartarate are the most commonly used anions for salt formation whereas, meglumine, calcium, potassium, sodium and zinc are the commonly used cations. Table 9.2 lists a spectrum of pharmaceutical alterations resulting from salt formation.

**Table 9.2:** Examples of modification of pharmaceutical agents through salt formation

| Salt-forming agent | Compound modified | Modification |
|---|---|---|
| Acetylaminoacetic acid | Doxycycline | Solubility |
| N-Acetyl-1-asparagine | Erythromycin | Solubility, activity, stability |
| N-Acetylcysteine | Doxycycline | Combined effect useful in pneumonia |
| Adamantoic acid | Alkylbiguanides | Prolonged action |
| Adipic acid | Piperazine | Stability, toxicity, organoleptic properties |
| N-Alkylsulfamates | Ampicillin | Absorption (oral) |

| | | |
|---|---|---|
| | Lincomycin | Solubility |
| Anthraquinone-l,5-disulfonic acid | Cephalexin | Stability, absorption |
| Arabogalactan sulfate (arabino) | Various alkaloids | Prolonged action |
| Arginine | Cephalosporins | Toxicity |
| | a-Sulfobenzylpenicillin | Stability, hygroscopicity, toxicity |
| Aspartate | Erythromycin | Solubility |
| Betaine | Tetracycline | Gastric absorption |
| Bis (2-carboxychromon- | 7-(Aminoalkyl) | Activity, prolonged prophylactic |
| 5-yloxy) alkanes | theophyllines | effect |
| Carnitine | Metformin | Toxicity |
| 4-Chloro-m-toluenesulfonic acid | Propoxyphene | Organoleptic properties |
| Decanoate | Heptaminol | Prolonged action |
| Diacetyl sulfate | Thiamine | Stability, hygroscopicity |
| Dibenzylethylenediamine | Ampicillin | Prolonged action |
| Diethylamine | Cephalosporins | Reduced pain on injection |
| Diguaiacyl phosphate | Tetracycline | Activity |
| Dioctyl sulfosuccinate | Vincamine | Organoleptic properties |
| Embonic (pamoic) acid | Kanamycin | Toxicity |
| | 2-Phenyl-3-methylmorpholine | Toxicity |
| Fructose 1,6-diphosphoric acid | Tetracycline | Solubility |
| | Erythromycin | Solubility |

| | | |
|---|---|---|
| acid | | |
| Glucose 6-phosphoric acid | Erythromycin | Solubility |
| 1-glutamine | Erythromycin | Solubility, activity, stability |
| Hydroxynaphthoate | Bephenium | Toxicity |
| 2-(4-Imidazolyl) ethylamine | Prostaglandin | Prolonged action |
| Isobutanolamine | Theophylline | Stability |
| Lauryl sulfate | Vincamine | Organoleptic properties |
| Lysine | α-Sulfobenzylpenicillin Cephalosporins | Toxicity, stability, hygroscopicity |
| Methanesulfonic acid | Pralidoxime (2-PAM) | Solubility |
| N-Methylglucamine | α-Sulfobenzylpenicillin | Toxicity, stability, hygroscopicity |
| | Cephalosporins | Reduced pain on injection |
| N-Methylpiperazine | Phenylbutazone | Toxicity, faster onset of action |
| Morpholine | Cephalosporins | Reduced pain on injection |
| 2-Naphthalenesulfonic acid | Propoxyphene | Organoleptic properties |
| Octanoate | Heptaminol | Prolonged action |
| Probenecid | Pivampicillin | Organoleptic properties |
| Tannic acid | Various amines | Prolonged action |
| Theobromine acetic acid | Propoxyphene | Activity |
| 3,4,5-Trimethoxybenzoate | Tetracycline | Organoleptic properties |
| | Heptaminol | Prolonged action |

| | | |
|---|---|---|
| Tromethamine | Aspirin Dinoprost (prostaglandin $F_{2a}$) | Absorption (oral)<br>Physical state |

## Prodrugs

While salt formation is limited to molecules with ionizable groups, prodrugs may be formed with any organic molecule having a chemically-reactive functional group. Prodrugs are synthetic derivatives (e.g. esters and amides) of drug molecules that may have intrinsic pharmacologic activity but usually must undergo some transformation in vivo to liberate the active drug molecule. Through the formation of a prodrug, a variety of side chains or functional groups may be added to improve the biologic and/or pharmaceutical properties of a compound. Some of the biologic response parameters that may be altered by prodrug formation are absorption due to increased lipophilicity or increased water solubility, duration of action via blockade of a key metabolic site, and distribution to organs due to changes in lipophilicity. Examples of biologic improvements are abundant in the steroid- and prostaglandin-prodrug literature. Pharmaceutical improvements resulting from prodrug formation include stabilization, an increase or decrease in solubility, crystallinity, taste, odour, and reduced pain on injection.

Erythromycin estolate is an example of a prodrug with improved pharmaceutical properties (Fig. 9.2). In aqueous solutions, protonated erythromycin is water-soluble, has a bitter taste, and is rapidly hydrolyzed in gastric acid ($t_{10\%}$ = 9 s) to yield inactive decay products. To overcome this problem, the water-insoluble lauryl sulfate salt of the propionate ester prodrug (estolate) was formed for use in both suspension and capsule dosage forms. Erythromycin propionate is inactive as an antimicrobial and must undergo ester hydrolysis to yield bioactive erythromycin. In an oral bioavailability comparison between Upjohn's enteric coated tablet formulation of erythromycin base, E-Mycin, and Dista's non-enteric Ilosone capsule formulation of erythromycin estolate, the lipophilic ester prodrug was absorbed four times more efficiently than the formulated free base, but hydrolyzed only 24% in serum to produce equivalent plasma levels of bioactive erythromycin base. Thus, a prodrug was used to overcome a pharmaceutical formulation problem without compromising bioavailability.

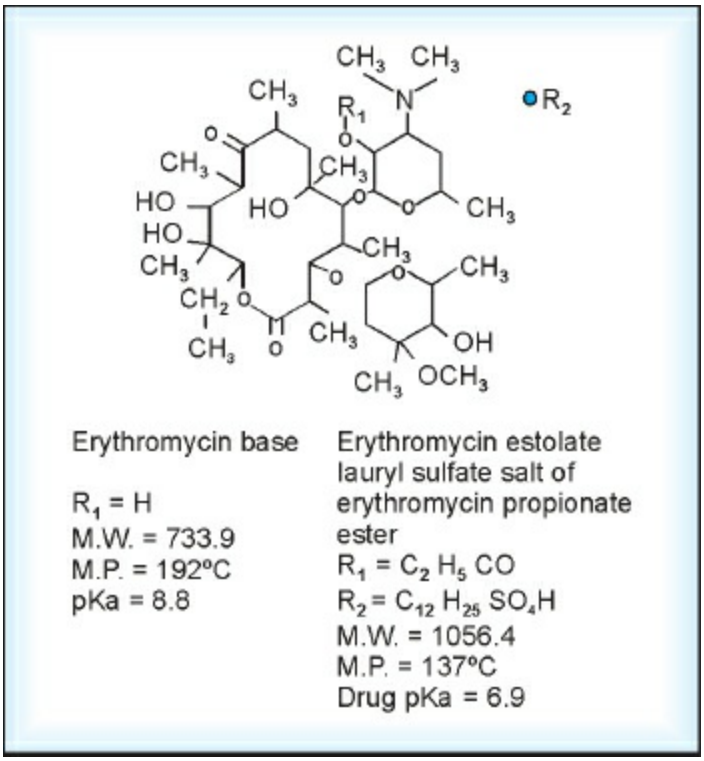



**Fig. 9.2:** Structures of erythromycin and its estolate prodrug

To date, most prodrugs have been esters or amides designed to increase lipophilicity. Unfortunately, this type-of modification often decreases water solubility and thus decreases the concentration gradient across the cell membrane, which controls the rate of drug absorption. In 1980, Amidon suggested the making of water-soluble prodrugs by adding selected amino acids that are substrates for enzymes located in the intestinal brush border. Assuming that enzyme cleavage was not rate limiting, and that the liberated drug molecule would remain in the lipophilic membrane, the resulting membrane transport of the parent compound should be very rapid, owing to the large concentration gradient of the liberated drug across the membrane. Using the lysine ester prodrug of estrone, a potential increase of five orders of magnitude in adsorption rate was found in vivo using perfused rat intestines.

Although, the modifications discussed may provide an increased bioavailability but chemical instability or a lack of synthetic feasibility may prohibit the commercial development of a modified drug molecule. Whatever the case, the molecular form of the drug advancing from this preliminary evaluation should have a substantial chance of successfully progressing

through the drug development process.

Once the optimum molecular form of a drug has been selected, formulation development commences, which prompts other disciplines to begin their task in the drug development process as depicted in Fig. 9.1. The objective of this phase is the quantitation of those physico-chemical properties that will assist in developing a stable, safe, and effective formulation with maximum bioavailability. Figure 9.3 lists the major areas of preformulation research in the order in which they are discussed in the following text. Keep in mind that these topics vary in importance according to the type of formulation sought for each individual drug candidate.

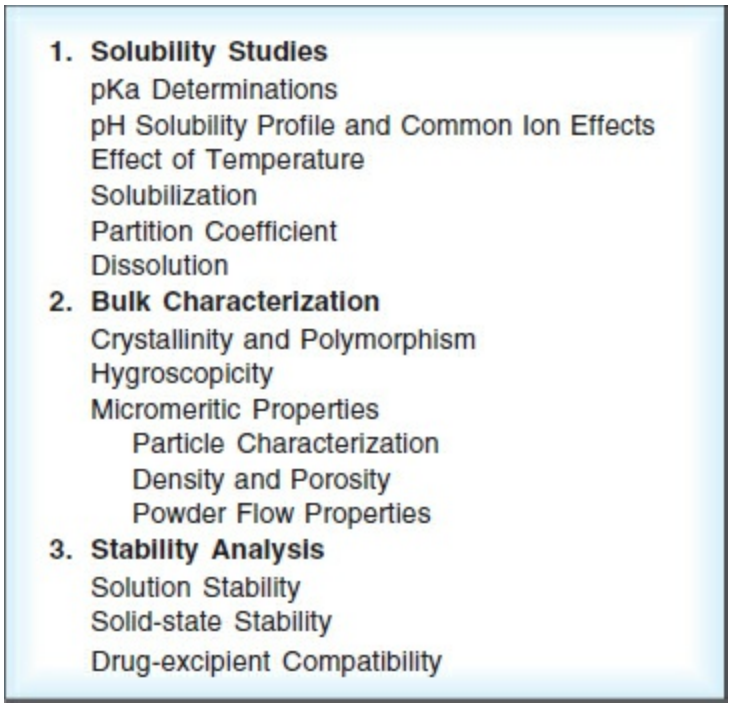



**Fig. 9.3:** Outline of the principal areas of preformulation research

## Solubility Studies

The solubility of a solid substance is defined as the concentration at which the solution phase is in equilibrium with a given solid phase at a stated temperature and pressure. The solubility of a candidate drug may be the critical factor determining its usefulness, since aqueous solubility dictates the amount of compound that will dissolve and, therefore, the amount available for absorption. The bioavailability of an orally-administered drug depends primarily on its solubility in the gastrointestinal tract and its permeability across cell membranes and this forms the basis for the Biopharmaceutical Classification System. In order to be transported across biological membranes, drug molecules are required to be present in a dissolved form. Therefore, low aqueous solubility can delay or decrease drug absorption.

Knowledge of the solubility of a drug is also important during manufacturing of an injectable formulation or when direct administration into the blood stream is desired. For conducting pharmacological, toxicological, and pharmaco-kinetic studies, a drug solution is preferred. Solubility studies also finds application in developing analytical methods for drugs.

Poor aqueous solubility not only limits a drug's biological application, but also challenges its pharmaceutical development. While solubility enhancement remains one of the primary areas of focus during the formulation development phase, there are several situations that may require solubility reduction. Enhancement of chemical stability, taste masking, and development of sustained release products, are examples of such situations. Table 9.3 depicts the solubility based classification of drug as per United State Pharmacopeia.

**Table 9.3:** Solubility based classification of drug

| **Descriptive term** | **Parts of solvent required for 1 part of solute** |
|---|---|
| Very soluble | Less than 1 |
| Freely soluble | From 1 to 10 |
| Soluble | From 10 to 30 |
| Sparingly soluble | From 30 to 100 |
| Slightly soluble | From 100 to 1,000 |

| | |
|---|---|
| Very slightly soluble | From 1,000 to 10,000 |
| Practically insoluble or insoluble | 10,000 and over |

Solubility experiments should have all factors defined, including pH, temperature, ionic strength, and buffer concentrations. The approach for measuring equilibrium solubility is based on the phase-solubility technique proposed by Higuchi and Connors.

An excess drug dispersed in solvent in a closed container is agitated at a constant temperature for a sufficient length of time. Samples of the slurry are withdrawn as a function of time, clarified by centrifugation, and assayed using nephlometry, HPLC, UV spectroscopy, fluorescence spectroscopy, or gas chromatography to establish a plateau concentration. The solid precipitate is also characterized to establish the equilibrium solid form of the drug.

The use of several types of shakers has been reported in the literature, including end-overend shakers, wrist action shakers, magnetic stirrers, vortexers, etc. Proper wetting of the solute has to be ensured by choosing an appropriate shaker. Generally, end-over-end shakers provide good solute-solvent exposure for small volumes of solvents that have a viscosity similar to that of water.

Problems encountered with sample workup can involve adsorption or incomplete removal of the excess drug during filtration or centrifugation steps. In addition, if excess drug is not a solid but an oil, sample preparation may be even more difficult. In particular, drugs capable of ionization may require different methods of removing excess drug, owing to altered adsorption properties. Filtered saturated solutions should be carefully examined using a high-intensity light beam to detect the presence of a finely dispersed oil or solid. Solutions can also be examined conveniently with a light microscope, with which particles or droplets size of 1pm or greater can be distinguished, if present in sufficient concentration.

Solubility values that are useful in a candidate’s early development are those in distilled water, 0.9% NaCl, 0.01M HCl, 0.1M HCl, and 0.1M NaOH, all at room temperature, as well as at *pH* 7.4 buffer at 37°C. These early results are useful in developing suspensions or solutions for toxicologic and pharmacologic studies. Furthermore, these studies identify those candidates which have a potential for bioavailability problems. Drugs having limited

solubility (<1%) in the fluids of the gastrointestinal tract often exhibit poor or erratic absorption unless dosage forms are specifically tailored for the drug. Solubility profiles are not predictors of biologic performance, but do provide rationale for more extensive in vivo studies and formulation development prior to drug evaluation in humans.

Preformulation solubility studies focus on drug-solvent systems that may form during the delivery of a drug candidate. For example, a drug for oral administration should be examined for solubility in media having isotonic chloride ion concentration and acidic pH. Even though the routes of administration may not be explicitly defined at this time, understanding the drug's solubility profile and possible solubilization mechanisms provides a basis for later formulation work. Preformulation solubility studies usually include determinations of *pKa*, temperature dependence, pH-solubility profile, solubility products, solubilization mechanisms, and rate of dissolution.

## *pKa Determinations*

Determination of the dissociation constant for a drug capable of ionization within a *pH* range of 1 to 10 is important since solubility, and consequently absorption, can be altered by orders of magnitude with changing pH. The Henderson-Hasselbalch equation provides an estimate of the ionized and un-ionized drug concentration at a particular pH.

For basic compounds:

$$pH = pKa + \log \frac{[\text{ionized drug}]}{[\text{un-ionized drug}]} \quad \ldots (1)$$

For acidic compounds:

$$pH = pKa + \log \frac{[\text{un-ionized drug}]}{[\text{ionized drug}]} \quad \ldots (2)$$

These equations can be manipulated into the form given by equation to yield the percentage of compound that will be ionized at any particular pH.

$$\%\,\text{Ionization} = \frac{10^{(pH-pKa)}}{1+10^{(pH-pKa)}} \times 100 \quad \ldots (3)$$

- For very weakly-acidic drug (*pKa* > 8) such as phenytoin and ethosuximide, or for very weakly-basic drugs, (*pKa* < 5), such as

diazepam and theophylline, the un-ionized form is present throughout GIT and their absorption is rapid.

- For a weakly-acidic drug (*pKa* ~ 3–7), such as aspirin and ibuprofen, the un-ionized form is present within the acidic contents of the stomach, but the drug is ionized predominately in the neutral media of the intestine. Such drugs are better absorbed from stomach.
- For a weakly-basic drug (*pKa* ~ 5–8), such as morphine and chloroquine the unionized form is present within the basic contents of the intestine, but the drug is ionized predominately in the acidic environment of the stomach. Such drugs are better absorbed from intestine.
- For strongly acidic drugs, such as cromolyn sodium or strongly basic drugs such as erythromycin and papaverine, the ionized form is predominant in both the stomach and intestine.

In general, the un-ionized drug molecule is the species absorbed from the gastrointestinal tract; however, rate of dissolution, lipid solubility, common-ion effects, and metabolism in the GI tract can shift or reverse predictions of the extent and site of absorption based on *pH* alone. A *pKa* value can be determined by a variety of analytical methods. Some of the reported methods for the determination of *pKa* are given in Table 9.4. Buffer, temperature, ionic strength, and cosolvent affect the *pKa* value and should be controlled for these determinations.

**Table 9.4:** Some reported methods for the determination of *pKa*

| |
|---|
| 1. Potentiometric titration |
| 2. UV spectotroscopy |
| 3. Solubility measurement |
| 4. HPLC techniques |
| 5. Capillary zone electrophoresis |
| 6. Foaming activity |

## *Ultraviolet (UV) Spectroscopy*

This method involves measuring the UV spectrum of the compound as a function of *pH*. Mathematical analysis of the spectral shifts can then be used

to determine the *pKa*(s) of the compound. This method is most suitable for compounds containing a UV chromophore that changes with the extent of ionization or for compounds in which the ionizing group is close to or actually within an aromatic ring, which usually results in large UV shifts upon ionization.

## *Potentiometric Titration*

It offers maximum sensitivity for compounds with *pKa* values in the range of 3 to 10 but is often hindered by precipitation of the unionized form during titration since a high drug concentration is usually required to obtain a significant titration curve. One way around this problem is to measure the apparent *pKa* of the compound in solvent and water mixtures, and then extrapolate the data back to a purely-aqueous medium using a Yasuda-Shedlovsky plot as shown in Fig. 9.4. The organic solvents most frequently used are methanol, ethanol, propanol, dimethylsulphoxide (DMSO), dimethyl formamide (DMF), acetone and tetrahydrofuran (THF). In general, the use of cosolvent yields higher *pKa* values for acids and lower values for bases than does pure water. However, methanol is by far the most popular because its properties are closest to water.

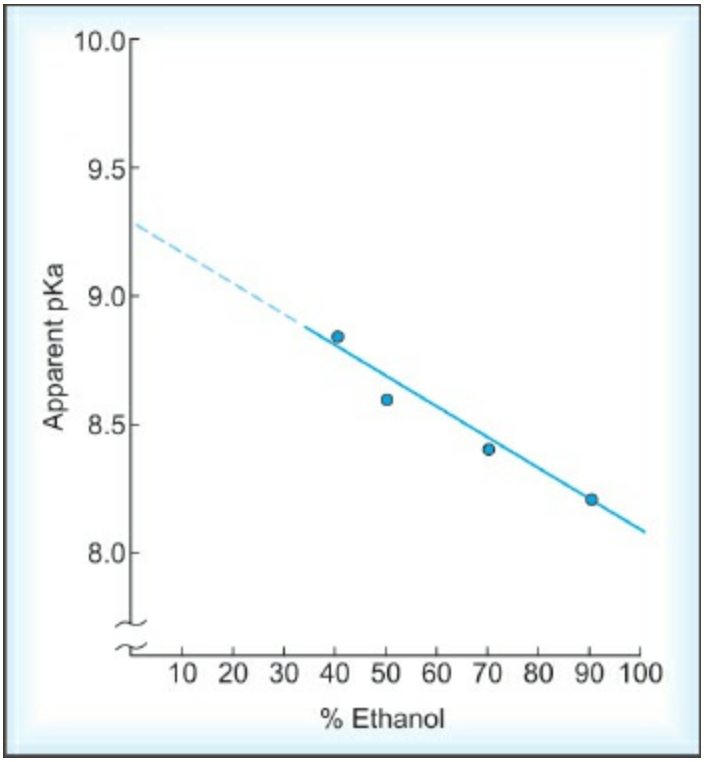


**Fig. 9.4:** The *pKa* determination for an organic amine drug candidate whose un-ionized form is exceedingly insoluble in water

## *Titrimetric Method*

A solution of weakly-acidic drug is titrated with a standard base and *pH* is noted upon each addition of titrant. The *pH* at half neutralization corresponds to *pKa*. If other solubility determining factors, such as solute-solute association or solubility products become significant owing to the titrant used to adjust the *pH*, then *pKa* values may be incorrectly determined by this procedure. Many references are available containing composite lists of *pKa* values for various functional groups and organic molecules, which facilitate initial estimates for dissociation constants. In addition, Albert and Serjeant have provided detailed descriptions of several experimental methods for determining *pKa* values.

## *pH-Solubility Profile and Common Ion Effects*

The solubility of an acidic or basic drug depends on the *pKa* of the ionizing functional group and the intrinsic solubilities for both the ionized and un-ionized forms. For a basic drug, the total molar solubility, $S_t$, is equal to the following summation:

$$S_t = [BH^+] + [B] \quad \ldots (4)$$

where, $BH^+$ is the protonated species and *B* is the free base. The *pH* at which both base and salt species are simultaneously saturated is defined as the $pH_{max}$ as mentioned below:

$$S_t,\ pH = pH_{max} = [BH^+]_S + [B]_s \quad \ldots (5)$$

where, the subscript (*s*) denotes saturation. For weak bases, in the *pH* region where the solubility of the protonated form is limiting, the molar solubility is:

$$S_t, pH < pH_{max} = [BH^+]_s + [B] = [BH^+]_s \left(1 + \frac{k_a}{[H^+]}\right) \quad \ldots (6)$$

Similarly, the solubility in the *pH* region where the free base is limiting is expressed as:

$$S_t, pH > pH_{max} = [BH^+] + [B]_s = [B]_s \left(1 + \frac{H^+}{k_a}\right) \quad \text{... (7)}$$

It therefore follows that the $pH_{max}$ is defined as:

$$pH_{max} = pKa + \log \frac{[B]_s}{[BH^+]_s} \quad \text{... (8)}$$

At a solution *pH* equivalent to $pH_{max}$, both the free base and salt form can exist simultaneously in equilibrium with a saturated solution. The $pH_{max}$ is verified by sampling the precipitated drug from the equilibrated solution and confirming the presence of both drug forms. Solubility expressions for acidic drugs are derived in a similar fashion.

When the ionized or salt form of a drug is the solubility-limiting species in solution, the concentration of the paired counter-ion is usually the solubility-determining factor. For a hydrochloride salt of a basic amine, the equilibrium between the solid and ionized species in solution is approximated by the following expression:

$$[BH{+}Cl^-] \text{ solid} = [BH^+] + [Cl^-] \text{ ... (9)}$$

where, $K_{sp}$ is the solubility product for the protonated species and chloride counter ion, or:

$$K_{sp} = [BH^+]\,[Cl^-] \text{ ... (10)}$$

If the contribution of the un-ionized species is negligible as compared to the protonated form, the total drug solubility decreases as the chloride ion concentration increases. In this case, the apparent solubility product is denned as:

$$K_{sp} = S_t[Cl^-]K_{sp} = S_t[Cl^-] \text{ ... (11)}$$

Experimental determination of a solubility product should include measurement of *pH* as well as assays of both drug and counter ion concentrations. To compute the thermodynamic solubility product, concentrations should be converted to activities with appropriate corrections for activity coefficient dependence on ionic strength.

In summary, aqueous solubility profiles for ionizable compounds over

large *pH* ranges with varying counter ion concentrations are functions of many variables, as given by the following expression for an organic amine drug:

$$S_t = f\,(pH, pKa, [B]_s, K_{sp}, \text{anions}) \ldots (12)$$

These parameters also depend on ionic strength, temperature, and the aqueous media composition. Consequently, pH-solubility profiles can appear dramatically different for compounds with similar functional groups. A particularly useful reference by Kramer and Flynn for organic hydrochlorides highlights *pH* solubility profile dependence on solvent composition. While most of this discussion has focused on solubility product limitations of basic compounds, similar analyses may be made for acidic drugs.

The *pH*-solubility profile for doxycycline (*pKa* 3.4) reported by Bogardus and Backwood illustrates a common ion effect for an amine hydrochloride salt. As shown in Fig. 9.5, the solubility in aqueous medium with *pH* 2 or less, logarithmically decreased as a function of *pH* (which was adjusted with hydrochloric acid) because of corresponding increase in the chloride ion concentration. In gastric juice, where the *pH* can range from 1 to 2 and the chloride ion concentration is between 0.1M and 0.15M, doxycycline hydrochloride dihydrate has a solubility of ~4 mg/ml, which is a factor of 7 less than its solubility in distilled water. For the hydrochloride salts of chlor-tetracycline, demeclocycline, and metha-cyeline, the apparent dissolution rates and solubilities were even less than their respective free base forms in media containing chloride ion. Consequently, the solubility product for each ionizable compound with either sodium or chloride ions should be evaluated to detect potential in vivo problems with dissolution and/or absorption.

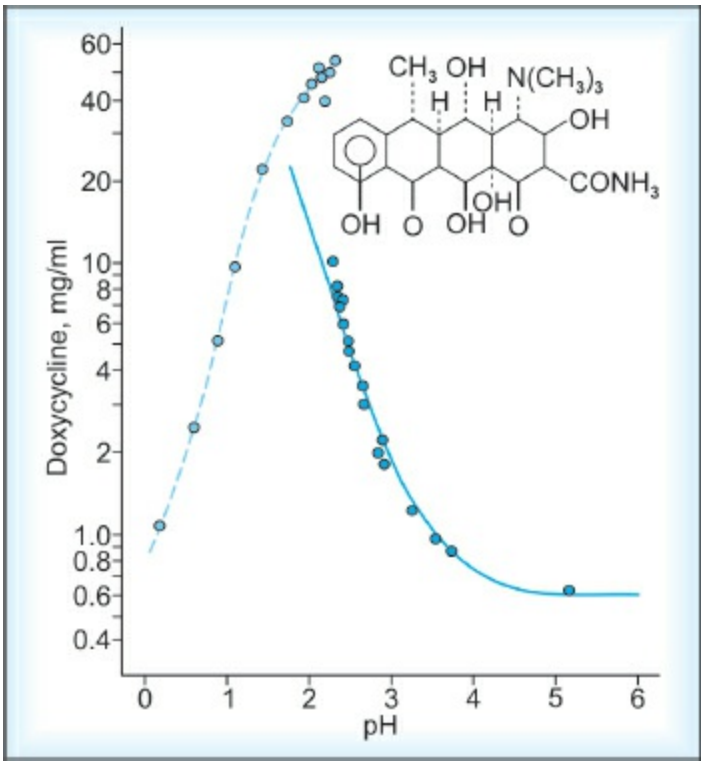



**Fig. 9.5:** The *pH*-solubility profile for doxycycline in aqueous hydrochloric acid at 25°C. At *pH* = 2.16, both doxycycline monohydrate and doxycycline hydrochloride dihydrate were in equilibrium with solution

Another point illustrated with doxycycline is that non-ideal behavior of a solution species can dramatically affect the solubility at certain *pH* values. Doxycycline was shown to form dimeric species involving self-association of the protonated form. This mechanism accounted for the large positive deviation from ideal behavior, in which actual solubility values are a factor of 10 higher at *pH* 2.0. Therefore, actual solubility profiles should be experimentally determined within the *pH* region of interest.

## *Effect of Temperature*

The heat of solution, $\Delta H_S$, represents the heat released or absorbed when a mole of solute is dissolved in a large quantity of solvent. Most commonly, the solution process is endothermic, or $\Delta H_S$ is positive, and thus increasing the solution temperature increases the drug solubility. For such solutes as lithium chloride and other hydrochloride salts that are ionized when dissolved, the process is exothermic (negative $\Delta H_S$) such that higher temperatures suppress the solubility.

Heats of solution are determined from solubility values for saturated solutions equilibrated at controlled temperatures over the range of interest. Typically, the temperature range should include 5°C, 25°C, 37°C, and 50°C. The working equation for determining $\Delta H_S$ is:

$$\ln S = \frac{-\Delta H_s}{R}\left(\frac{1}{T}\right) + C \qquad \text{... (13)}$$

where, *S* is the molar solubility at temperature T (°kelvin) and R is the gas constant. Over limited temperature ranges, a semilogarithmic plot of solubility against reciprocal temperature is linear, and $\Delta H_S$ is obtained from the slope. For non-electrolytes and un-ionized forms of weak acids and bases dissolved in water, heats of solution are usually in the range of 4 to 8 kcal/mole. Salt forms of drugs are often less sensitive to temperature and may have heats of solution between −2 and 2 kcal/mole.

In Fig. 9.6, examples of $\Delta H_S$ determinations are shown for the free base and hydrochloride salt of an organic amine as reported by Kramer and Flynn. Although $\Delta H_S$ for the free base is somewhat large, a 10° change in temperature produces a five-fold change in solubility. This finding would certainly affect solution dosage form design and storage conditions. In addition, solvent systems involving cosolvents, micelles, and complexation have very different heats of solution in comparison to water.

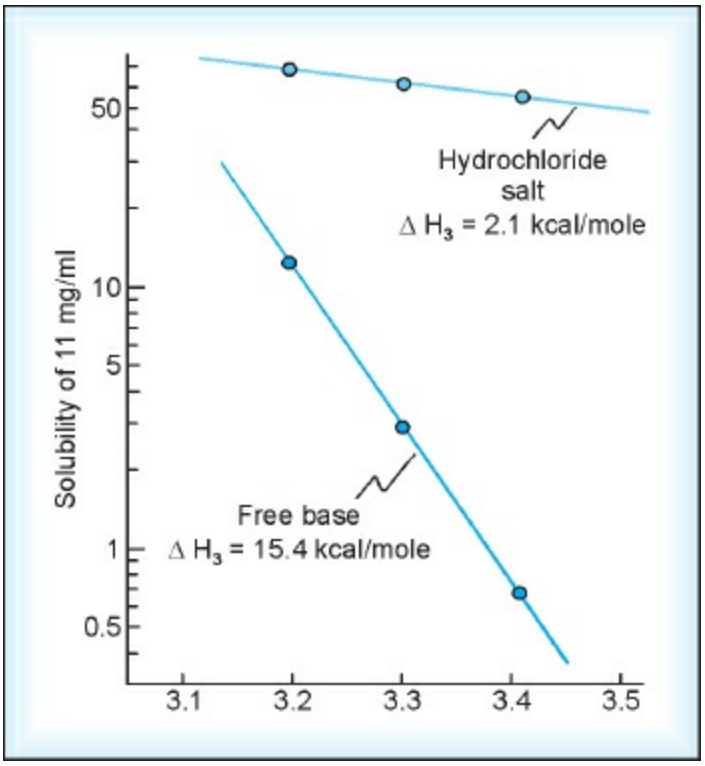



**Fig. 9.6:** Plot of hydrochloride and free base solubilities for etoxadrol, an organic amine, against reciprocal temperature

## *Solubilization*

For drug candidates with either poor water solubility or insufficient solubility for projected solution dosage forms, preformulation studies should include limited experiments to identify the possible mechanisms for solubilization. A general means of increasing solubility is the addition of a cosolvent to the aqueous system. The solubility of poorly soluble non-electrolytes can often be improved by orders of magnitude with suitable cosolvents such as ethanol, propylene glycol, and glycerin. These cosolvents solubilize drug molecules by disrupting the hydrophobic interactions of water at the non-polar solute/water interfaces. The extent of solubilization due to the addition of cosolvent depends on the chemical structure of the drug, that is, the more non-polar the solute, the greater is the solubilization achieved by cosolvent addition. This relationship is illustrated in Fig. 9.7 for hydrocortisone and hydrocortisone-21-heptanoate. The lipophilic ester is solubilized to a greater extent by addition of propylene glycol than by the more polar parent compound.

Cosolvent effects for dissociated drug molecules are usually much less, as shown by Kramer and Flynn. Some poorly-soluble drugs can be solubilized in micellar solutions such as 0.01M Tween 20, or via molecular complexes as with caffeine. However, these specific formulations are usually not developed during the preformulation phase.

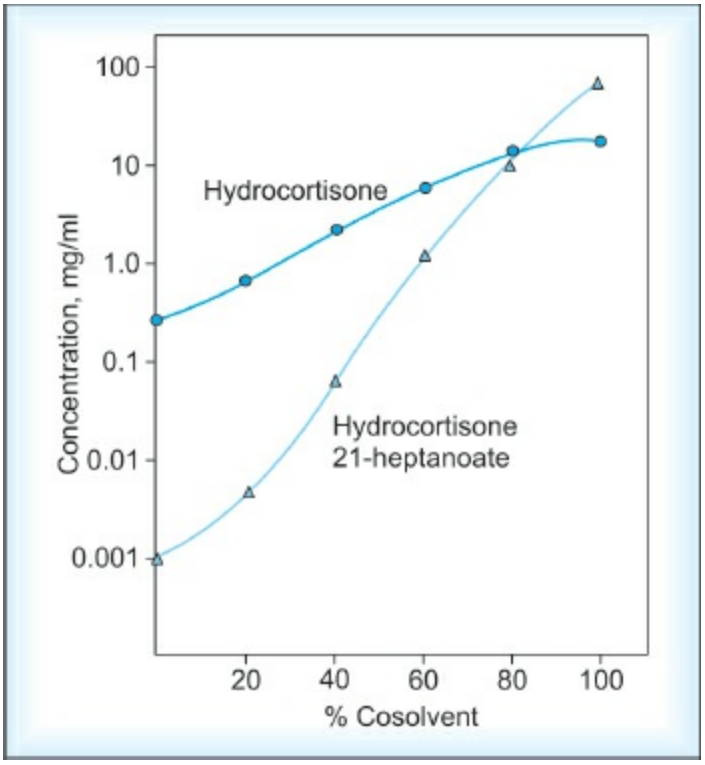



**Fig. 9.7:** Solubility of hydrocortisone and hydrocortisone 21-heptanoate in propylene glycol-water mixtures

## *Partition Coefficient*

A measurement of a drug's lipophilicity and an indication of its ability to cross cell membranes is the oil/water partition coefficient. The partition coefficient, log *P*, is defined as the ratio of un-ionized drug distributed between the organic and aqueous phases at equilibrium.

$$P_{o/w} = \left(\frac{C_{oil}}{C_{water}}\right)\text{equilibrium} \qquad \ldots (14)$$

It is worth noticing that this is a logarithmic scale, therefore, a log *P* = 0 means that the compound is equally soluble in water and in the partitioning solvent. If the compound has a log *P* = −2, then the compound is 100 times

more soluble in water, i.e. it is quite hydrophilic. A log $P = 5$, means that the compound is 100,000 times more soluble in the partitioning solvent.

For ionizable drugs, the distribution of species will depend on the *pH* and the concentration of the ionized drug in the aqueous phase will therefore have an effect on the overall observed partition coefficient. This leads to the definition of the distribution coefficient (log D) of a compound, which takes into account the dissociation of weak acids and bases. For ionizable drugs where the ionized species does not partition into the organic phase, the apparent partition coefficient, *D*, can be calculated from the following:

$$\textit{Acids:}\ \log_{10}D = \log_{10}P - \log_{10}(1+10^{pH-pKa})$$

$$\textit{Bases:}\ \log_{10}D = \log_{10}P - \log_{10}(1+10^{pKa-pH})$$

Figure 9.8 shows the effect of ionization on the partitioning of a proton pump inhibitor drug. The drug has a log P of 3.82 and three *pKa* values, i.e. ≤ 1, 5.26 and 8.63. At low *pH*, the benzimidazole and diethylamine nitrogens are protonated, the drug is ionized and resides in the aqueous phase. As the *pH* increases, deprotonation of the benzimidazole nitrogen takes place, the compound becomes less ionized and resides more in the octanol phase. At neutral *pH*, deprotonation of the diethylamine nitrogen renders the molecule neutral, and hence its lipophilicity is at a maximum. A further increase in the *pH* results in deprotonation of the second nitrogen to form an anion which, being ionized, is more hydrophilic, resulting in a decrease in log D.

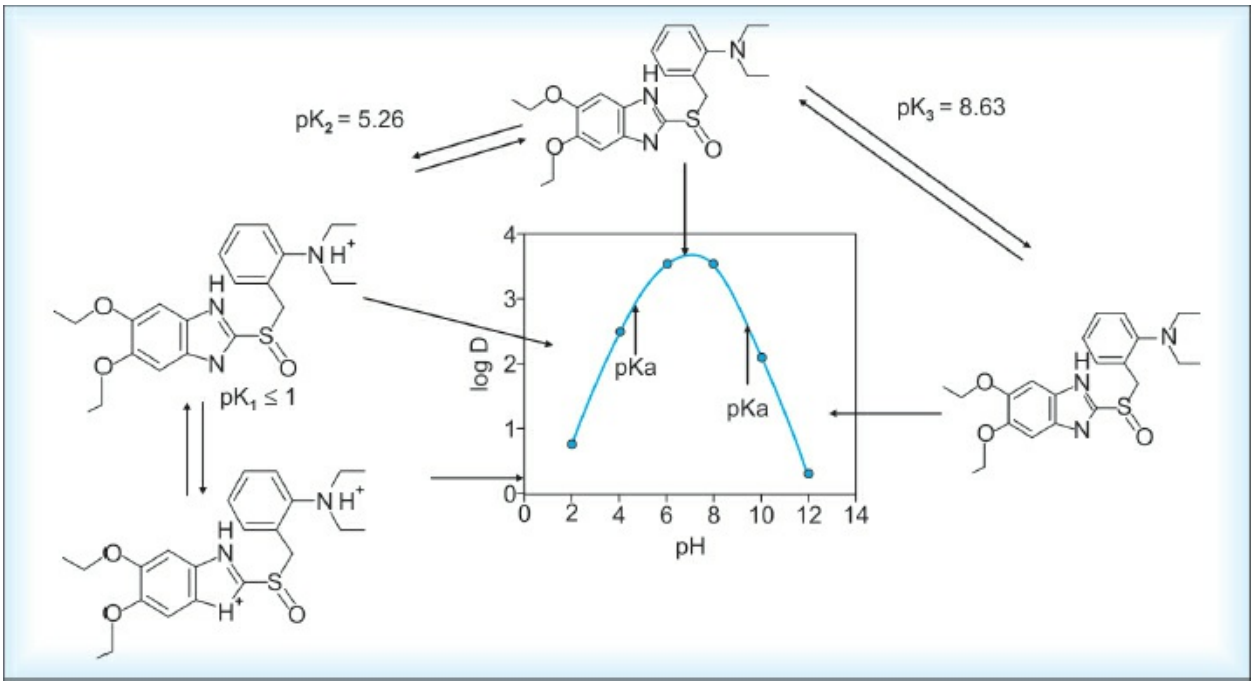



**Fig. 9.8:** Ionization and partitioning scheme for a proton pump inhibitor

Determination of log *P* in a realistic biological medium is virtually impossible. The octanol-water partition coefficient has been widely adopted as a model for defining the relative lipophilicity of a drug. The octanol-water system was selected because octanol is flexible, and contains a polar head and a nonpolar tail, which resembles biological membrane components. Hence, the tendency of a drug molecule to leave the aqueous phase and partition into octanol is viewed as a measure of how efficiently the drug will partition into, and diffuse across, biological barriers. Secondly, octanol and water are immiscible, but some water does dissolve in octanol in a hydrated state. This hydrated state contains 16 octanol aggregates, with the hydroxyl head groups surrounded by trapped aqueous solution. Unionized species dissolve in the aliphatic regions of octanol, whilst ionized species are drawn to the polar regions. Another reason to select octanol is that its solubility parameter (δ = 10.24) lies midway in the range for most of the drugs (δ = 8–12).

Solvents more polar than octanol, such as butanol and pentanol, have been termed hypodiscriminating and reflect transport across buccal membrane. Solvents such as chloroform and cyclohexane, are less polar than octanol, and such hyperdiscriminating solvents reflect transport across blood-brain barrier (Fig. 9.9).

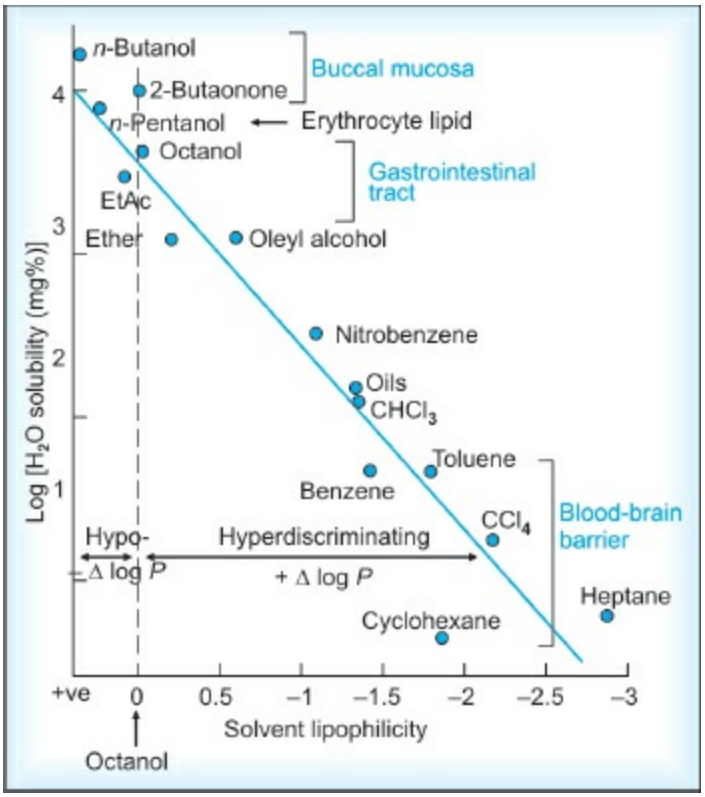



**Fig. 9.9:** Discriminating power of partitioning solvents as a function of their water capacity

The most common method for determining partition and distribution coefficients is the **shake flask method**. In this technique, the candidate drug is shaken between octanol and water layers previously presaturated with each other. Aliquot is taken and analyzed for the drug content. The value of the partition coefficient obtained from this type of experiment is affected by temperature, insufficient mutual phase saturation, *pH* and buffer ions and their concentration, as well as the nature of the solvents used and solute examined.

For a series of compounds, the partition coefficient can provide an empiric handle in screening of some biologic properties. Partition coefficient is an important indicator of permeability, because partitioning of a drug into lipophilic epithelial cells is a necessary step for passive diffusion. Additional factors that appear to influence permeability are polarity and molecular weight. Based on these factors Lipinski and others have proposed an empirical “Rule of Five,” which states that a drug candidate (only applies to drugs that undergo passive diffusion transport) is likely to have poor

permeability or poor oral absorption if a drug exceeds two or more of the following limits:

- Log P is greater than 5
- The molecular weight is greater than 500
- There are more than 5 hydrogen bond donors (number of NH + OH)
- There are more than 10 hydrogen bond acceptors (number of nitrogens + oxygens).

In principle, a drug with high log P tends to have higher membrane permeability as it can partition more easily into the lipid bilayer. However, drugs with excessively high log P may become sequestered in the cell, with little improvement in permeability across the membrane. Polar compounds with an excessive number of hydrogen bond donors and hydrogen bond acceptors form strong interactions with water, which has been correlated to poor permeability. Compounds with very high molecular weight also have poor permeability.

For drug delivery, the lipophilic/hydrophilic balance has been shown to be a contributing factor for the rate and extent of drug absorption. Although partition coefficient data alone does not provide understanding of in vivo absorption, it does provide a means of characterizing the lipophilic/hydrophilic nature of the drug.

## *Dissolution*

Dissolution of a drug particle is controlled by several physicochemical properties, including chemical form, crystal habit, particle size, solubility, surface area, and wetting properties. When coupled with equilibrium solubility data, dissolution experiments can help to identify potential bioavailability problem areas. For example, dissolution of solvate and polymorphic forms of a drug can have a significant impact on bioavailability and drug delivery.

The dissolution rate of a drug substance in which surface area is constant during dissolution is described by the modified Noyes-Whitney equation as follows:

$$\frac{dC}{dt} = \frac{DA}{hV}(C_s - C) \qquad \ldots (15)$$

where, $D$ is the diffusion coefficient, h is the thickness of the diffusion layer at the solid-liquid interface, $A$ is the surface area of drug exposed to the dissolution media, $V$ is the volume of media, $C_s$ is the concentration of a saturated solution of the solute in the dissolution medium at the experimental temperature, and $C$ is the concentration of drug in solution at time t. The dissolution rate is given by $dC/dt$. If the surface area of the drug is held constant and $C_s >> C$, then equation (15) can be rearranged and integrated to give the working equation:

$$\frac{W}{A} = kt \qquad \ldots (16)$$

where, the constant $k$ is defined as:

$$k = \frac{D}{h} C_s \qquad \ldots (17)$$

and $W$ is the weight (mg) of drug dissolved in time t. A plot of $W$ versus $t$ gives a straight line with the slope equal to the intrinsic dissolution rate constant $k$, usually expressed in units of mg/cm$^2$/min.

Experimentally, a constant surface area is obtained by compressing powder into a disc of known area with a die and punch apparatus. Either of the two systems shown in Fig. 9.10 can be used to maintain uniform hydrodynamic conditions ($k$ constant). The rotating disc method or Wood's apparatus permits the hydrodynamics of the system to be varied in a mathematically well-defined manner. The static disc method is used because it is conveniently available, but it contains an element of undefined turbulence, which necessitates calibration with standards. Potential problems with this method are transformations of the crystal form, such as polymorphic transformations or desolvation, during its compression into a pellet or during the dissolution experiment. Since many drug candidates are weak acids and bases, *pH* and common ion gradients at the solid-liquid interface can lead to erroneous conclusions, as discussed by Mooney and co-workers.

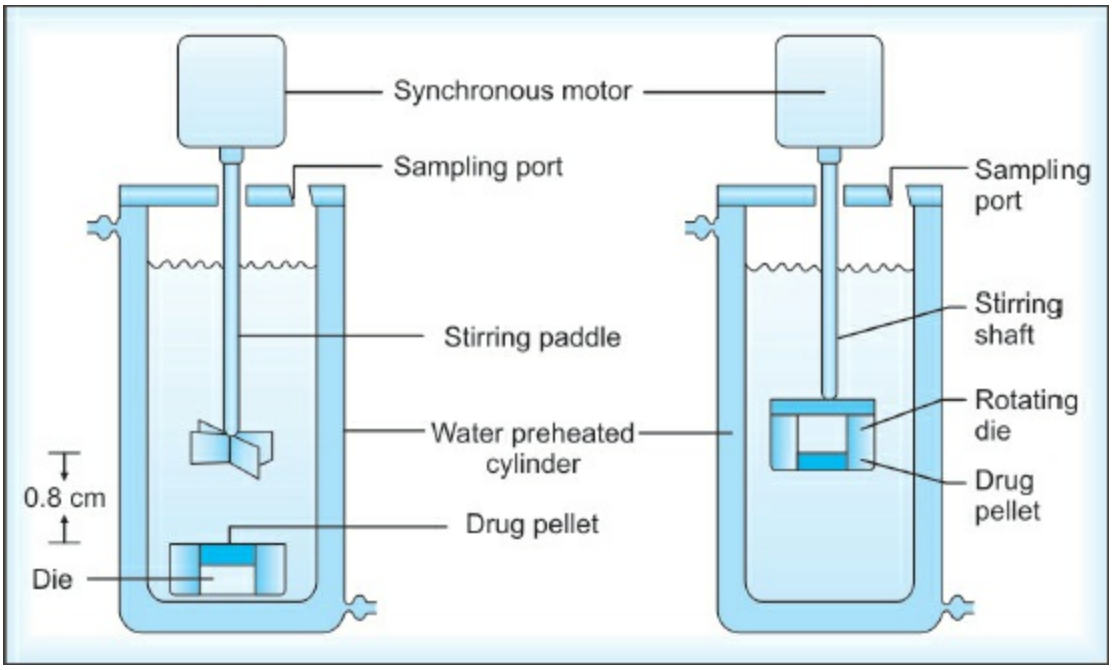



**Fig. 9.10:** Constant surface area dissolution apparatus. Left: Static-disc dissolution apparatus. Right: Rotating disc apparatus

Dissolution experiments with drug suspensions are further complicated by changing surface area, changing surface crystal morphology, and interstitial wetting. However, dissolution profiles with excess drug can be used to characterize metastable polymorphs or solvates. In Fig. 9.11, the conversion of the metastable form II to form I is shown to occur in an organic solvent medium, which clearly depicts form I as the thermodynamically stable form at room temperature. Static pellet dissolution rates also substantiated that form II was the higher energy form since its dissolution rate was significantly greater (Table 9.5).

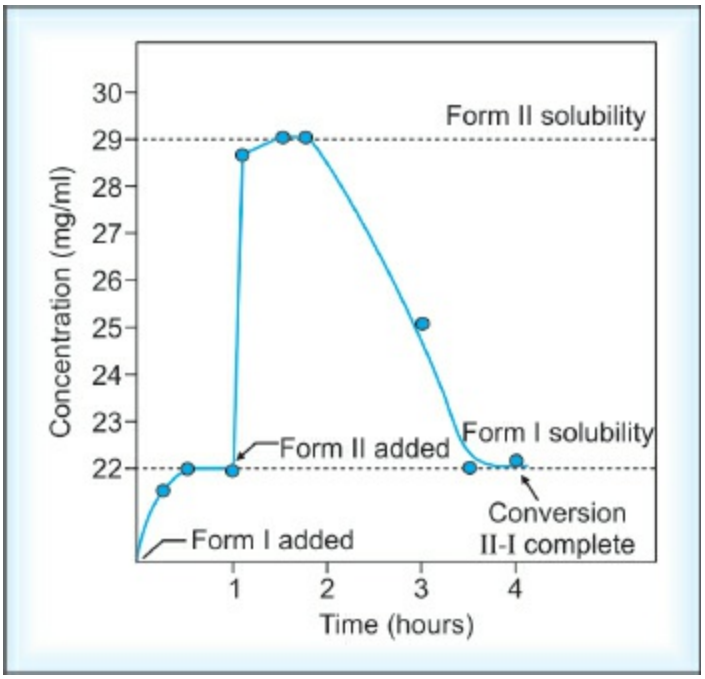



**Fig. 9.11:** Powder dissolution profiles for two polymorphic forms of an organic acetate salt in acetonitrile at 25°C

**Table 9.5:** Comparison of dissolution rates and solubility for two polymorphic forms of an organic acetate salt

| | | Ratio | |
|---|---|---|---|
| | **Form I** | **Form II** | **I/II** |
| Dissolution rate in ethanol ($mg/cm^2/min$) | 0.57 | 0.69 | 0.83 |
| Dissolution rate in acetonitrile ($mg/cm^2/min$) | 0.017 | 0.027 | 0.62 |
| Solubility in acetonitrile (mg/ml) | 0.22 | 0.29 | 0.75 |

## Bulk Characterization

In most instances, the synthetic process is developed in parallel with preformulation investigations. A drug candidate at this stage often does not has all of its solid forms identified, and there is a great potential for new polymorphs to emerge. Bulk properties for the solid form, such as particle size, bulk density and surface morphology, are also likely to change during process development. Therefore, comprehensive characterization of all preformulation bulk lots is necessary to avoid misleading predictions of stability or solubility, which depend on a particular crystalline form.

### *Crystallinity and Polymorphism*

Crystal habit and the internal structure of a drug can affect bulk and physicochemical properties, which range from flowability to chemical stability. *Habit* is the description of the outer appearance of a crystal, whereas the *internal structure* is the molecular arrangement within the solid. Several examples of habits of crystals are shown in Table 9.6 and some of them are represented in Fig. 9.12. A single internal structure for a compound can have several different habits, depending on the environment for growing crystals. Changes with internal structure usually alter the crystal habit while such chemical changes as the conversion of a sodium salt to its free acid form produce both a change in internal structure and crystal habit. Characterization of a solid form involves (1) verifying that the solid is the expected chemical compound, (2) characterizing the internal structure, and then (3), describing the habit of the crystal.

**Table 9.6:** Examples of crystal habits

| Habit | Description |
|---|---|
| Acicular | Elongated prism, needle-like |
| Angular | Sharp edged, roughly polyhedral |
| Bladed | Flattened acicular |
| Crystalline | Geometric shape fully developed in fluid |
| Dendritic | Branched crystalline |
| Fibrous | Regular or irregular thread-like |
| Flaky/platy | Plate or salt-like |

| | |
|---|---|
| Granular | Equidimensional irregular shape |
| Irregular | Lacking any symmetry |
| Nodular | Rounded irregular shape |
| Prismatic | Columnar prism |
| Spherical | Global shape |
| Tabular | Rectangular with a pair of parallel faces |

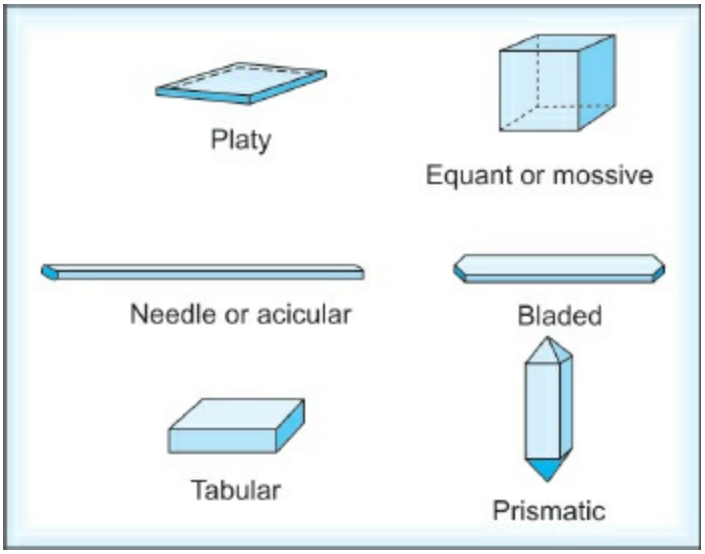



**Fig. 9.12:** Different habits of crystals

The internal structure of a compound can be classified in a variety of ways, as shown in Fig. 9.13. Many possibilities exist for transformations among these forms. Some phase transformations are thermodynamically favoured, and are therefore spontaneous, whilst some phase transformations are thermodynamically disfavored, occur only under stress, and require energy input from the environment. The major types of phase transformations are listed in Table 9.7 and the underlying mechanisms for phase transformations are listed in Table 9.8.

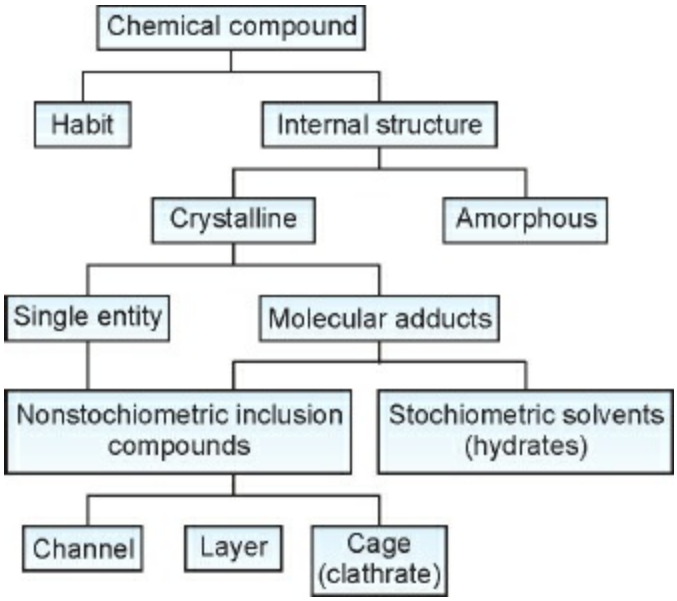



**Fig. 9.13:** Outline of differentiating habit and crystal chemistry of a compound

**Table 9.7:** Various types of phase transformations of a compound

| Type | Explanation of phase transformation |
|---|---|
| A | *Polymorphic transition:* Transition between the polymorphs. The crystalline phases include all types of crystail-line solids. The composition of the solid remains the same. |
| B | *Hydration/dehydration:* Transition between anhydrates and hydrates or hydrates of different stoichiometry. The compositions of the solids differ by the number of water molecules. |
| C | *Solvation/desolvation:* Transition between solvent-free crystal forms and solvates, solvates of different stoichiometry or solvates of different nature (i.e. different solvents are incorporated into the crystalline lattice). The compo-sition of the solids differ by the nature and number of solvent molecules. |
| D | *Salt/parent conversions or salt/salt exchange:* Transition between the salts (ionic adducts) and the parent unionized compounds (free acids or free bases) between the salts of |

| | |
|---|---|
| | different stoichiometry or between the different salts. The compositions of the solids differ by the nature and number of counterions. |
| E | *Co-crystal/parent conversions or co-crystal exchange:* Transition between the co-crystals (molecular adducts), and the parent compound (unionized compounds or salts), between the co-crystals of different stoichiometry or between the different co-crystals. The compositions of the solids differ by the nature and number of co-crystal formers. |
| F | *Amorphous crystallization/vitrification:* Transition between crystalline and amorphous phases. The crystalline phases include all types of crystalline solids. Since the compositions of the amorphous phases are usually less well-defined, the compositions of the solids change in most cases. |

**Table 9.8:** Mechanism of phase transformations

| **Mechanism** | **Types of phase transformations** |
|---|---|
| Solid-state | A-F |
| Melt | A, B and C (dehydration/desolvation only), D, E, F (vitrification only) |
| Solution | A-F |
| Solution-mediated | A-E, F (amorphous crystallization only); transitions occur only from the metastable state to the stable state under the defined conditions |

The first major distinction is whether the solid is crystalline or amorphous. Crystals are characterized by repetitious spacing of constituent atoms or molecules in a threedimensional array, whereas amorphous forms have atoms or molecules randomly placed as in a liquid. ***Amorphous forms*** are typically prepared by rapid precipitation, lyophilization, or rapid cooling of liquid melts. Since amorphous forms are usually of higher thermodynamic energy than the corresponding crystalline forms, solubilities as well as

dissolution rates are generally greater. Upon storage, amorphous solids tend to revert to more stable forms. This thermodynamic instability, which can occur during bulk processing or within dosage forms, is a major disadvantage for developing an amorphous form.

A **crystalline** compound may contain either a stoichiometric or non-stoichiometric amount of crystallization solvent. Non-stoichiometric adducts, such as **inclusions or clathrates**, involve entrapped solvent molecules within the crystal lattice. Usually, this adduct is undesirable, owing to its lack of reproducibility, and should be avoided for development. A stoichiometric adduct, commonly referred to as a **solvate**, is a molecular complex that has incorporated the crystallizing solvent molecules into specific sites within the crystal lattice. When the incorporated solvent is water, the complex is called a **hydrate**, and the terms hemihydrate, monohydrate, and dihydrate describe hydrated forms with molar equivalents of water corresponding to half, one, and two. A compound not containing any water within its crystal structure is termed **anhydrou**s.

Identification of possible hydrate compounds is important since their aqueous solubilities can be significantly less than their anhydrous forms. Conversion of an anhydrous compound to a hydrate within the dosage form may reduce the dissolution rate and extent of drug absorption.

An example of the in vivo importance of solvate forms is shown in Fig. 9.14, where the anhydrous and trihydrate forms of ampicillin were administered orally as a suspension to human subjects. The more soluble anhydrous form (10 mg/ml) produced higher and earlier peaks in the blood serum levels than the less soluble trihydrate form.

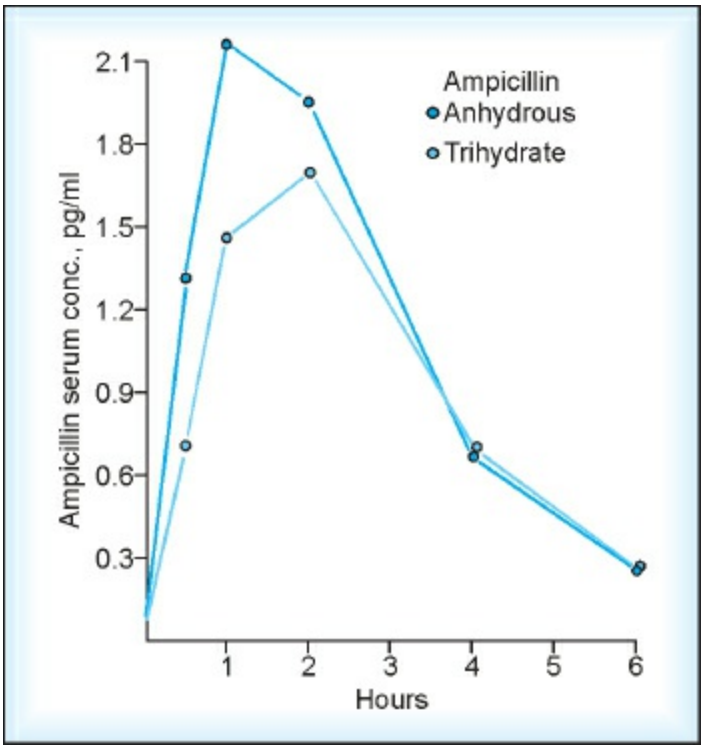



**Fig. 9.14:** Mean serum concentrations of ampiciNin in human subjects after oral administration of 250 mg doses of two solvate forms of the drug in suspension. Key: • anhydrous; and • trihydrate

**Polymorphism** is the ability of a compound (or element) to crystallize as more than one distinct crystalline species with different internal lattices. Chemical stability and solubility changes due to polymorphism can have an impact on a drug's bioavailability and its development program. Chloramphenicol palmitate exists in three crystalline polymorphic forms (A, B, and C) and an amorphous form. Aguiar and co-workers investigated the relative absorption of polymorphic forms A and B from oral suspensions administered to human subjects. As summarized in Fig. 9.15, "peak" serum levels increased substantially as a function of the percentage of form B polymorph, the more soluble polymorph.

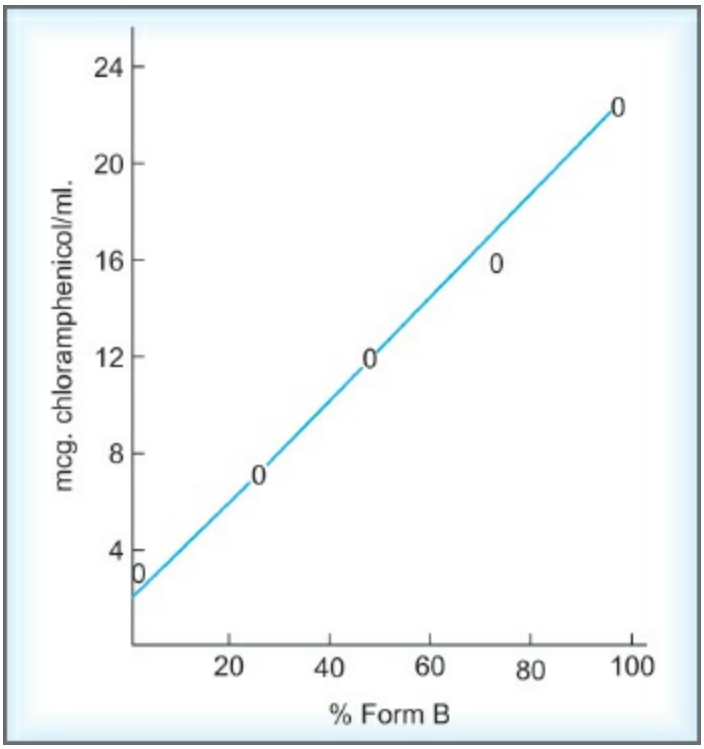



**Fig. 9.15:** Correlation of "peak" blood serum levels (2 h) of chloramphenicol vs percentage of concentration of polymorph B

Polymorphs can be classified as one of two types: *enantiotropic* (one polymorph can be reversibly changed into another by varying temperature or pressure, e.g. sulfur) or *monotropic* (one polymorphic form is unstable at all temperatures and pressures, e.g. glyceryl stearates). The thermodynamic difference between enatiotropic and monotropic polymorphic transitions is shown in Table 9.9.

**Table 9.9:** Thermodynamic difference between polymorphs

| **Enantiotropy** | **Monotropy** |
|---|---|
| Transition < melting I | Transition > melting I |
| I Stable > transition | I Always stable |
| II Stable < transition | – |
| Transition reversible | Transition irreversible |
| Solubility I higher < transition | Solubility I always lower |
| Solubility I lower > transition | – |

| | |
|---|---|
| Transition II → I endothermic | Transition II → I exothermic |
| $\Delta H_f^I < \Delta H_f^{II}$ | $\Delta H_f^I > \Delta H_f^{II}$ |
| IR peak I before II | IR peak I after II |
| Density I < II | Density I > density II |

There is no general way of relating enantiotropy and monotropy to the properties of the polymorphs, except by locating the transition temperature or the lack of one. At a specified pressure, usually 1 atmosphere, the temperature at which two polymorphs have identical free energies is the transition temperature, and at that temperature, both forms can coexist and have identical solubilities in any solvent as well as identical vapour pressures. Below the solid melting temperatures, the polymorph with the lower free energy, corresponding to the lower solubility or vapour pressure, is the thermodynamically stable form.

During preformulation, it is important to identify the polymorph that is stable at room temperature and to determine whether polymorphic transitions are possible within the temperature range used for stability studies and during processing (drying, milling, etc.). As discussed by Haleblian and McCrone, a polymorphic compound is best characterized by a complete pressure-temperature phase diagram showing melt-vapour, solid-vapour, and solid-melt curves. A free energy-temperature curve at 1 atmosphere should be constructed since temperature is usually a more critical variable than pressure in pharmaceutics. As previously discussed, chloramphenicol palmitate has three known polymorphic forms, which are thermodynamically described by a van't Hoff plot of free energy (as determined from solubility measurements) versus temperature (Fig. 9.16). Transition temperatures are shown by intersection of the extrapolated lines; 50°C for forms A and C, and 88°C for forms A and B. Form A is the stable form at temperatures less than 50°C.

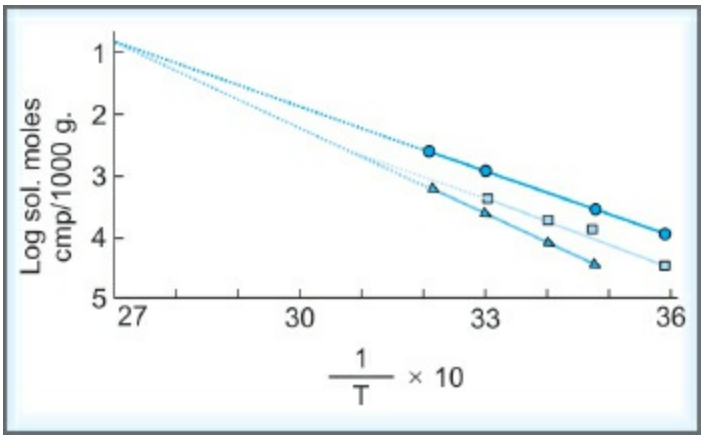



**Fig. 9.16:** The van't Hoff plot of solubility vs. reciprocal absolute temperature for polymorphs A, B, and C of chloramphenicol palmitate. Key: Polymorphs A (▲); B (●); and C (■)

Transition temperatures obtained by extrapolation of van't Hoff plots are susceptible to large errors. Direct measurements of transitions are preferred to support the extrapolated intersection points in the solubility-temperature diagrams. The most direct means for determining transition temperatures is microscopic observation of samples held at constant temperatures. Unfortunately, these solid-solid or solid-vapour-solid transitions usually occur slowly, owing to large activation energies and slow nucleation. To facilitate the conversion rate, a single polymorph or a mixture of forms can be granulated in a "bridging" solvent at various temperatures. The drug should be only sparingly soluble in the bridging solvent, and solvate formation should not occur. These experiments can be conducted quickly with a polarizing microscope, or samples can be stored in sealed containers at controlled temperatures and periodically examined by other suitable analytical methods.

A more difficult task in the study of polymorphism is the determination of the relative stability of metastable polymorph and prediction of its rate of conversion within a dosage form. For suspension dosage forms, the rate of conversion can depend on several variables, including drug solubility within the vehicle, presence of nucleation seed for the stable form, temperature, agitation, and particle size. Solid dosage forms such as capsules and tablets have similar complications due to the influence of particle size, moisture, and excipients. In short, the most effective means for evaluating the stability of a metastable polymorph in the dosage form is to initiate prototype formulation

work by screening a wide range of factors, including the presence and absence of seed crystals of the stable polymorphic form. Essential to this approach is the development of an analytical method that is sensitive to small amounts of stable polymorph in the presence of the metastable polymorphs and excipients. In most cases, the lower limit of detection for polymorph mixtures is in the range of 2 to 5%.

To screen for additional polymorphic forms of a particular drug, bridging solvents, supersaturated solutions, supercooled melts and sublimination have proven useful. Observation of the drug particles by light microscopy during or after processing by these techniques should provide substantial insight into the preferred crystalline forms of the compound without consuming inordinate quantities of material.

Many physico-chemical properties vary with the internal structure of the solid drug, including melting point, density, hardness, crystal shape, optical properties, and vapour pressure. Characterization of polymorphic and solvated forms involves quantitative analysis of these differing physicochemical properties. Several methods for studying solid forms are listed in Table 9.10 along with the sample requirements for each test. In the following sections, three of these techniques are discussed in detail, with particular emphasis on polymorphism.

**Table 9.10:** Analytical methods for characterization of solid forms

| **Method** | **Material required per sample** |
|---|---|
| Microscopy | 1 mg |
| Fusion methods (hot stage microscopy) | 1 mg |
| Differential scanning calorimetry (DSC/DTA) | 2–5 mg |
| Infrared spectroscopy | 2–20 mg |
| X-ray powder diffraction | 500 mg |
| Scanning electron microscopy | 2 mg |
| Thermogravimetric analysis | 10 mg |
| Dissolution/solubility analysis | mg to gm |

**Microscopy:** All substances that are transparent when examined under a 

microscope that has crossed polarizing filters are either isotropic or anisotropic. Amorphous substances, such as supercooled glasses and noncrystalline solid organic compounds, or substances with cubic crystal lattices, such as sodium chloride, are isotropic materials, which have a single refractive index. With crossed polarizing filters, these isotropic substances do not transmit light, and they appear black. Materials with more than one refractive index are anisotropic and appear bright with brilliant colours (birefringence) against the black polarized background. The interference colours depend upon the crystal thickness and the differences in refractive indices. Anisotropic substances are either uniaxial, having two refractive indices, or biaxial, having three principal refractive indices.

Most drugs are biaxial, corresponding to either an orthorhombic, monoclinic, or triclinic crystal system. Although one refractive index is easily obtained for biaxial systems, proper orientation of the crystal along its crystallographic axes is required to describe the crystalline form completely. Owing to the many possible crystal habits and their appearances at different orientations, these methods require a well-trained optical crystallographer to characterize fully even simple biaxial systems. Crystal morphology differences between polymorphic forms, however, are often sufficiently distinct so that the microscope can be used routinely by the less experienced microscopist to describe polymorphic crystal habits and observe transitions induced by heat or solvents.

**Hot-stage Microscopy (HSM):** The polarizing microscope fitted with a hot stage is a useful instrument for investigating polymorphism, melting points, transition temperatures, and rates of transition at controlled heating rates. HSM is a fusion technique whereby approximately one milligram of material is spread on a microscope slide which is then placed on the hot stage and heated. The sample can be heated at different rates in the sample chamber, and the atmosphere can be controlled. Sample chamber also allows the light from the microscope to pass through the sample. Direct observation by HSM can sometimes reveal subtle changes not readily detected by the other instrumental techniques. Figure 9.17 shows a simplified schematic diagram of a HSM. In conjugation with DSC, HSM facilitates differentiation of DSC endo-therms for polymorphic transitions. Figure 9.18 (HSM photographs) and 9.19 (DSC thermogram) shows the sequence of events recorded on heating a sample of carbamazepine (A). On heating, the sample melted (B) (first endo-therm recorded in the DSC thermogram). As the

sample was further heated compound recrystallized from the melt as acicular crystals (C). The acicular crystals continued to grow until the second crystal form of the compound melted (D) (second, large endo-therm on the DSC thermogram).

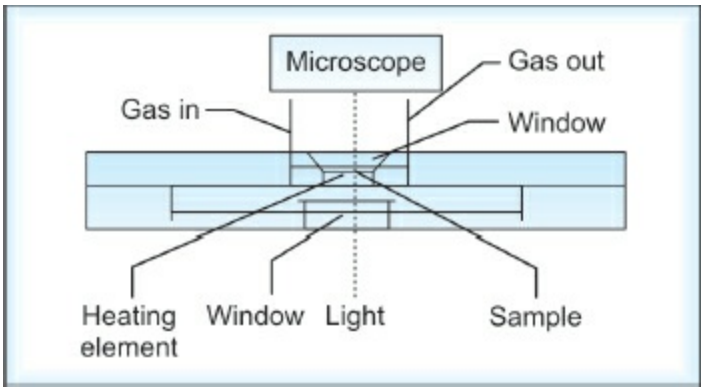



**Fig. 9.17:** Schematic diagram of a hot stage microscopy

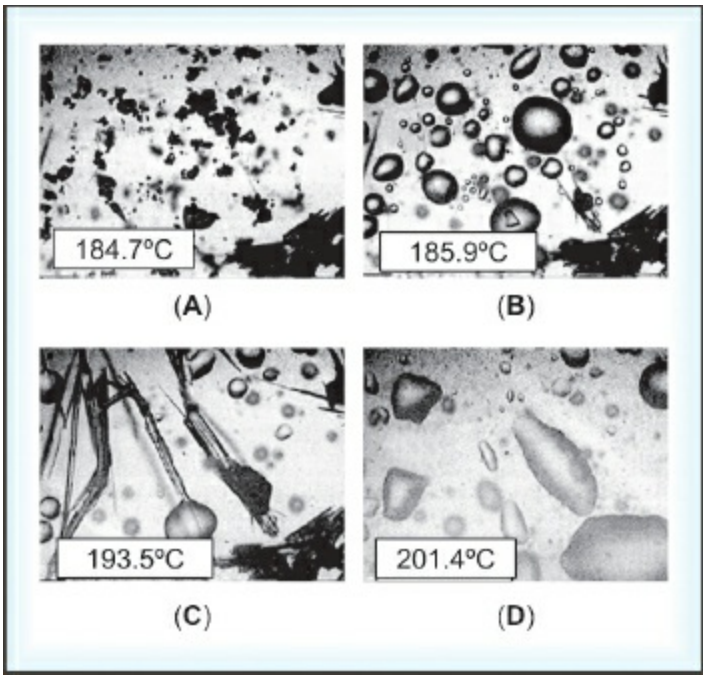



**Fig. 9.18:** Hot-stage microscopic photographs of carbamazepine

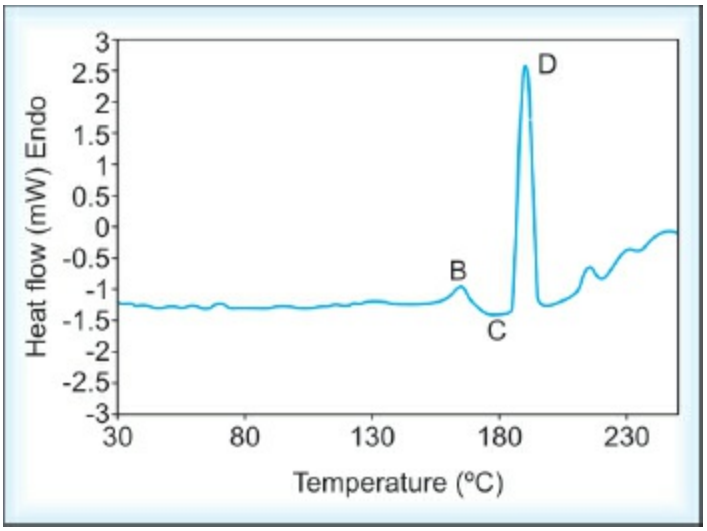



**Fig. 9.19:** DSC thermogram of carbamazepine

**Thermal Analysis:** Differential scanning calorimetry (DSC) and differential thermal analysis (DTA) measure the heat loss or gain resulting from physical or chemical changes within a sample as a function of temperature. Examples of endothermic (heat-absorbing) processes are fusion, boiling, sublimation, vapourization, desolvation, solid-solid transitions, and chemical degradation. Crystallization and degradation are usually exothermic processes. Quantitative measurements of these processes have many applications in preformulation studies including assessment of purity, polymorphism, solvation, degradation, and excipient compatibility.

For characterizing crystal forms, the heat of fusion, $\Delta H_f$, can be obtained from the area-under-the-DSC-curve for the melting endotherm. Similarly, the heat of transition from one polymorph to another may be calculated as shown by Guillory for several sulfonamides. A sharp symmetric melting endotherm can indicate relative purity, whereas broad, asymmetric curves suggest impurities or more than one thermal process. Heating rate affects the kinetics, and hence the apparent temperature of solid-solid transitions.

A variable with DSC experiments is the atmosphere in contact with the sample. Usually, a continual nitrogen purge is maintained within the heating chamber; however, the loss of a volatile counter ion such as ethanolamine or acetic acid during a polymorphic transition may produce misleading data unless the transition occurs within a closed system. In contrast, desolvation of a dihydrate species, as shown in Fig. 9.20, releases water vapour, which if unvented can generate degradation prior to the melting point of the anhydrous

form. During initial testing, a variety of atmospheres should be tried until the observed thermal process becomes fully understood.

Thermogravimetric analysis (TGA) measures changes in sample weight as a function of time (isothermal) or temperature. Desolvation and decomposition processes are frequently monitored by TGA. Comparing TGA and DSC data recorded under identical conditions can greatly aid in the interpretation of thermal processes. In Fig. 9.20, the dihydrate form of an acetate salt loses two moles of water via an endothermic transition between 70° and 90°C. The second endotherm at 155°C corresponds to the melting process, with the accompanying weight loss due to vapourization of acetic acid as well as to decomposition.

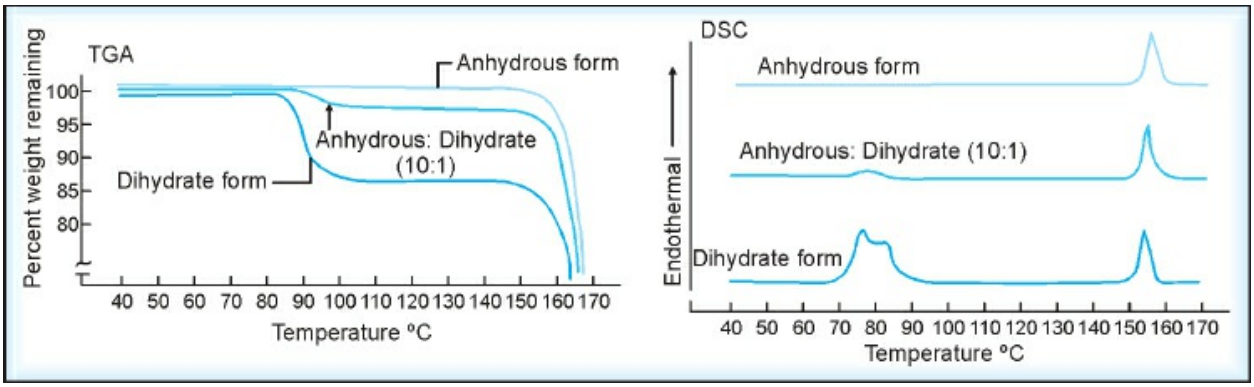



**Fig. 9.20:** Thermogravimetric (TGA) and differential scanning calorimetric (DSC) analysis for an acetate salt of an organic amine that has two crystalline forms, anhydrous and dihydrate. Anhydrous/dihydrate mixture was prepared by dry blending. Heating rate was 5°/min

TGA and DSC analysis can also be used to quantitate the presence of a solvated species within a bulk drug sample. For the above example, 10% of the dihydrate form was easily detected by both methods (Fig. 9.20). Both DSC and TGA are microtechniques and depend on thermal equilibration within the sample. Significant variables in these methods include sample homogeneity, sample size, particle size, heating rate, sample atmosphere, and sample preparation. Degradation during thermal analysis may provide misleading results, but may be detected by high-performance liquid chromatography (HPLC) analysis of samples heated under representative conditions for retention of drug or appearance of decay products (Fig. 9.21).

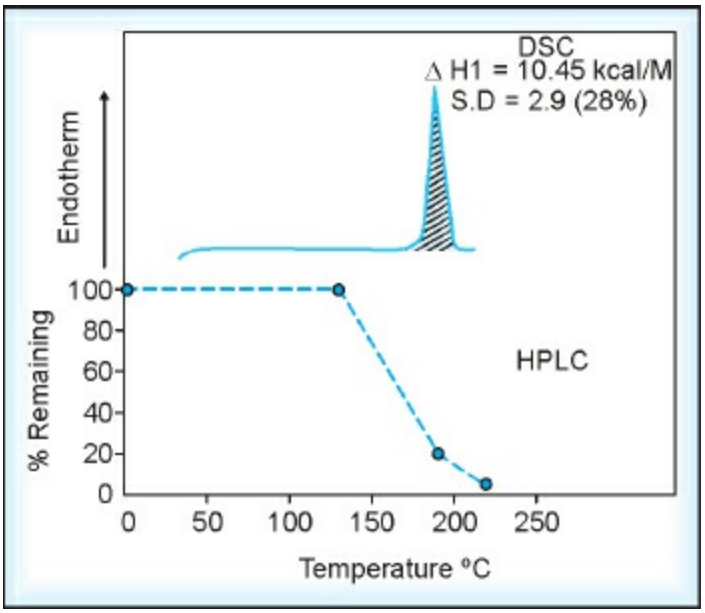



**Fig. 9.21:** Differential scanning calorimetric (DSC) analysis and HPLC stability analysis of an organic amine hydrochloride salt that undergoes decomposition upon melting

**X-Ray Diffraction:** An important technique for establishing the batch-to-batch reproducibility of a crystalline form is x-ray powder diffraction. Random orientation of a crystal lattice in a powder sample causes the x-rays to scatter in a reproducible pattern of peak intensities at distinct angles ($q$) relative to the incident beam. Each diffraction pattern is characteristic of a specific crystalline lattice for a given compound. An amorphous form does not produce a pattern. Mixtures of different crystalline forms can be analyzed using normalized intensities at specific angles, which are unique for each crystalline form. A typical standard curve is shown in Fig. 9.22 for polymorphic forms A and B of chloramphenicol palmitate.

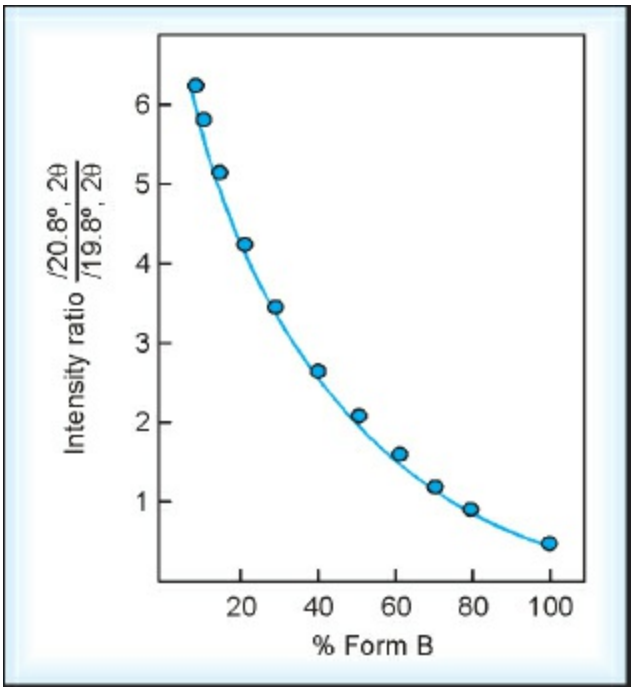



**Fig. 9.22:** X-ray intensity ratio as a function of composition of forms A and B of chloramphenicol palmitate

Single-crystal X-ray analysis provides a precise identification and description of a crystalline substance. Unit cell dimensions and angles conclusively establish the crystalline lattice system and provide specific differences between crystalline forms of a given compound. Other methods such as infrared spectroscopy, dilatometry, proton magnetic resonance (PMR), nuclear magnetic resonance spectroscopy (NMR), and scanning electron microscopy (SEM) have additional applications for studying polymorphism and solvation.

**Infrared (IR) Spectroscopy:** IR spectroscopy is also able to distinguish different polymorphic forms of a compound, since different arrangements of atoms in solid-state will lead to different molecular environments, and this leads to different stretching frequencies. The inclusion of solvent or water can be detected by broad −OH stretch associated with water. The use of IR spectroscopy to characterize three different polymorphic forms of carbamazepine has been exemplified in Table 9.11.

**Table 9.11:** IR bands of Form I, II and III of carbamazepine

| Carbamazepine form | N–H stretch 3490–3460 $cm^{-1}$ | C–O–R vibration 1700–1680 $cm^{-1}$ | C–H vibration 830–770 $cm^{-1}$ | –C = C–/–C = O vibration |
|---|---|---|---|---|
| Form I | 3489 $cm^{-1}$ | 1695 $cm^{-1}$ | 811, 800, 783 $cm^{-1}$ | Two bands |
| Form II | 3473 $cm^{-1}$ | 1688 $cm^{-1}$ | 815, 783, 770 $cm^{-1}$ | Single band |
| Form III | 3468 $cm^{-1}$ | 1680 $cm^{-1}$ | 810, 775 $cm^{-1}$ | Two bands |

## *Hygroscopicity*

Many drug substances, particularly water-soluble salt forms, have a tendency to adsorb atmospheric moisture. Adsorption and equilibrium moisture content can depend upon the atmospheric humidity, temperature, surface area, exposure, and the mechanism of moisture uptake, as described by Van Campen and co-workers. Deliquescent materials adsorb sufficient water to dissolve completely, as is observed with sodium chloride on a humid day. Other hygroscopic substances adsorb water because of hydrate formation or specific site adsorption. The European Pharmacopoeia Technical Guide has classified the degree of hygroscopicity into four classes based on the static method, after storage at 25°C for 24 h at 80% RH:

- *Slightly hygroscopic:* Increase in weight is ≥ 0.2 % w/w and < 2 % w/w.
- *Hygroscopic:* Increase in weight is ≥ 0.2 % w/w and <15 % w/w.
- *Very hygroscopic:* Increase in weight is ≥ 15% w/w.
- *Deliquescent:* Sufficient water is absorbed to form a solution.

With most hygroscopic materials, changes in moisture level can greatly influence many important parameters, such as chemical stability, flowability, and compactibility (Table 9.12).

**Table 9.12:** Effect of % relative humidity on appearance and flow property of a compound

| % RH | Moisture content; (w/w) | Appearance and flow properties |
|---|---|---|
| 0 | 0.31 | Free-flowing powder, passed easily through sieve |
| 11.3 | 0.24 | Free-flowing powder, passed easily through sieve |
| 22.5 | 0.27 | Less free-flowing powder |

| | | |
|---|---|---|
| 38.2 | 0.32 | Base of powder bed adhered to petri dish; however, material above this flowed |
| 57.6 | 0.34 | Less free-flowing |
| 75.3 | 0.62 | Material caked |
| Ambient | 0.25 | Base of powder adhered to petri dish |

To test for hygroscopicity, samples of bulk drug are placed in open containers with a thin powder bed to assure maximum atmospheric exposure. These samples are then exposed to a range of controlled relative humidity environments prepared with saturated aqueous salt solutions (Table 9.13). Moisture uptake should be monitored at time points representative of handling (0 to 24 h) and storage (0 to 12 weeks). Analytical methods for monitoring the moisture level (i.e. gravimetry, TGA, Karl Fischer titration, or gas chromatography) depend upon the desired precision and the amount of moisture adsorbed onto the drug sample.

Normalized (mg $H_2O$/g sample) or per-centage-of-weight-gain data from these hygroscopic studies are plotted against time to justify special handling procedures kinetically. A plot of normalized equilibrium versus relative humidity data may support the need for storage in a low-humidity environment or for special packaging with a desiccant. As these studies proceed, additional testing of powder flow, dissolution, or stability of “wet” bulk may be warranted to lend further support to the need for humidity controls.

**Table 9.13:** Relative humidity generated by various saturated salt solutions

| **Salt solution** | **% Relative humidity at 25°C** |
|---|---|
| Silica gel | 0 |
| Potassium acetate | 20 |
| Calcium chloride | 32 |
| Sodium bromide | 58 |
| Potassium bromide | 84 |

| | |
|---|---|
| Dipotassium hydrogen phosphate | 92 |
| Water | 100 |

## *Micromeritic Properties*

Bulk flow, formulation homogeneity, and surface area-controlled processes such as dissolution and chemical reactivity are directly affected by micromeritic properties of solids such as size, shape, and surface morphology. In general, each new drug candidate should be tested during preformulation with the smallest particle size as is practical to facilitate the preparation of homogeneous samples and maximize the drug's surface area for interactions. The micromeritic properties also have a significant impact on the processability and product quality of pharmaceutical dosage forms.

## *Particle Characterization*

The techniques most readily available for particle characterization include sieving, optical microscopy, electron microscopy, coulter counter and laser diffractometers. These techniques are described in detail in Chapter 2. Size characterization is simple for spherical particles, but not for irregular particles where the assigned size will depend on the method of characterization used. Table 9.14 lists particle size measurement methods and the corresponding size range.

**Table 9.14:** Particle sizing techniques and size range analyzed

| **Method** | **Size range (μm)** |
|---|---|
| Sieving (woven wire) | 20–125,000 |
| Sieving (electroformed) | 5–120 |
| Sieving (perforated plate) | 1,000–125,000 |
| Microscopy (optical) | 0.5–150 |
| Microscopy (electron) | 0.001–5 |
| Sedimentation (gravity) | 1–50 |
| Sedimentation (centrifugal) | 0.01–5 |
| Electrical zone sensing (e.g. coulter) | 1–200 |
| Laser-light scattering (Fraunhofer) | 1–1,000 |

| | |
|---|---|
| Laser light scattering (quasi-elastic) | 0.001–1 |

Kinetic processes involving drug in the solid state, such as dissolution and degradation, may be more directly related to the available surface area than to particle size. If drug particles have a shape that can be denned mathematically, then light microscopic size analysis or Coulter counter analysis with appropriate geometric equations may provide a reasonable estimation of the surface area. A more precise measurement of surface area is made by Brunauer, Emmett, and Teller (BET) nitrogen adsorption, in which a layer of nitrogen molecules is adsorbed to the sample surface at −196°C. Once surface adsorption has reached equilibrium, the sample is heated to room temperature, the nitrogen gas is desorbed, and its volume is measured and converted to the number of adsorbed molecules via the ideal gas law. Since each nitrogen molecule ($N_2$) occupies an area of $16A^2$, one may readily compute the surface area per gram for each preweighed sample. The BET isotherm for pharmaceutical powders is given by:

$$\frac{P}{V(P_0 - P)} = \frac{1}{CV_{mon}} + \left\{\frac{C-1}{CV_{mon}}\right\}\left\{\frac{P_{mon}}{P_0}\right\} \quad \text{... (18)}$$

where, $P$ is the partial pressure of the adsorbate, $V$ is the volume of gas adsorbed at pressure $p$, $V_{mon}$ is the volume of gas at monolayer coverage, $P_o$ is the saturation pressure and $c$ is related to the intercept.

Thus by plotting $P/V$ ($P_o - P$) versus $P/P_o$, a straight line of slope $C - 1/CV_{mon}$ and intercept $1/CV_{mon}$ will be obtained. The total surface area is thus:

$$S_t = \frac{V_{mon} N A_{cs}}{M} \quad \text{... (19)}$$

where, $N$ is the Arogadro's number, $A_{cs}$ is the cross-sectional area of the adsorbate and $M$ is the moleular weight of the adsorbate. It follows that the specific surface area is given by $S_t/m$, where $m$ is the mass of the sample.

While BET measurements are usually precise and quickly obtained with current commercial equipment, errors may arise from the use of impure gases and volatile surface impurities (e.g. hydrates).

Surface morphology may be observed by scanning electron microscopy

(SEM), which serves to confirm qualitatively a physical observation related to surface area. For example, bulk lots of drug recovered by different crystallization processes that have been used in an attempt to improve yield may result in surface morphologies that provide greater area for surface reactions such as degradation, dissolution, or hygroscopicity.

During preparation for SEM analysis, the sample is exposed to high vacuum during the gold coating process, which is needed to make the samples conductive, and concomitant removal of water or other solvents may result in a false picture of the surface morphology. Variable vacuum treatment of an identical sample prior to the gold coating step may confirm the effects of sample preparation on surface morphology. Most modern SEM instruments also provide energy dispersive X-ray spectroscopy analysis of surface metal ions, which may prove beneficial in deciphering an instability or incompatibility problem.

## *Density and Porosity*

Density and porosity are two important pharmaceutical properties that can be derived from the information on particle size distribution, particle shape, and surface area. Density can be defined as ratio of the mass of an object to its volume; therefore, the density of a solid is a reflection of the arrangement of molecules in a solid. Although the mass of a bulk powder sample can be determined with great accuracy, measurement of the volume is more complicated than it may first appear. The main problem arises in actually defining volumes of bulk powders, as may be seen from Fig. 9.23, in which three types of air spaces or voids can be distinguished:

1. Open intraparticulate voids—those within a single particle but open to the external environment.
2. Closed intraparticulate voids—those within a single particle but closed to the external environment.
3. Interparticulate voids—the air spaces between individual particles.

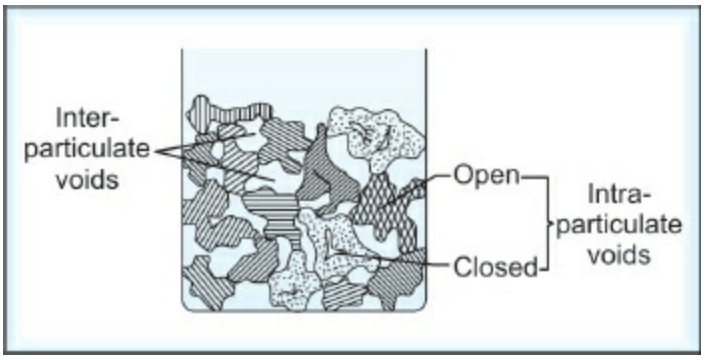



**Fig. 9.23:** Diagram of various intraparticulate and interparticulate air spaces in a bed of powder

Therefore, at least three interpretations of "powder volume" and corresponding "density" may be proposed:

1. The *true volume* ($\upsilon_t$)—the total volume of the solid particles, which excludes all spaces greater than molecular dimensions, and which has a characteristic value for each material. The density corresponding to true volume is termed as true density (sometimes called *theoretical density*) and defined as:

$$\frac{M}{v_t} = \rho_t \qquad \text{... (20)}$$

2. The *granular volume* (particle volume) ($\upsilon_g$)—the cumulative volume occupied by the particles, including all intraparticulate (but not interparticulate) voids. The density corresponding to granular volume is termed as granular density and defined as:

$$\frac{M}{v_g} = \rho_g \qquad \text{... (21)}$$

3. The *bulk volume* ($\upsilon_b$)—the total volume occupied by the entire powder mass under the particular packing achieved during the measurement. The density corresponding to granular volume is termed as granular density and defined as:

$$\frac{M}{v_b} = \rho_b \qquad \text{... (22)}$$

where, $M$ is the mass of the sample.

Comparing the density $p$ of a sample under specific test conditions with

the *true* of the material leads to the dimensionless quantity $p_r$, the *relative density*, where:

$$\rho_r = \frac{\rho}{\rho_t} \qquad \ldots (23)$$

During compressional processes, relative density increases to a maximum of unity when all air spaces have been eliminated.

## *Porosity*

The voids present in the powder mass may be more significant than the solid components in certain studies. For example, a fine capillary network of voids or pores has been shown to enhance the rate of liquid uptake by tablets, which in turn increases the rate of their disintegration. For this reason, a second dimensionless quantity, the ratio of the total volume of void spaces (vv) to the bulk volume of the material, is often selected to monitor the progress of compression. This ratio $\frac{V_v}{V_b}$ is referred to as the *porosity* ($E$) of the material:

$$V_v = V_b - V_t \qquad \ldots (24)$$

$$\text{Therefore, porosity, } E = \frac{V_b - V_t}{V_b} = 1 - \frac{V_t}{V_b}$$

Porosity is frequently expressed as a percentage:

$$E = 100 \times \left[1 - \frac{V_t}{V_b}\right] \qquad \ldots (25)$$

For example, a cylindric tablet of 10 mm diameter and 4 mm height weighed 480 mg and was made from material of true density 1.6 g $cm^{-3}$. The bulk volume $v_b$ is given by:

$$V_b = \pi \times \left(\frac{10}{10 \times 2}\right)^2 \times \frac{4}{10} cm^3$$

(volume of a cylinder is $\pi r^2 h$)

$$= (0.5)^2 \times 0.4 = 0.3142 \text{ cm}^3$$

The true volume of the solid is the true density divided by the mass, that is:

$$v_t = \frac{480}{1000} \div 1.6 = \frac{0.48}{1.6} = 0.3\ \text{cm}^3$$

Therefore, the porosity $E$ is:

$$E = 100 \times \left[1 - \frac{0.3}{0.3142}\right]$$
$$= 100(1 - 0.9548)$$
$$= 4.5\% \text{ (approximately)}$$

Several practical techniques are available for measuring the density of powder samples. Apart from X-ray diffraction methods, the nearest approach to true density is probably provided by a helium pycnometer. This works on the principle that within a sealed system containing helium (a nonadsorbing gas), the change in pressure caused by a finite change in volume of the system is a function of its total volume. A schematic diagram for a typical apparatus is shown in Fig. 9.24. During operation, the volume of the system is varied by means of the piston until a preset constant pressure is produced. This pressure is indicated by the sealed bellows pressure detector, which incorporates an integral electrical contact. The piston movement (U) necessary to achieve this pressure is read off from the scale. This value depends on the total volume of the system, which in turn is a function of

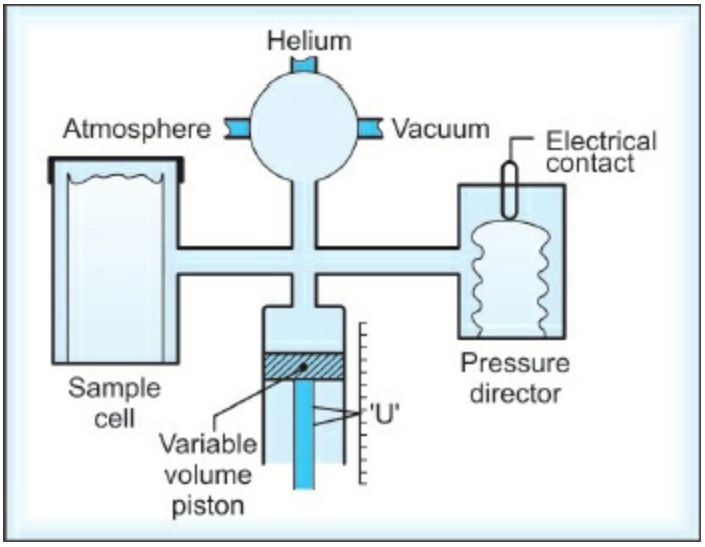



**Fig. 9.24:** Diagram of a helium pycnometer for determining the true volume of powders. The variable volume piston positions ($U_1$, $U_2$, and $U_s$) are read off from the scale and are used in equation (6) (*see* text)

the sample volume in the cell. The pycnometer is first calibrated using a sample of known volume $v_c$ (usually a stainless steel sphere). The operating equation for the instrument then becomes:

$$V_t = \frac{V_c}{U_1 - U_2} \times [U_1 - U_s] \qquad \ldots (26)$$

where, $U_1$, $U_2$ and $U_s$ are the variable volume scale readings for an empty cell, with standard volume, and with test powder, respectively.

Experimentally, the true density is also determined by suspending drug particles in solvents of various densities and in which the compound is insoluble. Wetting and pore penetration may be enhanced by the addition of a small quantity of surfactant to the solvent mixtures. After vigorous agitation, the samples are centrifuged briefly and then left to stand undisturbed until floatation or settling has reached equilibrium. The sample that remains suspended corresponds to the true density of the material. Liquid displacement by a powder pycnometer (specific gravity bottle method) can also be used, but unless special precautions are taken to ensure that no air remains in the sample, the results are prone to errors. For this reason, liquid displacement is probably best regarded as a measurement of granule density, especially if liquids that do not readily wet the powder are used, e.g. certain inert organic liquids, or mercury.

Apparent bulk density (g/ml) is determined by pouring presieved (40-mesh) bulk drug into a graduated cylinder via a large funnel and measuring the volume and weight “as is.” Tapped density is determined by placing a graduated cylinder containing a known mass of drug or formulation on a mechanical tapper apparatus, which is operated for a fixed number of taps (~1,000) until the powder bed volume has reached a minimum. Using the weight of drug in the cylinder and this minimum volume, the tapped density may be computed. Knowing the anticipated dose and tapped formulation density, one may use Fig. 9.25 to determine the appropriate size for a capsule formulation.

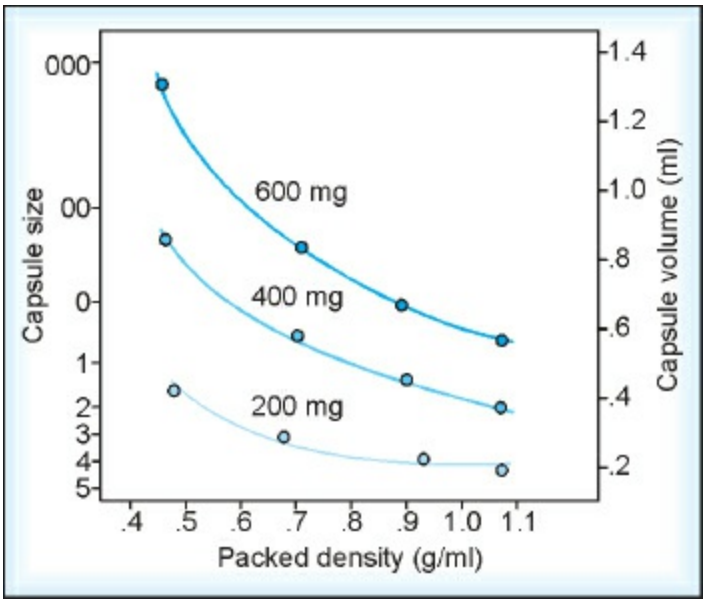



**Fig. 9.25:** Correlation between capsule size and packed density for different fill weights (200–600 mg)

## *Powder Flow Properties*

The flow properties of a material result from many forces. Solid particles attract one another, and forces acting between particles when they are in contact are predominately surface forces. There are many types of forces that can act between solid particles (.) frictional forces, (2) surface tension forces, (3) mechanical forces caused by interlocking of particles of irregular shape, (4) electrostatic forces and (5) cohesive or Van der Waals forces. All of these forces can affect flow properties of a solid. Pharmaceutical powders may be broadly classified as free-flowing or cohesive (non-free-flowing). Most flow properties are significantly affected by changes in particle size, density, shape, electrostatic charge, and adsorbed moisture, which may arise from processing or formulation. As a result, a free-flowing drug candidate may become cohesive during development, thus necessitating an entirely new formulation strategy. Preformulation powder flow investigations should quantitatively assess the pharmaceutical consequences of each process improvement and provide direction for the formulation development project team. This direction may consist of a formulation recommendation such as granulation or densification via slugging, the need for special auger feed equipment, or a test system for evaluating the improvements in flow brought about by formulation. This subject becomes paramount when attempting to

develop a commercial solid dosage form containing a large percentage of cohesive material.

A simple indication of the ease with which a material can be induced to flow is given by application of a *compressibility index* and Hausner ratio given by the equation:

$$\text{Carr's compressibility index} = \frac{(\text{Tap density} - \text{Bulk density})}{\text{Tap density}} \times 100\% \quad \ldots (27)$$

$$\text{Hausner's ratio} = \frac{\text{Tap density}}{\text{Bulk density}} \quad \ldots (28)$$

Scale of flowability for Carr's compressibility index and Hausner's ratio is given in Table 9.15, whereas, Table 9.16 lists compressibility data and flowability characterization for several pharmaceutical excipients.

Alternatively, free-flowing powders may be characterized by a simple flow rate apparatus consisting of a grounded metal tube from which drug flows through an orifice onto an electronic balance, which is connected to a strip chart recorder. Several flow rate (g/s) determinations at each of a variety of orifice sizes (⅛ to ½ inches) should be made. In general, the greater the standard deviation between multiple flow rate measurements, the greater is the weight variation in products produced from that powder. When several lots of a drug candidate are tested under dissimilar conditions, Eq. (1), proposed by Jones and Pilpel, may be used to show the dependence of flow rate ($W$) on true particle density ($\rho$), acceleration due to gravity ($g$), and orifice diameter ($D_0$). Both ($A$) and ($n$) are constants that are dependent upon the material and its particle size.

$$D_0 = A\left(\frac{4W}{60\pi\rho\sqrt{g}}\right)^{\frac{1}{n}} \quad \ldots (29)$$

**Table 9.15:** Scale of flowability for Carr's compressibility index and Hausner's ratio

| Carr's index | Hausner's ratio | Flowability |
|---|---|---|
| 5–15 | 1.05 – 1.18 | Excellent |
| 12–16 | 1.14 – 1.20 | Good |

| | | |
|---|---|---|
| 18–21 | 1.22 – 1.26 | Fair-passable |
| 23–35 | 1.30 – 1.54 | Poor |
| 33–38 | 1.50 – 1.61 | Very poor |
| >40 | >1.67 | Very, very poor |

**Table 9.16:** Compressibility and flowability of pharmaceutical excipients

| Material | % Compressibility | Flowability |
|---|---|---|
| Celutab | 11 | Excellent |
| Emcompress | 15 | Excellent |
| Star X-1500 | 19 | Fair-passable |
| Lactose monohydrate | 19 | Fair-passable |
| Maize starch | 26–27 | Poor |
| Dicalcium phosphate, dihydrate (coarse) | 27 | Poor |
| Magnesium stearate | 31 | Poor |
| Titanium dioxide | 34 | Very poor |
| Dicalcium phosphate, dihydrate (fine) | 41 | Very, very poor |
| Talc | 49 | Very, very poor |

## *Angles of Repose*

A simple practical technique for measuring resistance to particle movement is a quantity called *the angle of repose* of a powder. This is the angle, $\theta$ as defined by the equation:

$$\tan\theta = \frac{2h}{D} \qquad \ldots (30)$$

It is the maximum angle that can be obtained between the freestanding surface of a powder heap and the horizontal plane, as shown in Fig. 9.26A. Such measurements give at least a qualitative assessment of the internal cohesive and frictional effects under low levels of external loading, as might apply in powder mixing, or in tablet die or capsule shell filling operations.

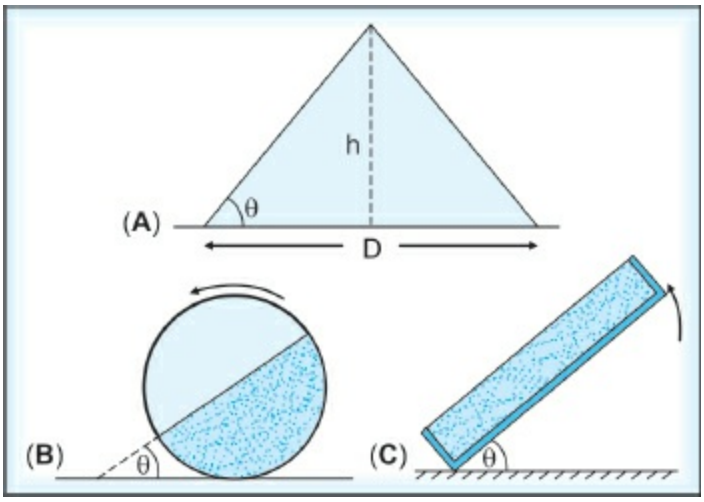



**Fig. 9.26:** Measurement of dynamic angles of repose, $\theta$, as defined from the dimensions of a conical bed of the powder (A), where tan$\theta$= twice the powder bed height (h)/powder bed diameter (D). (B and C) represent the rotating cylinder and tilted box methods of measurement, respectively

## *Static Angle of Repose*

The **fixed funnel method** employs a funnel that is secured with its tip at a given height, H, above graph paper that is placed on a flat horizontal surface. Powder or granulation is carefully poured through the funnel until the apex of the conical pile just touches the tip of the funnel. The diameter of the base of the conical pile is then determined to calculate the angle of repose.

The **fixed cone method** establishes the diameter of the cone base, D, by using a circular dish with sharp edges. Powder is poured onto the center of the dish from a funnel that can be raised vertically until a maximum cone height, H, is obtained. The repose angle is calculated as before. Angle of repose methods, which result in a so-called *dynamic* angle, are preferred, since they most closely mimic the manufacturing situation, in which the powder is in motion.

## *Kinetic or Dynamic Angle of Repose*

## *Rotating Cylinder Method*

A typical dynamic test involves a hollow cylinder half-filled with the test powder, with one end sealed by a transparent plate. The cylinder is rotated about its horizontal axis (Fig. 9.26B), until the powder surface cascades. The

curved wall is lined with sandpaper to prevent preferential slip at this surface.

### *Tilting Box Method*

A sandpaper-lined rectangular box is filled with the powder and carefully tilted until the contents begin to slide, as shown in Fig. 9.26C. The maximum angle that the plane of powder makes with the horizontal surface on rotation is taken as the angle of repose.Values of Φ are rarely less than 20°, and values of up to 40° indicate reasonable flow potential. Above 50°, however, the powder flows only with great difficulty, if at all. Values for angles of repose ≤30° usually indicate a free-flowing material and angles ≥40° suggest a poorly flowing material. As mentioned previously, flow of coarse particles is also related to packing densities and mechanical arrangements of particles. For this reason, a good auxiliary test to run in conjunction with the angle of repose test is the compressibility test, discussed previously (Fig. 9.27). From the angle of repose and compressibility values, a reasonable indication of a material's inherent flow properties should be possible.

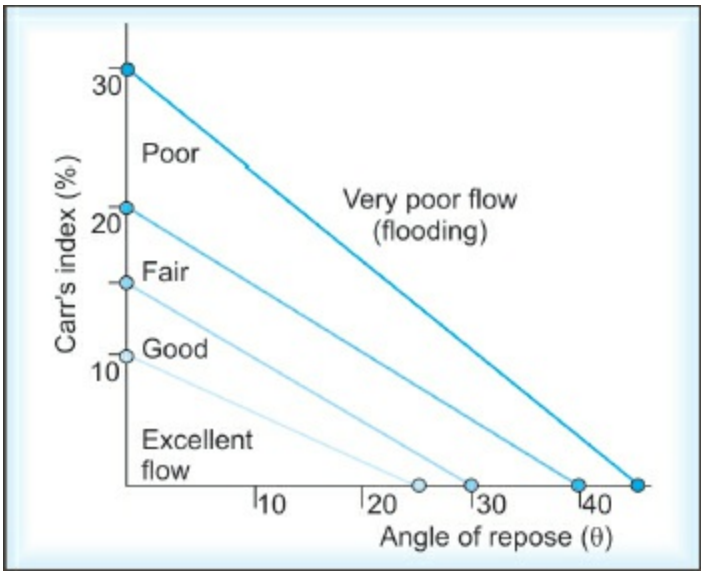



**Fig. 9.27:** Relationship between angle of repose, Carr's index and flow characteristics of powder

## Stability Analysis

Preformulation stability studies are usually the first quantitative assessment of chemical stability of a new drug. These studies include both solution and solid-state experiments under conditions typical for the handling, formulation, storage, and administration of a drug candidate.

Knowledge about the stability of a candidate drug in the solid and liquid state is extremely important in drug development. In the longer term, the stability of the formulation will dictate the shelf-life of the marketed product, however, to achieve this, careful preformulation is required. The major objectives of the preformulation team are, therefore, to identify conditions to which the compound is sensitive and interpret degradation profiles under these conditions. The major routes of drug degradation in solution are via hydrolysis, oxidation or photochemical reactions.

### *Solution Stability*

The decomposition of a drug occurs through several pathways, i.e., hydrolysis, oxidation, photolysis and racemization.

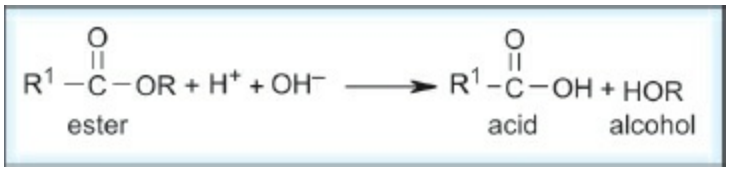
## Hydrolysis

Many pharamaceuticals contain ester or amide functional groups, which undergo hydrolysis in solution. Examples of drugs that tend to degrade by hydrolytic cleavage of an ester or amide linkage are anesthetics, antibiotics, vitamins, and barbiturates.

### *Ester Hydrolysis*

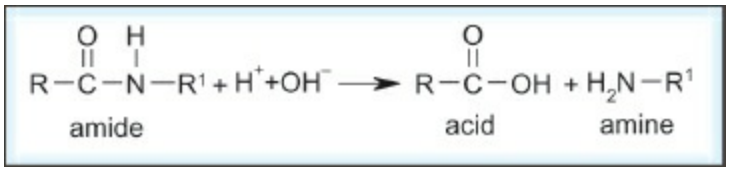
The hydrolysis of an ester into a mixture of an acid and alcohol essentially involves the rupture of a covalent linkage between a carbon atom and an oxygen atom. Although some of these hydrolyses can be effected in pure water, in the majority of cases, the presence of catalyst is needed to promote the reaction. These catalysts are invariably substances of polar nature, such as mineral acids, alkalies, or certain enzymes, all of which are capable of supplying hydrogen or hydroxyl ions to the reaction mixture.

In practice, the general scheme employed to denote ester hydrolysis is as follows:




### *Amide Hydrolysis*

Pharmaceutical compounds containing an amide group can undergo hydrolysis in a manner similar to that of an ester-type compound. Instead of the acid and alcohol that form as a result of ester hydrolysis, hydrolytic cleavage of an amide results in the formation of an acid and an amine.




Because of the relatively greater stability of amides as compared with structurally-similar esters, there is considerably less information in the literature on quantitative chemical kinetic studies pertaining to the hydrolytic

stability of such compounds. Pharmaceuticals such as niacinamide, phenethicillin, barbiturates, and chloramphenicol degrade by amide hydrolysis.

## Oxidation

The oxidative decomposition of pharmaceutical compounds is responsible for the instability of a considerable number of pharmaceutical preparations. For example, steroids, vitamins, antibiotics, and epinephrine undergo oxidative degradation. The most common form of oxidative decomposition occurring in pharmaceutical preparations is autoxidation, which involves a free radical chain process. In general, autoxidation may be defined as the reaction of any material with molecular oxygen. Free radicals are produced by reactions involving homolytic bond fission of a covalent bond, so that each atom or group involved retains one of the electrons of the original covalent bond. To test whether a compound is sensitive to oxygen, simply bubble air through the solution, or add hydrogen peroxide, and assess the amount of degradation that takes place.

## Photolysis

Consideration of the decomposition of pharmaceutical compounds resulting from the absorption of radiant energy in the form of light has become more important in recent years because of the complex chemical structure of many new drugs. Degradative reactions, such as oxidation-reduction, ring rearrangement, or modification and polymerization, can be brought about by exposure to light at particular wavelengths. According to the equation, $E = 2.859 \times 10^5/\lambda$ kcal per mole, the shorter the wavelength ($\lambda$) of light, the more energy ($E$) is absorbed per mole. Consequently, the radiations absorbed from the ultraviolet and violet portions of the light spectrum are more active in initiating chemical reactions than those absorbed from the other longer wavelength portions of the spectrum.

## Racemization

In such a reaction, an optically-active substance loses its optical activity without changing its chemical composition. This reaction can be important to the stability of pharmaceutical formulations, since the biologic effect of the dextro-form can be considerably less than that of the levo form. For example, levo-adrenaline is 15 to 20 times more active than dextro-adrenaline. Solutions of levo-adrenaline form a racemic mixture of equal parts of levo- and dextro-adrenaline with a pharmacologic activity just over half that of the pure levo-compound.

The primary objective of this phase of preformulation research is the identification of conditions necessary to form a stable solution. These studies should include the effects of pH, ionic strength, cosolvent, light, temperature, and oxygen. Solution stability investigations usually commence with probing experiments to confirm decay at the extremes of *pH* and temperature (e.g. 0.1N HCl, water, and 0.l N NaOH all at 90°C). These intentionally-degraded samples may be used to confirm assay specificity as well as to provide estimates for maximum rates of degradation. This initial experiment should be followed by the generation of a complete pH-rate profile to identify the pH of maximum stability. Aqueous buffers are used to produce solutions over a wide range of pH values with constant levels of drug, cosolvent, and ionic strength.

Since most solution pharmaceuticals are intended for parenteral routes of administration, this initial pH-rate study should be conducted at a constant ionic strength that is compatible with physiologic media. The ionic strength ($\mu$) of an isotonic 0.9% sodium chloride solution is 0.15, and several compendia contain formulae for isotonic buffer solutions. Ionic strength for any new buffer solution may be calculated from the following equation:

$$\mu = \frac{1}{2}\sum m_i Z_i^2 \qquad \text{... (31)}$$

where, $m_i$ (is the molar concentration of the ion having a valency, $Z_i$. Note that all ionic species (even the drug molecules) in a buffer solution must be considered in computing ionic strength.

Once the stability solutions are prepared, aliquots are placed in flint glass ampules, flame-sealed to prevent evaporation, and stored at constant

temperatures not exceeding the boiling point of the most volatile cosolvent or its azeotrope. Some of the ampules may be stored at a wide range of temperatures to provide data for calculating activation energies.

Some of these solution samples should be subjected to a light stability test, which includes protective packaging in amber and yellow-green glass containers. Control samples for this light test may be stored in cardboard packages or wrapped in aluminum foil.

Given that the potential for oxidation is initially unknown, some of the solution samples should also be subjected to further testing (1) with an excessive headspace of oxygen, (2) with a headspace of an inert gas such as helium or nitrogen, (3) with an inorganic antioxidant such as sodium metabisulfite, and (4) with an organic antioxidant such as butylated hydroxytoluene (BHT). Headspace composition can be controlled if the samples are stored in vials for injection that are capped with Teflon-coated rubber stoppers. After penetrating the stoppers with needles, the headspace is flooded with the desired atmosphere, and the resulting needle holes are sealed with wax to prevent degassing.

To generate a pH-rate profile, stability data generated at each pH and temperature condition are analyzed kinetically to yield the apparent decay rate constants. All of the rate constants at a single temperature are then plotted as a function of pH as shown in Fig. 9.28. The minimum in this curve is the pH of maximum stability. Often, this plot, as it approaches its limits, provides insight into the molecular involvement of hydrogen or hydroxide ions in the decay mechanism.

An Arrhenius plot is constructed by plotting the logarithm of the apparent decay rate constant versus the reciprocal of the absolute temperature at which each particular buffer solution was stored during the stability test. To justify extrapolation to “use” conditions, stability storage temperatures should be selected that incrementally ($\Delta$ t ~ 10° C) approach the anticipated “use” temperature. If this relationship is linear, one may assume a constant decay mechanism over this temperature range and calculate an activation energy ($E_a$) from the slope ($-E_a/R$) of the line as described following equation:

$$\ln k = \frac{-E_a}{R}\left(\frac{1}{T}\right) + C \qquad \text{... (32)}$$

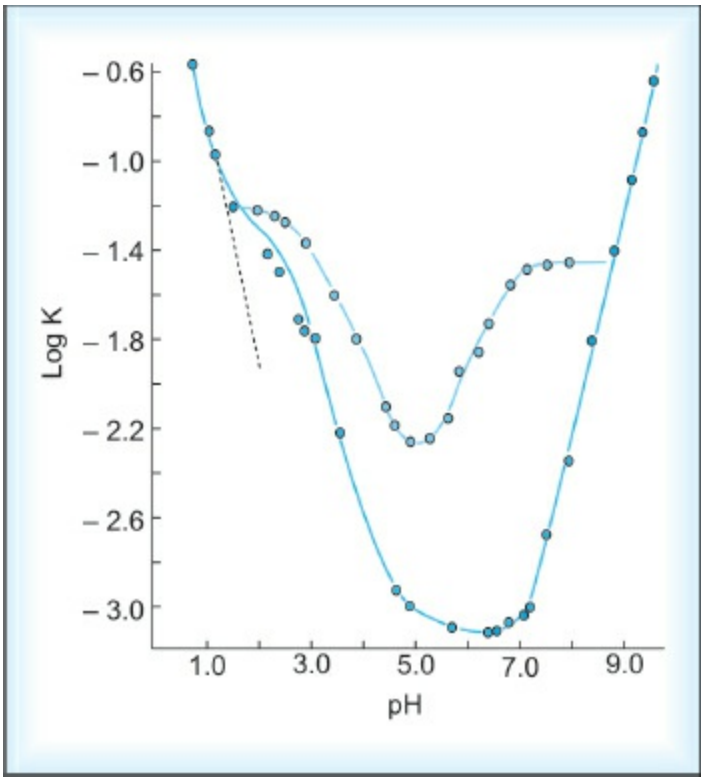



**Fig. 9.28:** The pH-rate profiles for ampicillin degradation in solution at 35°C and constant ionic strength ($\mu$ = 0.5). Dotted line is the apparent rate profile in the presence of buffer, while the solid line is the theoretic rate profile at zero buffer concentration

where, *C* is a constant of integration and *R* is the gas constant.

A broken or non-linear Arrhenius plot suggests a change in the rate-limiting step of the reaction or a change in decay mechanism, thus making extrapolation unreliable. In a solution-state oxidation reaction, for example, the apparent decay rate constant decreases with elevation of temperature because the solubility of oxygen in water decreases. At elevated temperatures, excipients or buffers may also degrade to give products that are incompatible with the drug under study. Often, inspection of the HPLC chromatograms for decay products confirms a change in the decay mechanism.

Shelf-life ($t_{10\%}$) for a drug at "use" conditions may be calculated from the appropriate kinetic equation, and the decay rate constant obtained from the Arrhenius plot. For a first-order decay process, shelf-life is computed from:

$$t_{10\%} = \frac{-\ln 0.90}{k_1} = \frac{0.105}{k_1} \qquad \ldots (33)$$

where, $t_{10\%}$ is the time for 10% decay to occur with apparent first-order decay constant $k_1$. Frequently, it is useful to present the pH-rate profile as a plot of pH versus $t_{10\%}$ shelf-life data.

Results of these initial solution stability studies dictate the subsequent course of action. If the compound is sufficiently stable, liquid formulation development may commence at once. If the compound is unstable, then further investigations may be necessary.

## *Solid-state Stability*

The primary objectives of this investigation are identification of stable storage conditions for drug in the solid state and identification of compatible excipients for a formulation. Contrary to the earlier solution stability profile, these solid state studies may be severely affected by changes in purity and crystallinity, which often result from process improvements. Repetitive testing of the initial bulk lot in parallel with newer bulk lots should be performed, and adequate material should be set aside for these studies.

In general, solid-state reactions are much slower and more difficult to interpret than solution-state reactions, owing to a reduced number of molecular contacts between drug and excipient molecules and to the occurrence of multiple-phase reactions. A kinetic analysis of slow solid-state degradation based on retention of intact drug may fail to quantitate clearly the compound's shelf-life, as assay variation may equal or exceed the limited apparent degradation, particularly at the low temperatures that are critical to establishing a room-temperature shelf-life. Usually, this situation may be corrected on analysis of the appearance of decay product(s), which may total only 1 to 5% of the sample. Additional analytic data from such studies as TLC, fluorescence, or UV/Visible spectroscopy may be required to determine precisely the kinetics of decay product(s) appearance, and to establish a room-temperature shelf-life for the drug candidate.

To study the many possible solid-state reactions, one may need more than a specific assay for the intact compound. Polymorphic changes, for example, are usually detected by or quantitative infrared analysis (IR). In the case of surface discolouration due to oxidation or reaction with excipients, surface

reflectance measurements on tristimulus or diffuse reflectance equipment may be more sensitive than HPLC assay. In any event, additional samples are required in the solid-state stability study to accommodate these additional tests.

To determine the solid-state stability profile of a new compound, weighed samples are placed in open screw-cap vials and are exposed directly to a variety of temperatures, humidities, and light intensities for up to 12 weeks (Fig. 9.29). Samples usually consist of three 5-to 10 mg weighed samples at each data point for HPLC analysis and approximately 10 to 50 mg of sample for polymorph evaluation by DSC and IR (~2 mg in KBr and ~20 mg in Nujol). To test for surface oxidation, samples are stored in large (25 ml) vials for injection, capped with a Teflon-lined rubber stopper and the headspace flooded with dry oxygen. To confirm that the decay observed is due solely to oxygen rather than to reduced humidity, a second set of vials should be tested in which the atmosphere is flooded with dry nitrogen. After a fixed exposure time, these samples are removed and analyzed by multiple methods to check for chemical stability, polymorphic changes, and discolouration.

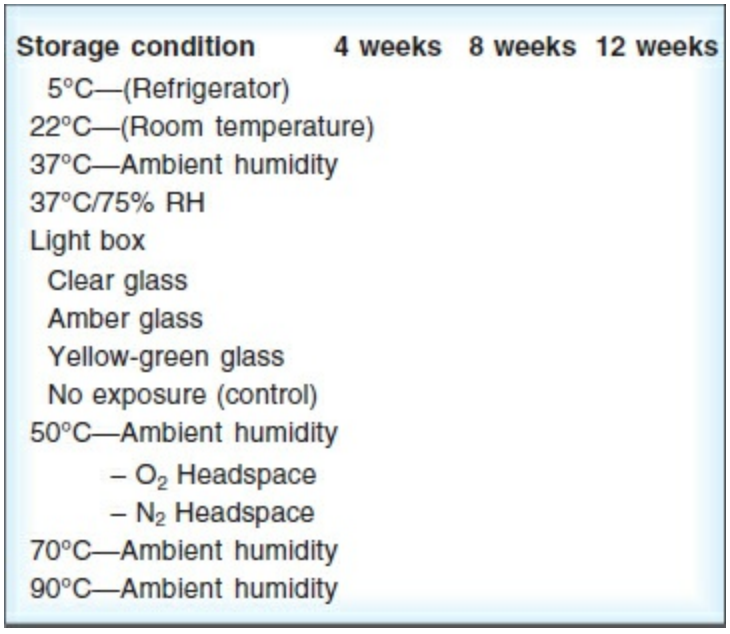



**Fig. 9.29:** Sample scheme for determining the bulk stability profile for a new drug candidate. At each point, bulk drug samples should consist of 3 × 10 mg for HPLC analysis and a 50 mg sample for polymorph analysis by DSC or IR

Once the results of this initial screen are tabulated, the decay process may

be analyzed by either zero-order or first-order kinetics, particularly if the amount of decay is less than 15 to 20%. The same kinetic order should be used to analyze the data at each temperature, if possible. Samples exposed to oxygen, light, and humidity may suggest the need for a follow up stability test at three or more levels of a given parameter for full quantitation of its involvement.

In the event that humidity is not a factor in drug stability, an Arrhenius plot may be constructed, and if linear, it may be extrapolated to "use" conditions for predicting shelf-life. If humidity directly affects drug stability, the concentration of water in the atmosphere may be determined from the relative humidity and temperature by using psychrometric charts. Stability data obtained at various humidities may be linearized w.r.t. moisture using the following apparent decay rate constant:

$$k_H = [gpl].k_0 \quad \text{... (34)}$$

where, [*gpl*] is the concentration of water in the atmosphere in units of grams of water per liter of dry air, and $k_0$ is the decay rate constant at zero relative humidity. For example, a 75% relative humidity atmosphere at 37°C is equivalent to 0.0405 grams of water per liter (*gpl*) of dry air. When the effect of moisture on chemical stability is examined in detail, a comparison to solution-state stability and hygroscopicity data may suggest an aqueous reaction occurring in the drug-saturated water layer on the crystal surface.

Another useful relationship for analyzing solid-state stability data assumes that a compound must partially liquefy prior to decomposition. Given that the mole fraction of the solid that has liquefied ($F_m$) is directly proportional to its decay rate, then:

$$\ln k_{app} \alpha \ln F_m = \frac{-\Delta H_{fus}}{R}\left[\frac{1}{T} - \frac{1}{T_m}\right] \quad \text{... (35)}$$

where, $\Delta H_{fus}$ is the molar heat of fusion, $T_m$ is the absolute melting point (°kelvin), $T$ is the absolute temperature of the stability study, and $R$ is the gas constant.

Once the bulk drug stability has been determined, compatibility with excipients commonly used to produce solid dosage forms must be established.

## *Drug-excipient Compatibility*

The successful formulation of a stable and effective dosage form depends not only on the active pharmaceutical ingredient but also on the careful selection of excipients. The proper selection of excipients is vital in the design of a quality drug product. Selection of the excipients and their concentration in a formulation is based not only on their functionality, but also on the compatibility between the drug and excipients. An incompatibility may result in changes in physical, chemical, microbiological or therapeutic properties of the dosage form.

Drug-excipient compatibility studies are conducted mainly to predict the potential incompatibility, and to provide justification for the selection of excipients in the formulation as required in regulatory filings. Knowledge gained from drug-excipient compatibility studies is important in the drug development process, is used to select the dosage form components, delineate stability profile of the drug, identify degradation products, and to understand the mechanisms of reactions.

Drug-excipient studies are usually conducted after gaining an understanding of solution-and solid-state stability, but before the formulation development activities.

An incompatibility in the dosage form can result in any of the following changes:

- Changes in organoleptic properties
- Changes in dissolution performance
- Physical form conversion
- An decrease in potency
- An increase in degradation products.

Several examples of incompatibilities in pharmaceutical excipients are known in the literature and can be used as a guiding reference (Table 9.17). Some examples of reactions of pharmaceutical excipients or their impurities with drugs are shown in Fig. 9.30.

**Table 9.17:** Incompatibilities due to pharmaceutical excipients

| Excipient | Incompatibility |
|---|---|

| | |
|---|---|
| Lactose | Maillard reactions, Claissen-Schmidt condensation reaction of its impurity 5-hydroxylmethyl-2-furfuraldehyde, and catalysis. |
| Microcrystalline cellulose | Maillard reaction, water sorption resulting in increased hydrolysis, adsorption of basic drugs, and non-specific incompatibilities due to hydrogen bonding capability. |
| Povidone and crospovidone | Oxidation attributable to peroxides, nucleophilic addition to amino acids and peptides, and hydrolysis of sensitive drugs due to moisture. |
| Hydroxypropyl cellulose | Oxidation of sensitive drugs due to residual peroxides. |
| Croscarmellose sodium | Weakly basic drugs can compete with the sodium counter-ion thus getting adsorbed, drug salt-form conversion. |
| Sodium starch glycolate | Adsorption of weakly-basic drugs and their salts due to electrostatic interactions, residual monochloroacetate may undergo nucleophilic reactions. |
| Starch | Terminal aldehydes in starch have been known to react with the hydrazine moiety, moisture-mediated reactions, drug adsorbtion, and may react with formaldehyde resulting in reduced functionality as a disintegrant. |
| Colloidal silicon dioxide | May act as a Lewis acid under anhydrous conditions, and may adsorb drugs. |
| Magnesium stearate | Exists in hydration states: mono, di and trihydrates, MgO impurity is known to react with ibuprofen, provides a basic pH environment and accelerate hydrolytic degradation, and magnesium metal may also cause chelation-induced degradation. |

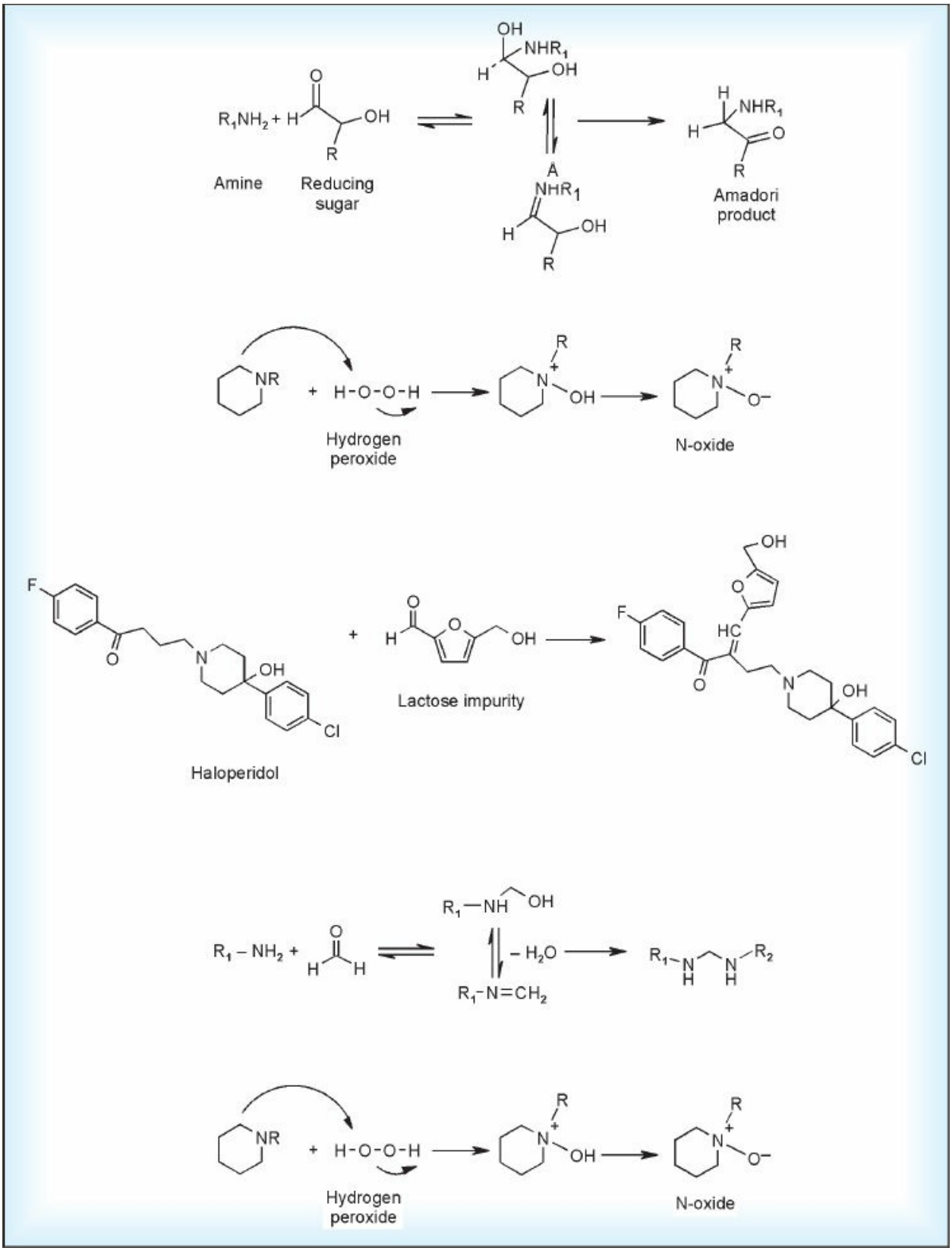



**Fig. 9.30:** Reactions of pharmaceutical excipients or their impurities with

drugs

## Method

A binary mixture of drug with the excipient being investigated is intimately mixed, the ratio of drug to excipient often being 1:1; however, other rates may also be investigated. Powder samples are then dispensed into clear, neutral glass ampoules and half of the ampoules are sealed without further treatment. To the other, 5% distilled water is added prior to sealing. Alternatively, the drug in suspension with excipients may be investigated. The ampoules are then stored at a suitable temperature and analyzed at various time points. The storage conditions used to examine compatibility can vary widely in terms of temperature and humidity, but a temperature of 50°C for storage of compatibility samples is considered appropriate. Some compounds may require higher temperatures to make reactions proceed at a rate that can be measured over a convenient time period. Figure 9.31 shows a representative of multiple drug-excipient compatibility testing protocol.

Compatibility protocol

| | Experiment | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 170 |
| Drug substance | 200 | 25 | 25 | 25 | 25 | 25 | 25 | 25 | 25 | 25 | 25 | 25 | 25 | 25 | 25 | 25 | 25 |
| Lactose | | 175 | | | | 170 | | | | 170 | | | 170 | | | | |
| Mannitol | | | 175 | | | | 170 | | | | 170 | | | 170 | | | |
| Microcrystalline cellulose | | | | 175 | | | | 170 | | | | 170 | | | 170 | | |
| Dibasic calcium phosphate dihydrate | | | | | 175 | | | | 170 | | | | 170 | | | 170 | |
| Magnesium stearate | | | | | | 5 | 5 | 5 | 5 | | | | | | | | |
| Sodium stearyl fumarate | | | | | | | | | | 5 | 5 | 5 | 5 | | | | |
| Stearic acid | | | | | | | | | | | | | | 5 | 5 | 5 | 5 |

**Fig. 9.31:** Protocol of multiple drug-excipient compatibility study

Methods of analysis also vary widely, ranging from thermal techniques (DSC), to chromatographic techniques (TLC, HPLC), to microcalorimetry. DSC has been used extensively for compatibility studies. Although only milligram quantities of drug are needed for a DSC experiment, the interpretation of the thermograms may be difficult, and conclusions may be misleading on the basis of DSC experiments alone. A scheme for interpreting drug-excipient compatibility data is shown in Fig. 9.32. Depending on the number of excipients to be investigated, compatibility tests can be speeded up by using factorial or reduced factorial design experiments.

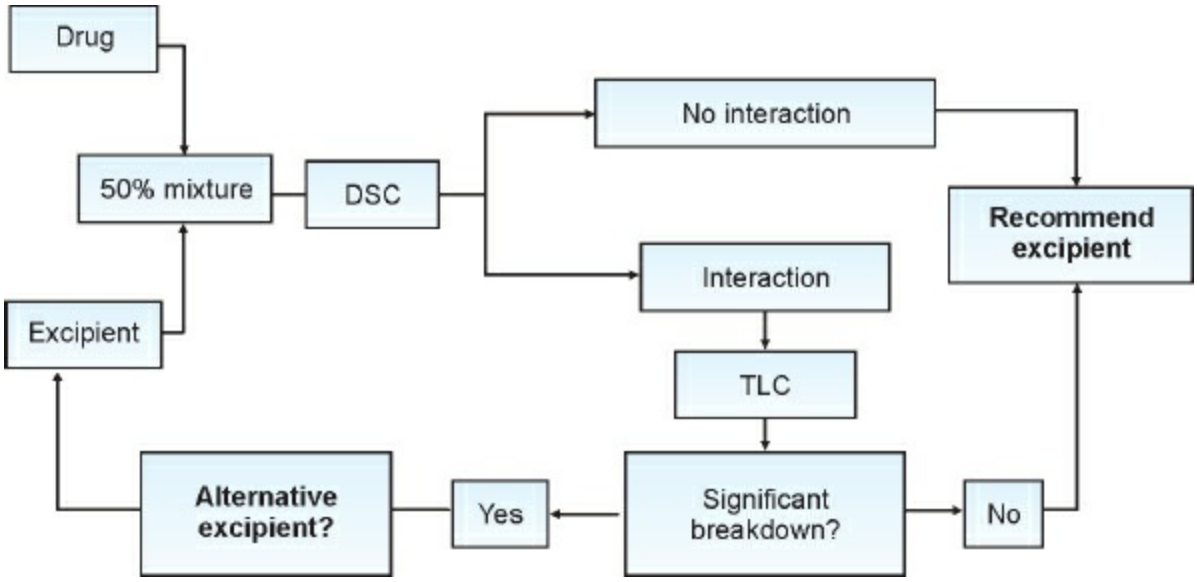



**Fig. 9.32:** Scheme to identify chemically compatible excipients

## Liquid Compatibilities

In general, such studies are performed by placing the drug in a solution of the additive. Usually, both amber and flint vials are used and in many cases autoclaved conditions are included. In case of emulsions, the preformulation studies include measuring the critical micelle concentration of the formulations. For preparations for oral use, compatibility testing with ethanol, glycerine, syrup, sucrose, buffers and preservatives are carried out to have an idea about the activation energy of the predominant reaction in solution.

## Formulation Recommendation

Upon completion of the preformulation evaluation of a new drug candidate, it is recommended that a comprehensive report be prepared highlighting the pharmaceutical problems associated with this molecule. This report should conclude with recommendations for developing phase I formulations. These reports are extremely important in preparing regulatory documents, and aid in developing subsequent drug candidates.

# 10:
# Biopharmaceutics and Pharmacokinetics

Biopharmaceutics is the study of the interrelationship of the physicochemical, physiologic and pharmaceutical properties of the active pharmaceutical ingredient (API) and the dosage form (in which the drug is fabricated) based on the biological performance of the drug (Table 10.1). Whereas physicochemical properties of the drug (and excipients) dictate the rate of drug release from the dosage form and the subsequent transport across biologic membranes, physiologic and biochemical realities determine its fate in the body.

**Table 10.1:** Biopharmaceutics consideration in design of drug product

| | | |
|---|---|---|
| **Active pharmaceutical ingredient (API)** | Stability Solubility pH and pKa Crystalline form (polymorph) Excipient interaction and compatibility | Impurities Salt-form Particle size Complexation |
| **Drug product** | Type of drug product (capsule, tablet, solution, etc.) Immediate or modified release Dosage strength Bioavailability | Stability Excipients Manufacturing variables |
| **Physiologic factors** | Route of administration Permeation of drug across cell membranes Binding to macromolecules | Blood flow Surface area Biotransformation Pharmacokinetics |
| **Pharmacodynamic and pharmacokinetic considerations** | Bioavailability Therapeutic objective Adverse reaction | Dose Toxic effects |

| | | |
|---|---|---|
| **Manufacturing consideration** | Production methodology and technology Quality control/quality assurance Specification of raw material | Cost Stability testing |
| **Patient consideration** | Compliance, labeling, and product acceptance | Cost |

As shown in Table 10.1, biopharmaceutics allows for rational design of drug products based on the physicochemical characteristics of the active drug substance, the desired drug product, and considerations of the anatomy and physiology of the human body. The optimal delivery of the active moiety to the site of action depends on an understanding of specific interactions between the formulation variables and the biologic variables. In drug product development, it is invariably an iterative process, whereby the pharmaceutical or biologic system is systematically perturbed to yield specific information concerning the effect of one on the other.

In biopharmaceutics, the study of **'liberation'** of drug from the drug product at the desired site is the first and the most important step. Liberation of drug from dosage form broadly includes the factors that influence the (1) stability of the drug within the drug product, (2) rate of drug release from the drug product, (3) rate of dissolution of the drug at the absorption site and (4) availability of the drug at its site of action. Liberation is thus controlled by the characteristics of the drug and the drug product.

Whereas physicochemical parameters of the drug and the dosage form can be accurately and precisely measured in vitro, meaningful quantitative estimates of drug absorption can be obtained only through in vivo pharmacokinetics. Pharmacokinetic techniques provide the means by which the processes of drug **absorption**, **distribution**, **metabolism**, and **excretion** are quantified in the intact organism. Inherent in the design of a suitable drug product is knowledge of the pharmacodynamics of the drug, including the desired onset time, duration, and intensity of drug action, collectively termed as **response**.

The process of drug liberation (L), absorption (A), distribution (D),

metabolism (M), and excretion (E) and assessment of drugs therapeutic response (R), i.e. **LADMER** system, are the major biopharmaceutic and pharmacokinetics topics for research and regulatory considerations in drug development. Figure 10.1 shows a schematic diagram of the LADMER system. On both sides of the diagram are the five processes, liberation, absorption, distribution, metabolism, excretion and response.

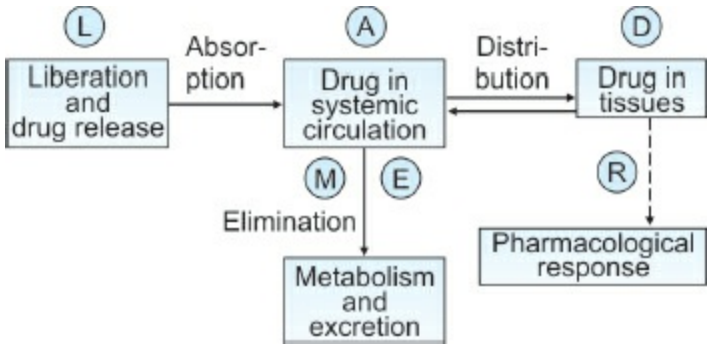



**Fig. 10.1:** Schematic diagram of the LADMER system

A drug administered in a dosage form by any route of administration must *liberate* from the dosage form and must get dissolved for *absorption*. Irrespective of the route of administration, in order to get absorbed, the drugs must pass through biological membranes which act as lipid barriers sans for intravenous route, where drug is in direct contact with the systemic circulation. Different transport mechanisms are employed to penetrate into and to permeate through these barriers. Once absorbed, the drug may directly enter the central compartment (systemic circulation) if the route of drug administration is intramuscular, subcutaneous, intraperitonial, intracutaneous, nasal, ocular or pulmonary. Drugs administered per-orally and some of the drugs administered rectally are confronted with enzymes as they pass through the liver. In the liver, some drugs are activated and most of the drugs are inactivated or metabolized, the process called as first-pass metabolism. In the systemic circulation, the drug exists either in free form or bound to plasma proteins. The free or unbound form of the drug fraction is bioavailable and elicits therapeutic response. The protein-bound fraction is not permanently trapped but is in equilibrium, and is released from the protein as the free drug is eliminated from the plasma. The drug may enter the peripheral compartment (***distribution***) by again passing through a lipid barrier until it reaches the biophase. After producing the therapeutic response, the drug again passes through lipid barrier and enters the central compartment from

where the drug by again passing through a lipid barrier is ***metabolized*** in the liver, kidney, tissue or plasma. It then passes either via biliary excretion into the intestine or through the kidney in the form of urine. ***Elimination*** is not only by urinary and bilary means but also through salivary glands and lungs. Reabsorption takes place in the kidney by tubular reabsorption but also in the intestine after enterohepatic cycling if the drug or its metabolite is in absorbable form. All these factors are involved in determining whether the drug administered will produce the therapeutic ***response*** or yield only a sub therapeutic effect, or even show toxic effects.

Knowledge and understanding of LADMER system enable the scientists to design a drug product controlling these factors. Onset of action, intensity of effect and duration of effect are controllable and the sum of all these phenomena is the quantitative characteristic of a drug product's effect. For most of the drug products, a relatively rapid and quantitative absorption, and a slow elimination are required, thus maintaining a therapeutic drug concentration for a long period of time. In some cases, this goal may easily be achieved if the drug is soluble, highly unionized, absorbed by passive diffusion, and has a long elimination half life. If this is not the case, many manipulations are necessary to create a drug product of the desired characteristics.

In this chapter, the basic pharmacokinetic concepts and techniques are reviewed; the physicochemical and biologic factors that influence drug absorption, elimination, and accumulation are discussed; and finally, the specific application of these general principles to the design and evaluation of drug dosage forms is illustrated. Other texts and monographs should be consulted for historical perspective and more extensive treatments of these topics.

## ABSORPTION

Absorption is defined as the amount of unchanged drug that reaches the general circulation. Hence, the drug which is metabolized or chemically transformed at the site of application or in transit is, by definition, not absorbed. Systemic drug absorption from a drug product consists of a succession of rate processes as shown in Fig. 10.2. For solid oral dosage forms such as tablet and capsule, the rate processes include (1) dosage form disintegration, (2) deaggregation and subsequent drug release, (3) dissolution of the drug and (4) absorption through cell membranes into the systemic circulation. In the above sequence, the slowest step determines the rate at which the drug reaches the systemic circulation. Disintegration of a solid oral drug product is usually more rapid than drug dissolution and absorption, with the exception of controlled-release products, where drug release is rate limiting. For poorly aqueous soluble drugs, the rate at which the drug dissolves (dissolution) is often the slowest step, and therefore, exerts a rate-limiting effect on drug bioavailability. In contrast, for a drug that has a high aqueous solubility, the rate at which the drug permeates (absorption) is the slowest or rate-limiting step.

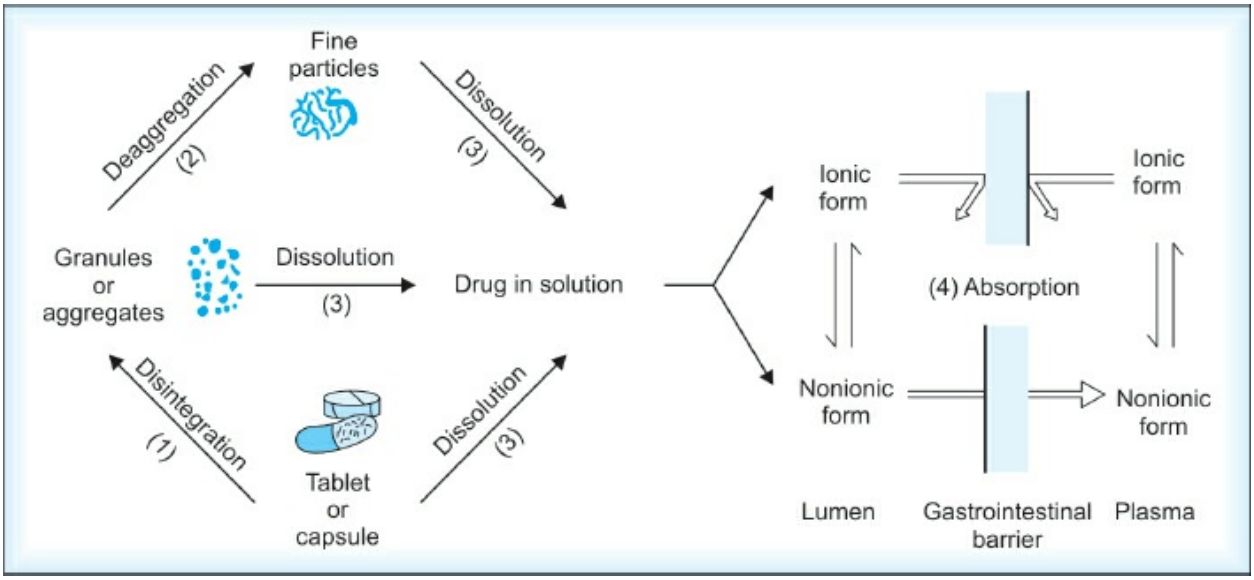



**Fig. 10.2:** Summary of processes involved in systemic drug absorption following the oral administration of a tablet or capsule drug product

## Drug Absorption Mechanism

To gain access into the systemic circulation, a drug must pass from the absorption site through one or more layers of cells or biological membranes. The permeability of a drug through biological membranes is intimately related to the size, physicochemical characteristics and molecular structure of the drug, and the physical and biochemical properties of the cell membranes. For absorption into the cell, a drug must traverse the cell membrane. The process of movement of a drug across a cell is called as transcellular absorption. Some polar molecules may not be able to traverse the cell membrane, but instead, traverse through “tight junctions” between cells, a process known as paracellular drug absorption. Some drugs are probably absorbed by a mechanism involving one or more of these processes.

### *Passive Diffusion*

Passive diffusion is the major transcellular process for nearly 90% drugs. It is the process by which molecules spontaneously diffuse from a region of higher concentration to a region of lower concentration without the expense of external energy (Fig. 10.3). Drug molecules move randomly forward and backward across a membrane. If the regions across a membrane have the same drug concentration, no net transfer of drug occurs. For a region that has a higher drug concentration, the number of forward-moving drug molecules will be higher compared to the number of backward-moving molecules, resulting in a transfer of molecules to the region with the lower drug concentration.

Thus, the driving force for passive diffusion is the difference in drug concentrations on either side of the cell membrane.

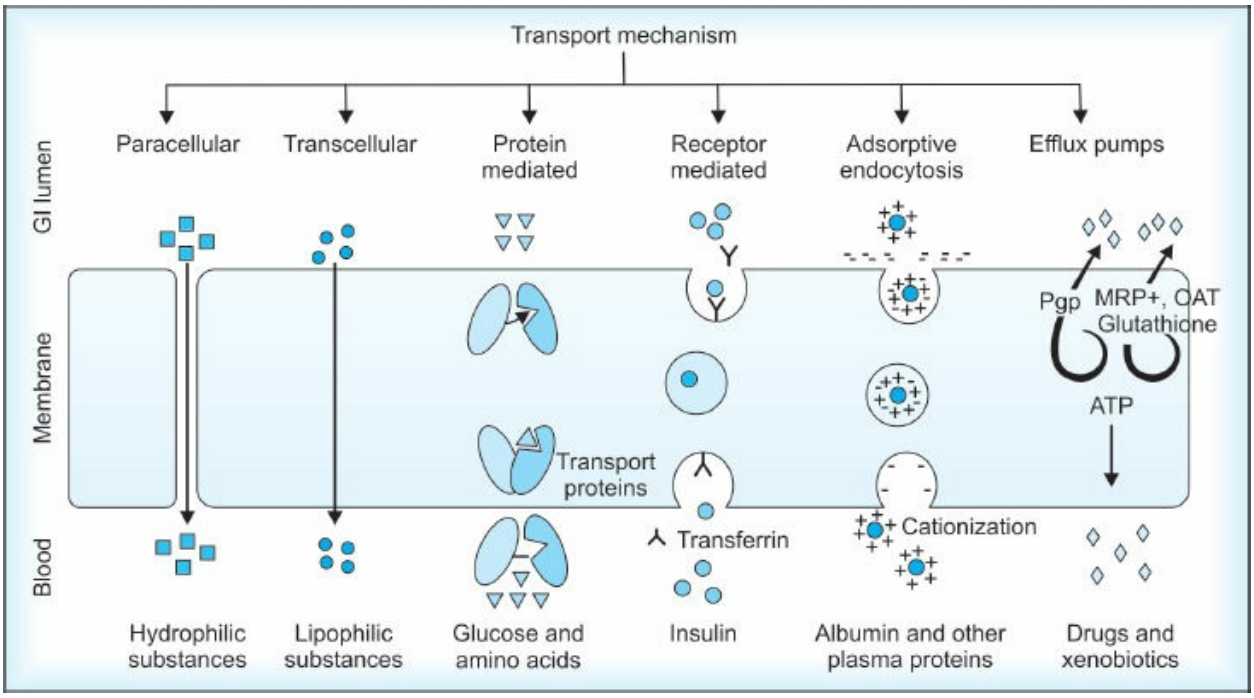



**Fig. 10.3:** Hypothetical representation of drug absorption processes

Passive diffusion is based on Fick's Law of Diffusion which can be expressed as follows:

$$\frac{dQ}{dt} = \{DAK/h\}\left(C_{GI} - C_p\right) \quad \ldots (1)$$

where, $dQ/dt$ = rate of diffusion; $D$ = diffusion coefficient; $K$ = partition coefficient; $A$ = surface area of membrane; $h$ = membrane thickness; and $C_{GI}$-$C_p$ = difference between the concentrations of drug in the GI tract and plasma.

Drug distributes rapidly into a large volume after entering the blood resulting in a very low plasma drug concentration with respect to the concentration at the site of drug administration. A large concentration gradient is maintained driving drug molecules into the plasma from the GI tract. As shown by Fick's Law of Diffusion, lipid solubility of the drug, surface area and thickness of the membrane influence the rate of passive diffusion of drugs.

## *Carrier-mediated Transport*

For a lipophilic drug with molecular weight less than 500 daltons, the lipid cell membrane is not a barrier to drug diffusion and absorption. Such

molecules are also absorbed by paracellular transportation. However, for the absorption of a hydrophilic drug, especially ions and nutrients required by the body, specialized carrier-mediated transport systems are present in the body. The carrier-mediated absorption is structure-specific, capacity-limited and site specific.

## *Active Transport*

It is a carrier-mediated transcellular process that is characterized by the transport of drug molecules from regions of low drug concentrations to regions of high concentration, i.e. against the concentration gradient. Active transport is an energy-consuming system whereby a carrier binds the drug to form a carrier-drug complex that shuttles the drug across the membrane and then dissociates the drug on the other side of the membrane (Fig. 10.3). The active transport is important for biliary and renal secretion of many drugs and metabolites, and is also involved in the GI absorption of some drugs. Drugs of structure similar to a natural substrate that is actively transported may compete for absorption sites on the carrier. A few drugs such as (i) levodopa and methyldopa via L-amino acid transporter, (ii) enalapril via peptide transporter, and (iii) 5-fluorouracil via pyrimidine transporter, are absorbed by this process. In contrast to passive diffusion, active transport is a capacity-limited (saturable) process because only a certain number of carriers are available and the binding sites on the carrier may become saturated at high drug concentrations.

## *Facilitated Diffusion*

Facilitated diffusion is a carrier-mediated transport system in which the drug moves along a concentration gradient (Fig. 10.3). Facilitated diffusion is non-energy requiring, saturable, structurally-selective for the drug and shows competition kinetics for drugs of similar structure. Facilitated diffusion seems to play a very minor role in drug absorption. Example of such a transport system includes absorption of Vitamin $B_{12}$ by using intrinsic factor as a carrier.

## *Carrier-mediated Intestinal Transport*

Various carrier-mediated transporters such as P-Glycoprotein (P-Gp) and amino acid transporter are present at the intestinal brush border and

basolateral membrane. P-Gp appears to reduce apparent intestinal epithelial cell permeability of various lipophilic or cytotoxic drugs. Amino acid transporter has a role in the absorption of many oral cephalosporins.

### *Vesicular Transport*

Vesicular transport or endocytosis is a minor absorption process of engulfing particles or fluids by the cell (Fig. 10.3). Phagocytosis refers to the engulfment of larger particles or macromolecules, whereas pinocytosis refers to the engulfment of small solutes or fluid. During vesicular transport, the cell membrane invaginates to surround the material, and then engulfs the material into the cell. Subsequently, the cell membrane containing the material forms a vesicle or vacuole within the cell. Orally-administered Sabin polio vaccine and various large proteins are proposed to be absorbed by vesicular transport. The processes of moving macromolecules out of a cell, termed as exocytosis is exemplified by the transport of insulin from insulin-producing cells of the pancreas.

## Factors Affecting Drug Absorption

In general, drug absorption from dermal, vaginal, rectal, parenteral, or GI absorption sites into the systemic circulation occurs by passive diffusion across the biologic membranes. These membranes form lipoidal barriers that separate the body's interior from its exterior environment. The rate of diffusive movement, *dA/dt*, across a homogeneous membrane is governed by Fick's law, that is:

$$\frac{dA}{dt} = D_c P C S \frac{dC}{dX} \qquad \text{... (2)}$$

where, $D_c$ is the diffusion coefficient of the drug through the membrane, $P_C$ is the partition coefficient between the membrane and the donor medium containing the drug, $S$ is the membrane surface area, $dC$ is the concentration differential across the membrane, and $dX$ is the membrane thickness. In actual practice, concentrations on the receptor side of the membrane are low because of continuous blood flow. Thus, when the concentration on the donor side is relatively high, Eq. (2) reduces to:

$$\frac{dA}{dt} = P_m S C \qquad \text{... (3)}$$

where, $C$ is the drug concentration at the absorption site, and $P_m$ is the permeability constant defined by:

$$P_m = \frac{D_c P_c}{dX} \qquad \text{... (4)}$$

For solid dosage forms, drug concentration at the absorption site is a function of the dissolution rate of the drug in the medium at that site. The dissolution is given by the Noyes-Whitney equation:

$$\frac{dC}{dt} = \frac{D_c S}{h}(C_s - C) \qquad \text{... (5)}$$

where, $C$ is the concentration at time t, $D_c$' is the diffusion coefficient of drug in the medium, $S$ is the surface area of drug particles, $h$ is the thickness of the diffusion layer surrounding the particles, and $C_s$ is the solubility of the drug in the diffusion layer.

### *Physical Factors*

## *Solubility*

Drug absorption requires that molecules be in solution at the absorption site. Dissolution of solid dosage forms in GI fluids is a prerequisite to the delivery of a drug to the systemic circulation following oral administration. Dissolution depends, in part, on the solubility of the drug substance in the surrounding medium. Polar solutes are more soluble in water than in organic phases, while the reverse is true for non-polar solutes. Ionized species have a greater aqueous solubility than their un-ionized counterparts. The total solubility of acids or bases in aqueous medium is therefore pH-dependent. For drugs absorbed by passive diffusion, those exhibiting low aqueous solubility tend to have a slower oral absorption rate than those exhibiting high aqueous solubility.

For drugs intended for topical application (e.g. vaginal, rectal, dermal), the solubility of the drug in the vehicle is important. For a given vehicle, the highest driving force for absorption is obtained when the drug concentration in the vehicle equals its solubility. Concentrations below saturation decrease the absorption efficiency, while concentrations exceeding drug solubility serve as reservoirs to maintain a saturated solution.

## *Particle Size*

Surface area of drug particles is another parameter that influences drug dissolution, and in turn, drug absorption. Particle size is a determinant of surface area. Small particles with greater surface area dissolve more rapidly than larger particles, even though both have the same intrinsic solubility. Particle size appears to have little influence on the absorption of drugs with high aqueous solubility, but it may have a pronounced effect on the absorption of drugs with low aqueous solubility. The absorption of griseofulvin, a neutral compound with low aqueous solubility (15μg/ml), is poor and erratic. Increasing particle surface area through micronization markedly improves absorption, as illustrated in Fig. 10.4. Further reduction of particle size by formation of fine solid dispersions in polyethylene glycols results in an approximate doubling of absorption efficiency, compared with conventionally-micronized formulations.

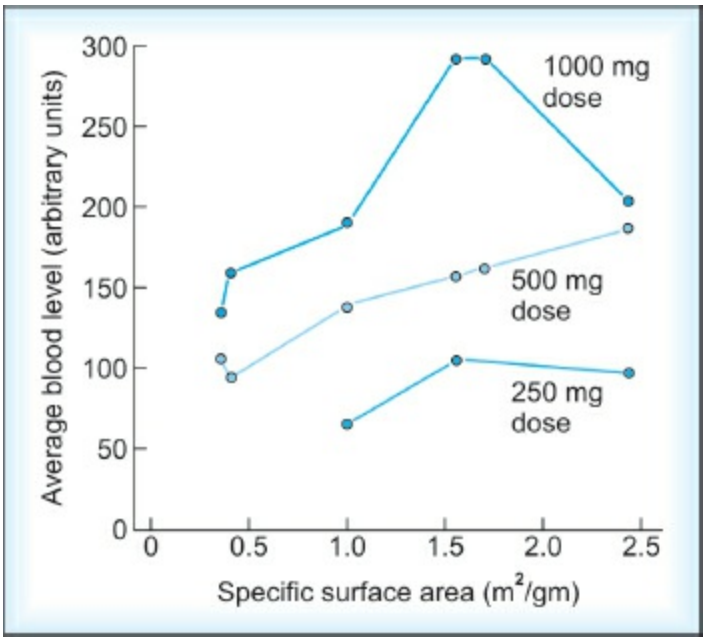



**Fig. 10.4:** Effect of specific surface area on the absorption of griseofulvin in humans

## *Crystal Form*

Polymorphs are crystal forms differing in the packing and orientation of molecules under different crystallizing conditions. The physicochemical properties of these crystal forms (e.g. density, solubility, melting point) are influenced by the intermolecular forces present. For example, polymorphs with weak attractive forces (thus, in a high-energy state) exhibit greater solubilities than those with strong attractive forces. Hence, differences in dissolution and absorption rates between polymorphs of a given compound may also be observed. The rate of absorption of chloramphenicol appears to be directly related to the solubility of the different crystal forms of its palmitate ester. When differences in crystal energy are small, effects of polymorphism on absorption may not be observed. Three polymorphs of chlorpropamide have been shown to yield comparable serum drug concentrations.

## *Dissociation Constant*

The un-ionized species of acidic or basic compounds in solution penetrates lipoidal membranes of the GI tract more efficiently than the ionized species. The rate of GI absorption of a drug, therefore, is directly related to the

concentration of its unionized species at the absorption site, which is a function of the *pKa* of the compound and pH of the environment. The pH of the GI tract ranges from approximately 1.2–3.5 in the stomach, 5.0 to 6.0 in the duodenum, 6.5 to 8.0 in the jejunum and large intestine. Over the pH range 1 to 8, the un-ionized fraction changes dramatically for acids with *pKa* values between 2.5 and 7.5 (decreasing with increasing pH) and for bases with *pKa* values between 5 and 11 (increasing with increasing pH). For these compounds, pH-dependent absorption is expected. Weak acids with *pKa* values greater than 7.5 and bases with *pKa* values less than 5 have pH-independent absorption. In general, drugs with *pKa* values from 5 to 7 are more readily absorbed than acidic drugs with high *pKa* values. Amphoteric compounds manifest the least absorption difficulties when they are presented as zwitterions, while neutral compounds do not exhibit pH-dependent absorption.

## *Chemical Factors*

## *Lipophilicity*

Biologic membranes, being lipoidal in nature, are usually more permeable to lipid-soluble substances. Transport across these membranes therefore depends, in part, on the lipid solubility of the diffusing species. Lipid solubility of a drug is determined by the presence of non-polar groups in the structure of the drug molecule as well as by ionizable groups that are affected by local pH. The unionized drug species exhibit a greater lipid solubility than the ionized species. The relative lipophilic to hydrophilic properties of the entire molecule, described by the partition coefficient, determine whether the molecule readily undergoes passive diffusion across the GI or other biologic membranes. In general, the lipid/water partition coefficient of a molecule is a useful index of its propensity to be absorbed by passive diffusion. A high lipid solubility, however, does not necessarily favour absorption unless the water solubility is relatively low so that the drug is not “trapped” in the aqueous phase. On the other hand, when the water solubility is too low, a significant concentration cannot be achieved at the membrane surface, and absorption may be inefficient despite a favorable partition coefficient.

The absorption of a drug may often be enhanced through appropriate structural modifications that serve to alter the relative lipophilicity/hydrophilicity of the compound, e.g., esterification of a water-

soluble acid. Another approach to enhanced absorption of compounds with poor lipid solubility is that of inclusion of adjuvants in the dosage form, which rather than altering the lipid solubility of the drug in question, possibly enhance absorption by altering the permeability of the absorbing membrane. For example, salicylate may interact with calcium or magnesium ions in the rectal membrane and thus facilitate rectal absorption of theophylline.

## *Stability*

Chemical integrity must be maintained until the compound is delivered to its intended site of absorption or application. Obviously, chemical instability in the dosage form, or instability prior to transport across the initial biologic barrier, invariably affects bioavailability.

Salt formation is a chemical modification that usually enhances the aqueous solubility; however, aqueous solubility per se may not be the sole determinant of bioavailability. For example, salts of weak acids may precipitate on initial contact with the gastric environment following oral administration. Hence, dissolution of the precipitate is a prerequisite to absorption. In a static situation, the concentration and partition characteristics of the drug at the pH of the mucosal surface where absorption occurs can be independently studied. The dynamics of transport along the GI tract and across the GI membranes, however, are such that the effect of salt formation on the bioavailability of the sample drug is usually unpredictable. The observed effect is a function of the dissolution rate of the salt or its precipitate, gastric pH, gastric-emptying time, intestinal motility, *pKa* of the drug, and so on.

Chemical instability is often a function of pH. Compounds that are highly labile in the neutral range are seldom useful as drugs. Whereas instability in the alkaline range is seldom encountered under physiologic conditions, stability in acid is a concern, particularly for drugs intended for oral administration. For example, gastric instability of penicillin G is a major factor in its poor and erratic bioavailability following oral administration. This problem led to the synthesis of the acid-resistant phenoxyalkylpenicillins, of which penicillin V is a member.

Prodrug formation is commonly used to enhance absorption of a drug by chemical modification. An ideal prodrug is one that is quantitatively absorbed and biotransformed to the parent drug during its transport to the site of action

(e.g. systemic circulation, brain, epidermis). Prior chemical degradation, including the formation of the active moiety, adversely affects bioavailability.

## *Metabolic Factors*

In addition to the physical and chemical factors that may affect drug absorption are metabolic factors. Drugs may be exposed to presystemic biotransformation when administered by a nonintravascular route.

## *Gastrointestinal Tract*

The drug absorption pattern followed in the GI tract is depicted in Table 10.2. Pre-systemic elimination of drugs occurs most often following oral administration. Biotransformation can occur in the gut lumen, gut wall, or liver. Drug metabolizing enzymes that are found in the upper intestine probably originate in secretions of Paneth's cells, or from cells shed into the lumen from the mucosal lining. Such enzymes have been implicated in the hydrolysis of phthalate esters and pivampicillin. These enzymes are inactivated by gut flora in the lower bowel.

**Table 10.2:** Drug absorption in the gastrointestinal tract

| Anatomic area (pH) | Function and anatomy | Affect on drug absorption |
|---|---|---|
| Oral cavity (pH 7) | Saliva, contains ptyalin, digests starches; mucin, lubricates food and may interact with drugs. | Buccal and sublingual absorption occurs for lipid-soluble drugs. |
| Esophagus (pH 5–6) | Prevents acid reflux from the stomach. | Very little drug dissolution occurs. |
| Stomach fasting state (pH 2–6) fed state (pH 1.5–2) | Acid secretion is stimulated by gastrin and histamine; break-down of large food particles. | Absorption not efficient; basic drugs are solubilized rapidly in acid. |
| Duodenum (pH 6–6.5) | Enzymatic digestion of | Main site for drug |

| | | |
|---|---|---|
| | protein and peptide; trypsin, chymotrypsin, and carboxy-peptidase are involved in the hydrolysis of proteins into amino acids; amylase is involved in the digestion of carbohydrates; Pancreatic lipase hydrolyzes fats into fatty acids; presence of microvilli | absorption; Immense surface area for the passive diffusion of drugs; ester prodrugs undergo hydrolysis; proteolytic enzymes degrade proteins; acid drugs dissolve in alkaline pH; bile helps to dissolve hydrophobic drugs. |
| Jejunum (pH 6.5–7) | Digestion of proteins and carbohydrates; preferred for in vivo drug absorption studies | Drugs generally absorbed by passive diffusion. |
| Ileum (pH 7) | Has fewer contractions than duodenum. | Drugs generally absorbed by passive diffusion. |
| Colon (pH 5.5–7) | Contains both aerobic and anaerobic microorganisms that may metabolize some drugs; lack of microvilli and more viscous lumen contents. | Limited drug absorption; theophylline and metoprolol are absorbed and drugs are good candidates for an oral sustained-release dosage form. |
| Rectum (pH 7) | The rectum is about 15 cm long, ending at the anus; rectum is perfused by the superior, middle, and inferior hemorrhoidal veins. | Variable drug absorption; Absorption of drug from suppository or drug solution placed into rectum; drug moves directly into systemic |

| | |
|---|---|
| | circulation when absorbed via lower hemorrhoidal veins; drug undergo first-pass metabolism prior to systemic absorption when absorbed via superior hemorrhoidal veins. |

The intestinal flora represents a highly diverse and relatively potent source of drug metabolic activity. At least 400 different species of bacteria are present in the human intestinal tract. These bacteria are mostly anaerobes involved in reductive reactions. Drugs containing nitro-groups may be reduced to amines, which can be toxic. Sulfa drugs, lactulose, and some cathartics may be activated by these bacteria. These reactions may be complementary to the metabolism occurring subsequently in the gut wall or liver, but they may also be regenerative (e.g. hydrolysis of glucuronide, sulfate or acylamide metabolites excreted in the bile). Diet, disease, and drugs contribute to differences in the number, type, and location of these bacteria. Since most are restricted to the lower bowel, the potential for bacterial degradation is the greatest for drugs administered rectally. Drugs that are rapidly absorbed following oral ingestion may not be exposed to bacteria in the lower intestine. On the other hand, bio-inactivation by gut flora may further decrease the bioavailability of compounds that are not efficiently absorbed in the upper GI tract.

The metabolism of a drug during transit through the gut wall also influences bioavailability. Drug-metabolizing enzymes are known to be located in the endoplasmic reticulum, mitochondria (monoamine oxidase), and cytosol (N-acetyltransferase). Some enzymes such as phenol and estrone sulfokinases exist throughout the GI tract, while others may be more localized in the jejunal mucosa (steroid alcohol sulfokinase). These enzyme systems fall into two categories (1) those that catalyze such pre-conjugation reactions as C-oxidation, hydroxylation, dealkylation, N-and S-oxidation, reduction, and hydrolysis, and (2) conjugative or synthetic reactions. Since many of these reactions may also occur in the liver, it is usually difficult to

quantify the relative contribution of either site to the overall metabolic scheme of a drug. It can be said, however, that pre-conjugation reactions that depend upon cytochromes P450 or P448 are quantitatively unimportant in the gut wall, where as synthetic ones such as O-sulfation are more highly developed. Although the pylorus, duodenum, and jejunum have the greatest metabolic activity, biotransformation can occur throughout the alimentary tract from the buccal mucosa to the rectum. Metabolism is a major source of variation in bioavailability and therapeutic response.

Physiologic changes in the GI tract, other drugs, diet, or disease may alter drug-metabolizing enzyme activity in the gut wall. Various sites along the GIT where drug may be enzymatically metabolized or chemically altered are shown in Fig. 10.5. Monoamine oxidase activity is affected by thyroid activity, systemic progesterone levels, and iron deficiency. At the mucosal level, induction of pre-conjugative reactions is usually unimportant, but drug competition for conjugation or enzyme inhibition is clinically relevant. Sulfation depends on the systemic supply of inorganic sulfate that can be depleted by such drugs as salicylamide. Interactions between sympathomimetics, which are sulfated, and other drugs which are metabolized similarly are potentially dangerous. A classic example of enzyme inhibition involves tyramine, which is normally deaminated in the gut wall by monoamine oxidase. Antidepressant monoamine oxidase inhibitors (e.g. iproniazid, isocarboxazid, nialamide, phenelzine) block this metabolic pathway and thus expose patients to severe hypertensive crisis following ingestion of tyramine-rich foods.

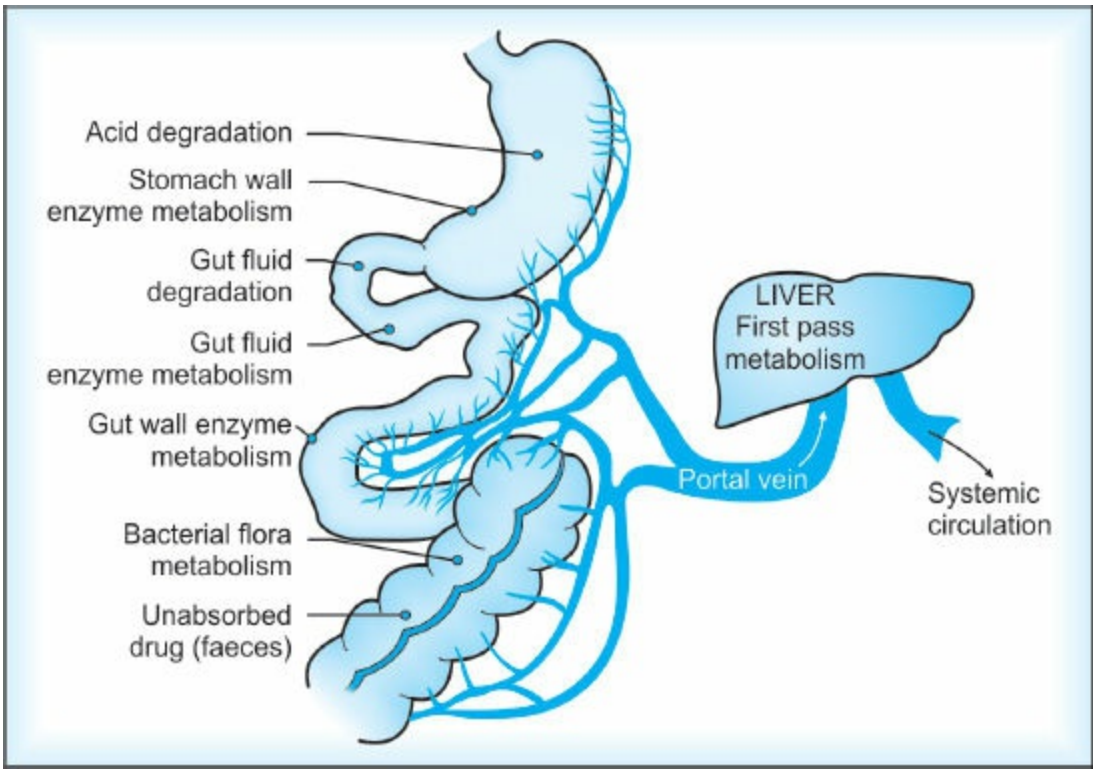



**Fig. 10.5:** Diagrammatic sketch indicating sites along the GIT where drug may be enzymatically metabolized or chemically altered

Foodstuffs are known to contain several compounds that may induce microsomal drug oxidation in the gut wall. Some plant indoles and poly-cyclic aromatic hydrocarbons produced in meat cooked over charcoal have been implicated in the enhanced metabolism of ethoxycoumarin and phenacetin by O-deethylation.

Although little is known about the effect of disease on the metabolic activity of the gut wall, celiac disease has been shown to reduce the conjugation of ethinyl estradiol but increase the sulfation of methyldopa.

## *Liver*

The most important site of presystemic drug metabolism is the liver. The liver receives its blood supply from the hepatic artery and hepatic portal vein. Approximately 75% of the hepatic blood flow stems from the portal vein, which drains all, but the lowest 10 cm and the uppermost 55 cm of the GI tract. Hence, drugs absorbed from the intestinal tract and the upper portions of the rectum must pass through the liver before reaching the systemic circulation. Drugs absorbed from the intestinal tract into the lymphatic

system may bypass the liver.

The types of metabolic reactions encountered in the liver are similar to those occurring in the gut wall. Unlike the gut wall, mixed function oxidases have a major role in pre-conjugative reactions in the liver, where glucuronidation is the most prevalent conjugative or synthetic reaction. A drug or its metabolites may undergo one or more of these reactions to form products having different pharmacologic activites. The net effect of first-pass hepatic metabolism of a drug is reduction of its bioavailability. Examples of drugs that undergo significant first-pass hepatic metabolism include antiarrhythmics, (e.g. lidocaine and verapamil), β-blockers (e.g. propranolol and metoprolol), centrally-acting analgesics (e.g. propoxyphene and pentazocine), and antidepressants (e.g. imipramine and amitriptyline). More comprehensive review of this subject may be found in the literature.

## *Lungs*

Since drug absorption following most routes of administration (oral, rectal, inhalation, intramuscular, buccal, transdermal, subcutaneous) places a drug in the venous side of the systemic circulation, the agent must pass through the lungs before reaching the arterial portion of the system. Drugs exposed to metabolic activity at the site of application, during the absorption process, or upon first pass through the liver risk further biotransformation upon entry into the lungs. The extensive capillary network in the lungs exposes the blood to an endothelial surface area of roughly 70 to 125 $m^2$. A large number of pinocytotic vesicles containing drug-metabolizing enzymes are found in lung tissues. On a mass basis, the lungs represent a smaller organ and contain proportionately more fibrous tissue than the liver. Blood flow to the lungs, however, is approximately three times that to the liver. Thus, while the intrinsic clearance of the lungs is normally smaller than that of the liver, total clearance by the lungs may be significant because of blood flow.

The lungs have been implicated in the metabolism of a number of compounds. For example, benzphetamine, aminopyrine, ethyl-morphine, and imipramine undergo demethylation, while ethoxycoumarin and coumarin undergo O-deethylation and aromatic hydroxylation, respectively. Reductive reactions in the lungs are not well understood, but dehydrogenation of steroids has been reported. The reduction of nitro groups to the corresponding amines is operative and is applicable for chloramphenicol. Hydrolytic

enzymes are numerous in the lungs and play an important role in the metabolism of endogenous compounds. Their role in the metabolism of xenobiotics, however, is not well understood. The major conjugative enzyme systems detected in the lungs are glutathione S-transferase, UDP glucuronyl-transferase, sulfotransferase, and N-acetyl-transferase, with glutathione conjugation being the most important.

## *Other Tissues*

Little is known about the pre-systemic metabolism of drugs that are administered intramuscularly, subcutaneously, nasally, or dermally; however, many of the enzymes identified in the liver, gut wall, or lungs may be present at these sites of administration as well, and could contribute to the pre-systemic elimination of agents so administered. The metabolic potential of the skin has been extensively evaluated. The skin contains many enzymes found in the liver as well and is capable of oxidation, reduction, hydrolytion, and conjugation reactions. Most of the enzyme activity appears to be localized in the epidermal layer, which constitutes only about 3% of the total skin. While enzyme activity in whole-skin homogenates is low compared to the liver, if the enzymes are concentrated in the epidermal layer, activities are actually close to those of liver enzymes. Such activity may result in first-pass metabolism (during absorption) of topically-applied drugs that are intended for systemic action.

## *Physiologic Factors*

The physiologic conditions at the site of drug application, residence time of the drug at the site, and shunting through body fluids also influence drug entry into the systemic circulation.

## *Site of Application/Route of Administration*

The large effective surface area of the GI tract exerts a great influence on the absorption of orally-administered drugs. The GI mucosal surface is a mass of folded tissue covered by projections of columnar epithelial cells called *villi* and *microvilli*. These structures increase the effective surface area of the intestinal tract 600-fold over its simple tube-like appearance. Surface area decreases in a distal direction, suggesting that passive drug absorption is less efficient as the drug migrates toward the colon. Active transport processes, more prevalent in the ileum than in the upper small intestine, may

compensate for the decrease in passive absorption.

The buccal cavity is richly supplied with capillaries. Venous return bypasses the liver. Many of the chemical and metabolic processes encountered with oral administration are avoided when drugs are administered by the buccal route.

The rectal cavity has a surface area of 200 to 400 $cm^2$. In general, drugs absorbed in the lower region of the rectum enter directly into the systemic circulation. Those absorbed in the upper region pass through the liver first. Anastomoses among the rectal veins complicate this picture.

The nasal mucosa represents a site of drug administration that has absorption characteristics similar to those of the intramuscular administration. An ample blood supply, neutral pH, and low enzymatic activity contribute to these characteristics. If other than local effects are intended, drug potency is of prime concern since maintaining relatively large amounts of drug at the absorption site is difficult.

Until recently, dermal application of drugs was intended for local effects only. As transport of substances through skin is now better understood, lipophilic drugs that are reasonably potent are being incorporated into transdermal dosage forms with the intent of establishing its therapeutic blood levels. Since skin of various portions of the body has been shown to have different permeability characteristics for a given drug, selection of an site appropriate application is important. The rate of penetration depends on diffusion of drug from the vehicle to the skin surface, surface pH considerations, and diffusion through the stratum corneum and supportive tissues. Drug may also move along hair follicles, sweat glands, sebaceous glands, or aqueous channels.

Bioavailability is usually good following intramuscular and subcutaneous drug administration. Mostly, a large muscle mass is chosen for intramuscular injection. Muscle is richly supplied with blood, and absorption of drug is efficient. Though usually not a problem, the potential for metabolism at these injection sites exists. On the other hand, blood flow is more restricted in subcutaneous tissue; therefore, absorption from a subcutaneous injection is likely to be sustained.

## *Residence Time*

The time for which a drug remains at its site of absorption may affect its bioavailability. The effect is minimal with intramuscular or subcutaneous injection, or with dermal application, since the drug is confined to the site of application. Drug may be available for absorption from these sites for a protracted period. Because of leakage, swallowing, or expectoration, absorption following the intranasal and buccal administration may be more variable and less complete.

Drugs taken orally enter the stomach, where gastric emptying regulates movement into the intestinal tract. Once drug has moved into the small intestine, transit should proceed unimpeded until drug reaches the colon. The mouth-to-caecum transit time in healthy humans has been shown to be as short as 1.5 to 3.5 h. Hence, within 3 h of ingestion, drug may enter the large intestine, where absorption may be inherently less efficient or impeded by the presence of fecal material. Depending on the dosage form and presence of food, however, the stomach and small intestine may not be completely free of drug for up to 5 and 20 h, respectively. The physicochemical nature of the drug, type of formulation, and sites of absorption along the GI tract will determine the effect of delayed gastric emptying or decreased intestinal motility on drugs bioavailability.

With regard to rectal drug administration, residence time in the rectal cavity is governed by leakage, defecation, or upward migration of the drug into the lower bowel. The first two factors would reduce or terminate drug absorption, while the latter would increase drug exposure to first-pass metabolism in the liver.

## *Shunting and Recycling*

The plasma concentration-time profile of a drug may be altered by shunting, a process that may occur following drug administration. Shunting refers to the entry of a drug into a body fluid (other than blood) before or after entry into the systemic circulation. Depending on the route of administration and the efficiency of the shunting process, the drug might conceivably never appear systemically. Usually, however, the drug is recycled, through absorption or mixing, into the systemic circulation. One example of shunting involves secretion of the drug from blood into the parotid or submaxillary fluids.

The liver may also secrete the drug from blood into bile. Bile is then

stored in the gallbladder, which empties periodically, albeit irregularly, into the small intestine. Drug may then be reabsorbed, and the process recurs. A notable example of this is the drug indomethacin, for which it has been estimated that 50 to 60% of a dose administered orally, rectally, or intravenously may be recycled through the bile.

Finally, the lymphatic system is a pathway through which the drug, having been absorbed gastrointestinally or parenterally, may be shunted, only to reappear later in the systemic circulation. Nearly all tissues of the body have lymphatic channels that drain excess fluid directly from the interstitial spaces. The lymphatic system is a major pathway for absorption of fatty substances from the GI tract. About one tenth of the fluid filtering from arterial capillaries enters the terminal lymphatic capillaries. This shunt is particularly important for substances with high molecular weight because lymphatic capillaries are much more permeable than venous capillaries. Lymph generated in the lower half of the body, and in the left half of the upper body, eventually re-enters the systemic circulation at the juncture of the left internal jugular and subclavian veins, whereas lymph from the right half of the upper body re-enters contralaterally. The flow of lymph is slow as compared to blood flow dynamics. Hence, the initial effect of drug shunting through the lymphatics is a reduction of drug concentration in the blood, followed later by a sustaining effect.

## DISTRIBUTION

Drug distribution is the reversible transfer of a drug between the blood and the extravascular fluids and tissues. Once a drug has entered the vascular system, it becomes distributed throughout the various tissues and body fluids in a pattern that reflects the physiochemical nature of the drug and the ease with which it penetrates different membranes. Distribution usually occurs much more rapidly than elimination.

Drug molecules are distributed throughout the body by means of the circulation of blood. The entire blood volume (about 6 L) is pumped through the heart each minute, and within minutes after a drug enters the blood stream it is diluted into the blood volume. A drug that is restricted to the vascular space and can freely penetrate erythrocytes has a true volume of distribution of 6 L. If the drug cannot penetrate the RBCs, the available space is reduced to about 3L (plasma volume). Certain body fluids may be relatively inaccessible to drugs in the bloodstream; which include cerebrospinal fluid (CSF), bronchial secretions, pericardial fluid and middle ear fluid.

The degree of access of antibiotics to these fluids may be a limiting factor in treating diseases. Drug concentrations in body fluids also depend upon the degree of binding in the fluid. Drug concentrations are lower in CSF and saliva, which are usually protein free, and in extravascular fluid (ECF), which has lower albumin concentration compared to plasma. The rate of distribution to each organ is determined by the blood flow perfusing the organ and the ease with which the drug molecules cross the capillary wall and penetrate the cells of a particular tissue.

Blood flow is the rate-limiting step in the distribution of most drugs. Accordingly, rapid equilibration of lipid-soluble drugs is observed between the blood and kidney, liver, heart and brain, all of which are highly perfused with blood. Less rapid equilibration is found in skeletal muscles and adipose tissue, which receive a considerably smaller volume of blood per unit mass per minute. Redistribution of the drug into less well-perfused tissues, rather than metabolisim and excretion, may limit the duration of effect of certain drugs at highly perfused sites. For example, thiopental produces anaesthesia within seconds after administration because of rapid equilibriation between blood and brain. The duration of effect is short lived despite the fact that it is slowly metabolized. The rapid decline in brain drug concentration is the

result of distribution into other tissues.

## Volume of Distribution

The volume of distribution ($V_d$) is not a "real" volume but an artifact. It is a parameter of a particular model used to fit experimental data. The volume of distribution depends on many factors such as blood flow rate in different tissues, lipid solubility of the drug, partition coefficient of the drug type of tissue, pH of surrounding and binding to biological materials. The $V_d$ is often proportionate to the body weight. However, obesity and oedema produce abnormal deviations. In obesity, $V_d$ of hydrophilic drugs is lower and in edematic patients, the $V_d$ of hydrophilic drugs is larger than expected from body weight.

Since blood is the only accessible body fluid in which we can determine the drug concentrations, and since in the case of intravascular administration we know the amount of drug injected directly into the bloodstream, we are able to determine the $V_d$ as soon as the drug has equilibrated (steady state) by dividing the dose administered intravenously by the drug concentration in the blood ($C$).

$$V_d = Dose/C \quad \ldots (6)$$

$V_d$ is a proportionality constant. It can be smaller or larger than the true physiologic fluid spaces of the body, depending upon whether the affinity of the drug is highest for plasma constituents or for other tissues. For a normal 70-kg man, the plasma, blood, extracellular fluid and total body water volume is about 3, 6, 12 and 42 L, respectively. $V_d$ can be calculated from the plasma concentration versus time data by means of non-compartmental methods for a bolus IV dose:

$$V_{ss} = \frac{Dose_{IV} AUMC_{0\to\infty}}{AUC_{0\to\infty}} \quad \ldots (7)$$

Here, $V_{ss}$ represents the volume in which a drug would appear to be distributed during steady state, $AUC$ is the area under the plasma concentration-time curve, and $AUMC$ is the area under the first moment of plasma concentration-time curve, i.e. the area under the curve of the product of time, $t$, and plasma concentration, over the time span zero to infinity.

$V_d$ at steady state can also be determined by compartmental methods, i.e. by using the coefficients and exponents of a multiexponential fit to the data.

## CNS Drug Distribution

The capillaries in the brain are different in their permeability characteristics from those found in the rest of the body. They possess a cellular sheath that makes them much less permeable to water-soluble substances. This sheath constitutes the so-called blood brain barrier (BBB). The penetration rate of a drug into the brain depends on its degree of ionization in the plasma and its lipid solubility. Highly lipid soluble drugs, like thiopental, reach the brain almost immediately after administration. More polar compounds like barbital, penetrate the CNS at a lower rate.

The rapid penetration of lipid-soluble drugs in the CNS is of great importance for anticonvulsant or psychotropic drugs, which must act on the brain, but facile penetration may produce unwanted side effects with drugs intended to affect other systems. The CSF concentrations of lipid-soluble drugs usually reflect free drug concentrations in the plasma. Drug levels are usually higher in the brain tissue than in the CSF. The BBB may severely limit treatment with antibiotics, cancer chemotherapeutics and other polar drugs.

## Placental Drug Distribution

Since the thalidomide tragedy, there has been a keen interest in the passage of drugs across the placenta. The membranes separating foetal capillary blood from maternal blood resemble cell membranes elsewhere in the body. Many drugs of moderate-to-high lipid solubility, including sulphonamides, barbiturates, anticonvulsants, narcotic analgesics, antibiotics and steroids can be detected in appreciable concentrations in foetal blood or tissues shortly after administration to mother.

Although its development has been greatly hampered by ethical and experimental difficulties, the theory of maternal equilibrium rates is considerable. The shortest possible time for a drug to equilibrate between maternal blood and foetal tissue has been estimated to be about 40 min. Drugs like tubocurarine whose passage across the placenta is impeded by low lipid solubility, large molecular size, ionization, and protein binding require hours for equilibration. The higher the blood levels of a drug in pregnant patients on chronic medication, the greater the risk to the foetus. Since some degree of foetal exposure is likely to occur with virtually all drugs, and since the consequence of such an exposure are unknown, many advocate that drug administration during pregnancy must be severely restricted.

## Drug Protein Binding

Binding is usually a reversible interaction between a small molecule such as drug or metabolite and a protein or other macromolecule. Only unbound drug can be transported across capillaries, distributed to tissues, gain access to metabolizing enzymes, and interact with receptors to elicit a pharmacologic effect. Binding is a distributive process, which delays drug elimination. Tissue binding cannot be measured directly in the intact, living animal or human although its influence on drug distribution and elimination has been deduced and discussed.

Blood plasma contains 93% water and the remaining 7% consists of different dissolved compounds, primarily proteins. The main protein fraction is albumin which constitutes approximately 5% of the total plasma. Proteins are found not only in the plasma but also in tissues. Many drugs, particularly those having low molecular weight, are bound to plasma and tissue proteins.

Human serum albumin has a molecular weight of approx. 67,500 daltons and is composed of 20 different amino acids. The different kinds of amino acids and their relative position in the protein molecule determine the binding of drugs. Amino acids such as arginine, histidine and lysine contain basic groups and are responsible for binding acidic drugs. Amino acids such as aspartic acid, glutamic acid, and tyrosine are responsible for binding basic drugs due to the presence of acidic groups.

Drugs can be bound to albumin either by van der waal’s forces, hydrophobic bonds, hydrogen bonds, or ionic bonds. Acidic drugs are strongly bound to albumin, usually one or two molecules per molecule of albumin. Basic, positively-charged drugs are weakly bound to a large number of sites on the albumin molecule. Often, interactions between basic drugs and proteins are without clinical importance. Although albumin constitutes the largest proportion of plasma proteins and is predominantly responsible for binding of drugs, there is another protein factor, α-acid glycoprotein (AAG), which has received considerable attention during the past years, AAG, also known as orosomucoid, is an α-globulin, having a molecular weight of 41,000–45,000 daltons and consists of a linear polymer of amino acids with branching chains of carbohydrates. The plasma concentration of AAG is normally only 0.6–0.8 g/L. AAG has only one high-affinity binding site and binds only basic, highly lipophilic drugs.

Drug binding to protein is usually rather non-specific, i.e. many drugs bind to the same binding sites (receptors) on the protein molecule. The drug with the higher affinity will therefore displace a drug of lower affinity from its binding site by competition. Since only the non-protein bound fraction is free for pharmacological action, intensity of pharmacological response, side effects and toxicity increase in displacement from protein binding sites. This is of importance only in drugs bound to protein. If, for instance, a drug is 98% bound to proteins and another drug given simultaneously competes for the same binding site and displaces some of the first drug so that it becomes only 94% bound, the level of free, nonprotein bound drug increases from 2 to 6%, i.e. 200% more of free drug is available. The extent of protein binding is determined in vitro by dialysis, ultracentrifugation, ultrafiltration, sephadexgel filtration, molecular filtration, electrophoresis, or by agar plate test. The extent of protein binding is usually given in percentage. However, one must be aware that percentage bound is a function of the capacity of the protein and the concentration of the drug bound in the environment.

## *Drug Protein Binding Equilibria*

Binding is a function of the affinity of the protein for drug as well as the concentration of drug and protein. By far, the most important binding proteins in plasma are albumin for acidic drugs and $\alpha_1$-acid glycoprotein for basic drugs. The interaction between drug and protein can be represented by the law of mass action, such that:

$$D + P \underset{K_d}{\overset{K_a}{\rightleftharpoons}} DP$$

$$\text{Drug} + \text{protein} \rightleftharpoons \text{drug} - \text{protein}$$

With '*n*' type of binding site, protein binding can be described mathematically by the equation:

$$r = \frac{K_a[P][D]}{K_a[P][D]+[P]} = n\frac{K_a[D]}{K_a[D]+1} \quad \ldots (8)$$

where, [D] is free-drug concentration, [P] is concentration of free protein with '*n*' binding sites per molecule. Thus [*nP*] is the total concentration of protein binding sites, and [*rP*] is the concentration of bound drug or bound protein with '*r*' drug molecules bound per protein molecule. Typically, there may be 1–4 binding sites per protein molecule (Fig. 10.6). $K_a$ = association rate constant and $K_d$ = dissociation rate constant.

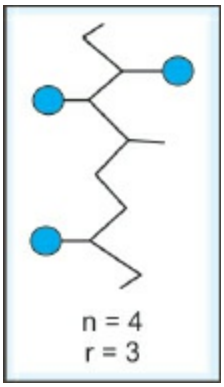



**Fig. 10.6:** Binding sites on protein molecule

Transforming Eq. 8 into linear form, we get

$$r + rK_a[D] = nK_a[D] \quad \text{... (9)}$$

$$r = nK_a[D] - rK_a[D] \quad \text{... (10)}$$

This can be rearranged to give:

$$\frac{r}{[D]} = nK_a - rK_a \quad \text{... (11)}$$

Thus, plotting $r/[D]$ versus r should give a straight line. This is called a **Scatchard plot** (Fig. 10.7).

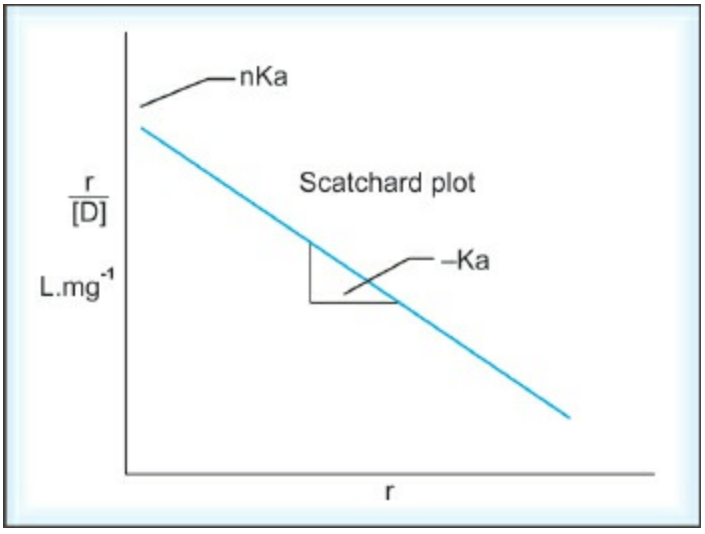



**Fig. 10.7:** Plot of $r/[D]$ Versus $r$

Alternate rearrangement gives:111

$$\frac{1}{r} = \frac{1}{nK_a[D]} + \frac{1}{n} \quad \text{... (12)}$$

Thus, a plot of 1/*r* vs 1/[*D*] should also give a straight line. This is the **double reciprocal plot** (Fig. 10.8).

With **one type of binding site**, these plots produce straight lines which can be used to determine $K_a$ and *n* values. With more than one type of binding site, these plots are curved.

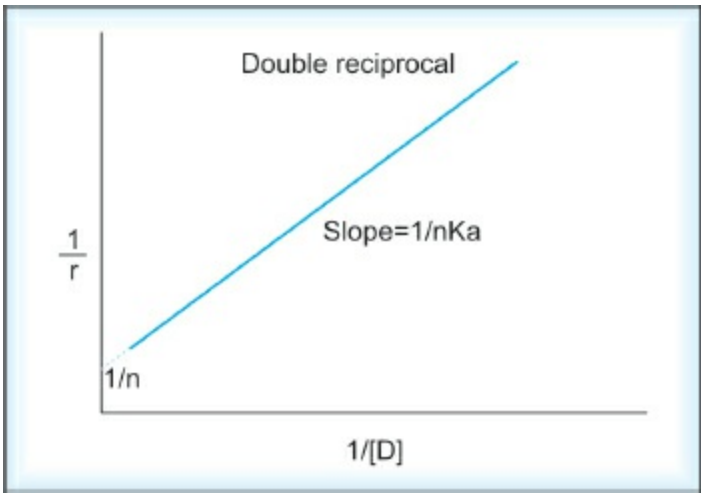



**Fig. 10.8:** Plot of 1/*r* *Versus* 1/[*D*]

Drug-protein binding is of importance to the distribution equilibrium of drugs. Plasma proteins may exert a “buffer and transport function” in the distribution process. Only the free, non-protein bound fraction of a drug can leave the circulatory system and diffuse into the tissue. The equilibrium between free and bound drug acts as a “buffer” system since a relatively constant concentration of free drug can be maintained over a relatively long period of time due to dissociation of drug-protein complex. The transport function of plasma proteins is of importance for drugs of low aqueous solubility. Often, hydrophobic drugs are found to be considerably bound to plasma proteins and also to erythrocytes.

## METABOLISM (BIOTRANSFORMATION)

Most drugs are cleared from the body, at least partially, through biotransformation (metabolism). The term 'biotransformation' of drugs can be defined as the chemical conversion of drugs to other compounds in the body. However, this definition excludes degradation due to any inherent chemical instability of drugs in biological media. Biotransformations are usually mediated by enzymes in the body's organs, tissues, or biofluids, but occasionally they may occur by non-enzymic reactions and even by a combination of both enzymatic and non-enzymatic processes. Biotransformation reactions may be catabolic (e.g. hydrolysis, reduction, and oxidation) or anabolic (e.g. conjugation with glucuronic acid, glycine, or sulfate). The maximum fraction of the dose that may be metabolized is given by the relative difference between the plasma and renal clearances of the drug, $(C_L\text{-}Cl_r)/C_L$. Metabolism, unlike excretion, does not result in the removal of drug mass from the body. Rather, new chemical entities are formed, and the distribution, metabolism, and excretion of each metabolite are unique and usually independent of the parent compound. When a metabolite possesses intrinsic pharmacological activity, as with the N-demethylated metabolite of a number of antiepileptic drugs, characterization of its pharmacokinetic profile as well as its formation kinetics is essential to a complete understanding of the pharmacokinetic/pharmacodynamic relationships.

The numerous biotransformation reactions are conventionally classified into two main types: ***Phase I and Phase II*** (Fig. 10.9), that generally occur sequentially and the combination of several phase I and phase II reactions yield a range of metabolites. Phase I reactions introduce or reveal a functional group (e.g. oxidation, reduction or hydrolysis) within substrate that can serve as a site for subsequent conjugation (i.e. Phase II). Phase I enzymes are found in several subcellular compartments including the cytoplasm and mitochondria. Reductases which metabolize quinine, nitro-and keto-functional groups can be found associated with the cytoplasm and endoplasmic reticulum. Microsomal monooxygenases such as flavin containing micro-genases and more importantly cytochrome P450 (CYP450) are quantitatively the most important oxidative enzymes for drug biotransformation. CYP450 represents the superfamily of haem-containing

proteins that catalyse the metabolism of many lipophilic and endogenous substances and exogenous compounds. CYP4A4, 2D6, 2C9, 2C19 are involved in oxidative biotransformation of majority of xenobiotics used therapeutically in humans. There is considerable interindividual variability in the hepatic expression of CYP450. For any individual, the pathway and the rate of a given compound metabolic clearance is a function of that individual's unique phenotype w.r.t. the forms and amounts of P450 species expressed. Interindividual variability in drug biotransformation activity is a consequence of complex interplay among genetic (genotype, phenotype, gender, race, stroke, race/ethnic background), environmental (diet, disease, concurrent medication, xenobiotic exposure) and developmental (somatic growth in children, adolescents) factors. The term pharmacogenetics was proposed to describe the study of genetically-determined variations in drug response. Phase II reactions further increase water solubility, and thereby facilitate elimination. Phase II enzymes include the glucouronosyl transferases (UGTs), sulpho-transferases and acetyl transferases, glutathione S-tranferases and methyl transferases.

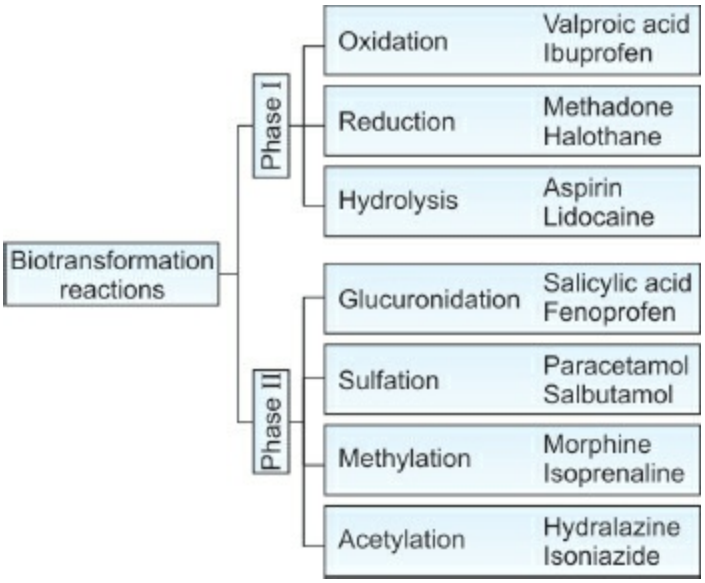



**Fig. 10.9:** Figure depicting examples of some drugs undergoing phase I and phase II reactions

## Stereoselectivity in Drug Metabolism

Since majority of the chiral synthetic or semisynthetic drugs are marketed as racemates, consideration must be given to the potential stereoselective metabolism of chiral xenobiotics. Both phase I and phase II metabolic reactions are potentially capable of discriminating between enantiomers. For drug enantiomers with low hepatic extraction, a difference in intrinsic clearance will be directly reflected in the hepatic clearance. This in turn results in stereoselectivity in the resultant plasma concentration after both oral and IV drug administration. In contrast, for highly extracted enantiomers, a difference in the intrinsic clearance may not result in a significant alteration in stereoselectivity of plasma drug concentrations. Warfarin provides perhaps the best example of some of the clinical implications of stereoselective drug metabolism.

## Chronopharmacokinetics and Drug Metabolism

Diurnal variations have also been shown to impact the activity of selected phase I and phase II enzymes capable of metabolizing xenobiotics. In the case of acetaminophen conjugation, diurnal variation in the activity of glucuronosyltransferase and sulfotrans-ferase isoforms appears to explain the chrono-pharmacokinetics of this compound and perhaps other drugs as well. The basic mechanism by which diurnal variation alters the activity of drug metabolizing enzymes are not completely understood but may involve neuroendocrine regulation. Also, other physiological alterations known to have a circadian pattern could potentially impact drug metabolism by altering the free fraction of drug available for hepatic clearance and/or the efficiency of presentation of xenobiotics to the liver.

## First-pass Effect

Drug absorbed from the stomach, duodenum, jejunum, ileum or following deep rectal administration is carried by the total circulation to the liver where it is subjected to biotransformation before reaching the systemic circulation. Substantial drug biotransformation may also occur in the GI epithelium during the course of absorption. The process of drug biotransformation prior to becoming systemically available is referred to as first-pass effect (FPE) or pre-systemic clearance and represents one means by which systemic availability of drugs is reduced. Incomplete release from the dosage form, chemical degradation, physical complexation, microbial transformation and accelerated GI transit may also contribute to FPE of drug and hence be responsible in part for poor oral bioavailability. The quantitative importance of FPE depends on the rate of drug absorption, rate of biotransformation and biotransformation capacity for a particular drug. In its simplest form, the extent of FPE can be estimated by comparing AUC-time curve following i.v and oral administration after correcting for the doses administered assuming complete absorption. By determining AUC-time curve for the parent drug and its metabolite, it is possible to differentiate between FBE and incomplete absorption. For example if absorption is complete but there is a FPE, then the ratio of AUC(p.o.) to AUC(IV) for the parent drug would be less than unity, while the AUC(p.o.) by AUC(IV) for the metabolite would be approximately one. Conversely, if there is incomplete absorption and no FPE, AUC(p.o.) to AUC(IV) for the parent compound and metabolite would be approximately the same and both will be less than one.

## EXCRETION

Excretion of the drugs is the final loss of drug from the body. Drugs may be eliminated from the systemic circulation by different pathways, i.e. via urine, bile, intestines, saliva, alveolar air, sweat and milk. The two major pathways of excretion are via the kidney into the urine and via liver into the faces. The major organ for the excretion of drugs is the kidney. The mammalian kidney is composed of many units called *nephrons* (Fig. 10.10). A nephron consists of an individual renal tubule and its glomerulus. The glomerulus is formed by the invagination of a tuft of capillaries into the dilated blind end of the nephron (Bowman's capsule). The capillaries are supplied by an afferent arteriole and drained by an efferent arteriole. The renal tubule consists of the proximal convoluted tubule (pars convoluta), which drains into the straight portion of the proximal tubule (pars recta), which forms the first part of the loop of Henle. The thick ascending limb of the loop of Henle reaches the glomerulus of the nephron from which the tubule arose and passes close to its afferent arteriole. The final portion of the tubule is the distal convoluted tubule. Distal tubules coalesce to form collecting ducts that pass through the renal cortex and medulla and empty into the pelvis of the kidney. The vascular supply of the tubules is essentially a portal one in that the blood that perfuses the peritubular capillaries has initially traversed the glomerular capillaries.

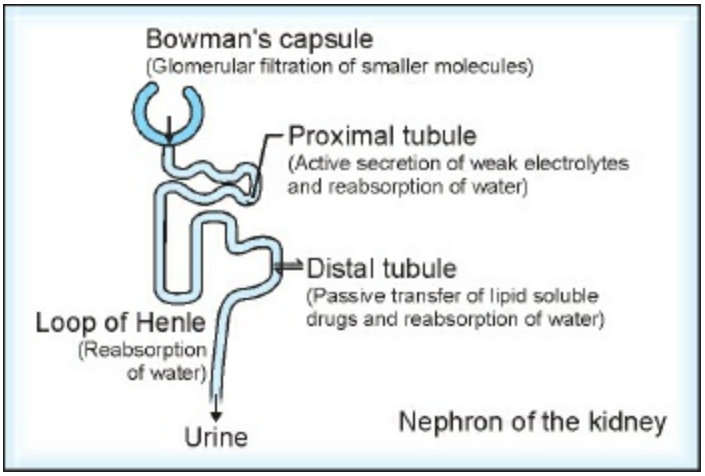



**Fig. 10.10:** Structure of nephron

## Renal Excretion Mechanisms

The nephrons carry out the vital functions of the kidney through three major processes: glomerular filtration, tubular reabsorption of water and filtered substances from the lumen of the tubule into the plasma, and tubular secretion of substances from the plasma across the tubular membrane into the tubular lumen. The processes of tubular reabsorption and secretion may be governed by either passive mechanisms (simple diffusion) or carrier-mediated mechanisms (facilitated diffusion and active transport). While the renal processes serve primarily to maintain the extracellular fluid volume and osmolality, to conserve important solutes, and to help regulate acid-base balance, they are also important in controlling the excretion of exogenous compounds such as drugs.

### *Glomerular Filtration*

Approximately 25% of the cardiac output goes to the kidneys, 10% of which is filtered by the glomerulus. A fluid conforming closely to that of an ideal ultrafiltrate of plasma moves across the glomerular membrane. It has nearly the same composition as plasma w.r.t. water and low molecular weight solutes. Glomerular filtration separates gross particulate matter and such colloidal materials as proteins from the ultrafiltrate. The availability of a drug for glomerular filtration depends on its concentration in plasma and only free drug (drug not bound to macromolecules or red blood cells) can be filtered. The filtrate contains the drug at a concentration equal to that in plasma water, i.e. $f_u C_1$.

The rate at which plasma is filtered is called the *glomerular filtration rate* (GFR) and the rate at which a drug is filtered is equal to the concentration of unbound drug in the plasma multiplied by GFR. The GFR can be measured in intact animals and humans by measurement of the excretion and plasma concentration of a substance that is freely filtered through the glomeruli, and is neither secreted nor reabsorbed by the tubules. In addition, the substance should be physiologically-inert and non-toxic; neither destroyed, synthesized nor stored within the kidney and preferably, easily measured in plasma and urine. Inulin, a fructose polymer of molecular weight 5,200 daltons, appears to meet all the criteria, and its renal clearance provides an index of GFR.

## *Passive Transport across the Renal Tubule*

Passive transport of exogenous compounds, such as drugs, in the renal tubule involves simple diffusion along a concentration gradient. The concentration gradient between urine and plasma is the driving force for diffusion. The rate of movement is also governed by the diffusivity of the molecule through the tubular membrane, the membrane/aqueous phase partition coefficient for the molecule, the thickness of the membrane at the site of diffusion, and the area of the membrane through which the molecule passes. Biological membranes, being lipoidal in nature, are more permeable to lipid-soluble substances, and transmembrane diffusion depends in part on the lipid solubility of the diffusing compound. In the case of acids and bases, the un-ionized species exhibit greater lipid solubility than the ionized species, and is either the sole diffusing species or the more rapidly diffusing species. The diffusing species of an acid or base is governed by the concentration gradient across the membrane and the *pKa* of the compound. In the specific case of passive diffusion across the renal tubular membrane, the concentration gradient depends on urinary pH since intracellular and blood pH are essentially constant.

Diffusion across the tubular membrane is also affected by flow-related changes in urinary concentration. The observation that excretion of $N^1$-methylnicotinamide (a quaternary ammonium compound) is enhanced by increases in urinary flow suggests that ionized substances diffuse passively across the renal tubular membranes.

Although diffusion across the tubular membrane may occur in either direction for some organic bases, diffusion from blood to the tubular lumen is highly improbable for organic acids. Passive reabsorption from filtrate to blood occurs for many drugs. The urinary excretion rate of amphetamine (*pKa* 9.77) fluctuates with changes in urinary pH as the total amount excreted under alkaline urine conditions is lower than under acidic urine conditions. The renal clearance of phenobarbital (*pKa* 7.2) increases with increasing urinary flow. At any given rate of flow, the clearance is higher when the urine is alkaline than when it is acidic. In general, a weak base whose *pKa* is 6 or below is expected to be extensively reabsorbed at all urinary pH values. Little or no reabsorption is expected for a strong base whose *pKa* is close to 12 throughout the range of urinary pH. Reabsorption is expected to vary with changes in pH for bases with *pKa* values of 6 through 12. Reabsorption of

acidic drugs with *pKa* values between 3 and 7.5 varies with urinary pH; acids with *pKa* ≤2 are not reabsorbed, and those with *pKa* > 8 are extensively reabsorbed throughout the range of urinary pH. Urinary flow rate often affects the excretion of compounds whose tubular reabsorption is pH-sensitive.

## *Carrier-mediated Transport Across the Renal Tubule*

*Carrier-mediated transport* is a term used to describe transfer across a biological membrane that is at a higher rate than could be attributed to diffusion alone. Such transport may involve either facilitated diffusion or active transport. The term *active transport* is usually applied only to those systems in which a substance is transported across a biological membrane against a concentration gradient at the expense of energy derived from cell metabolism. Active transport processes figure prominently in the renal excretion of many drugs and their metabolites. These processes are characterized by (1) susceptibility to interference by metabolic or competitive inhibitors and (2) a maximal capacity of the transporting mechanism ($T_m$). Although some organic ions undergo simultaneous active bidirectional transport, i.e., reabsorption and secretion, the predominant transport of most organic ions in the renal tubules is secretory. Thus, active secretion usually adds substances to the filtrate. Substances added to the filtrate in this manner may be subsequently reabsorbed. Active secretion occurs in the proximal portion of the renal tubule, and reabsorption can occur all along the tubule.

The existence of renal tubule transport systems for the secretion of organic acids and bases is well documented. These systems involve separate and independent mechanisms, and are not subject to inhibition by the same competitive inhibitors. There is not a high degree of specificity within either systems. Substances transported by the same system compete with each other. Probenecid was synthesized in the early 1950s as an inhibitor of renal transport mechanisms. It has since been shown that probenecid itself is actively secreted by the organic acid mechanism. It has been used as the classic inhibitor of the organic anion transport system, and its ability to inhibit secretion of a compound is taken as evidence for secretion of that compound through the acid transport system.

$N^1$-methylnicotinamide is the classic example of an organic base that is actively transported by the renal tubule. It was one of the earliest bases for

which secretion was demonstrated, and it has been used throughout the years to define the characteristics of the organic cation transport system. For drugs cleared by tubular secretion, it makes no difference what fraction is bound to plasma protein, provided that the binding is reversible. Secretion can be so extensive that all of the drug, whether in red blood cells or bound to plasma proteins, can be removed. Only the unbound drug is capable of crossing the cells lining the tubule; transport of all drugs out of the blood requires rapid dissociation of the drug-protein complex and movement of drug out of the blood cells.

A compound may be subject to all of the renal excretion processes. Most organic acids and bases, however, are thought to be excreted by a three-component mechanism: glomerular filtration, active secretion, and passive reabsorption. The amount of a substance excreted by the kidney is equal to the amount filtered by the glomerulus plus the net amount transferred by tubules. Observed renal clearance, therefore, essentially consists of two components: glomerular filtration and net tubular transport. If a drug is not bound to macromolecules in the plasma, then its renal clearance either equals glomerular filtration (GFR) if there is no net tubular secretion or reabsorption, exceeds GFR if there is net tubular secretion, or is less than GFR if there is net tubular reabsorption.

While a renal clearance equal to GFR indicates the absence of a net tubular involvement, it does not preclude the involvement of compensating tubular processes. When a compound is bound to macromolecules, the degree of binding must be determined at relevant concentrations before meaningful statements can be made about its renal clearance. When a compound is excreted by glomerular filtration with possible passive tubular reabsorption, renal clearance should be calculated from the non-protein-bound fraction in plasma. For a compound with renal clearance higher than GFR, e.g. one involving tubular secretion, the total plasma concentration of the compound is used in the calculation of clearance. In those experimental situations whereby the contributions of glomerular filtration and tubular secretion to the renal clearance of a compound are separable, the filtration contribution should be corrected for protein binding in all cases.

Changes in renal clearance may result in changes in plasma half-life ($t_{½}$), in the overall elimination rate constant, $k_{10}$, in the fraction of renal elimination of a compound ($f_r$), and in the fraction metabolized ($f_m$). A

decrease in renal clearance (e.g. increased reabsorption, inhibition of secretion, decreased glomerular filtration) may result in an increase in $t_{½}$ and a decrease in $k_{10}$. For a drug that is metabolized, as well as excreted unchanged in the urine, decreased renal clearance may also result in a larger fraction of drug metabolized ($\uparrow f_m$) and a smaller fraction recovered in the urine unchanged ($\downarrow f_r$). Opposite changes in the respective parameters would occur for an increase in renal clearance.

## Experimental Techniques

A number of techniques have been employed for studying renal processes both in vitro and in vivo. In vitro techniques used to examine renal tubular transport mechanisms include renal slices, suspensions of renal tubules, perfused isolated renal tubules, and perfused isolated kidneys. In vivo methods for studying renal excretion processes include clearance techniques, the stop-flow technique, micropuncture, and micro-perfusion. Clearance techniques have the advantages of being technically simple, of enabling quantification of kidney function as a whole, and of being applicable to humans. Limitations include the inability to separate reabsorption from secretion for substances undergoing both processes and the inability to define transport mechanisms or to localize function to specific nephron segments.

In the clearance techniques used to study the renal excretion of a compound, an appropriate marker for GFR (e.g. inulin) is infused simultaneously with the compound. Blood and urine samples are collected and analyzed for the reference material and compound being studied. Renal clearances of each are then calculated. When the ratio of the clearance of the compound being studied to that of the reference is greater than 1.0, net secretion is indicated; when the ratio is less than 1.0 (after correcting for plasma protein binding), net reabsorption is indicated. A change in the ratio with a change in urinary pH implies some passive transport, as does a change in ratio with changing urinary flow rates. A change in the ratio with increasing plasma concentration of the study compound, or in the presence of a competitive or metabolic inhibitor of transport, indicates the presence of a carrier-mediated transport process. Urinary excretion of the study compound due to glomerular filtration is equal to the unbound plasma concentration of the compound multiplied by the renal clearance of the marker (GFR); net tubular transport is equal to the total urinary excretion minus excretion due to glomerular filtration.

In the absence of active reabsorption processes, appropriate adjustment of urinary pH to preclude significant passive reabsorption of a compound enables quantitation of the secretory component (i.e., “net tubular transport” equals secretion). If there is no passive diffusion of compound from plasma to the tubular lumen, the secretory component can be further evaluated in terms of saturation (Michaelis-Menten) kinetics.

## Enterohepatic Circulation

### *Biliary Excretion*

Drug elimination from the liver can occur by two distinct mechanisms: Hepatic metabolism and biliary excretion. Hepatic clearance is the a quantitative measure of the overall ability of the liver to eliminate the drug and is the sum of the hepatic metabolic clearance and the biliary clearance.

While all compounds may be excreted in bile, the importance of biliary clearance to the elimination of a compound depends on the degree to which it concentrates in bile relative to plasma. Biliary clearance is given by Eq. (13) as follows:

$$\text{Biliary clearance} = \text{Bile flow} \times \frac{\text{Concentration in bile}}{\text{Concentration in plasma}} \quad \ldots (13)$$

Since bile flow in man is relatively constant—between 0.5 and l.0 ml/min—biliary clearance is proportional to the ratio of bile to plasma drug concentration. Compounds are classified according to the degree to which they concentrate in bile. Biliary clearance cannot exceed bile flow for such compounds as electrolytes and proteins, whose concentration ratios of bile to plasma are usually equal to or less than unity. On the other hand, a biliary clearance of 500 ml/min is not uncommon for drugs.

For the biliary clearance of a compound to be significant, the compound must be actively secreted into bile and achieve a concentration gradient relative to the blood. Separate secretory mechanisms appear to exist for acids, bases, and neutral compounds. Several physicochemical features of a molecule are important in determining the extent to which it is secreted in bile. First, compounds secreted in bile usually have molecular weights exceeding 300 to 400 daltons. Moreover, the molecular weight threshold appears to depend on species: about 325 daltons in the rat, 400 daltons in the guinea pig, 475 daltons in the rabbit, and 500 daltons in man. Second, compounds excreted in bile are usually polar in nature. Molecular structure may also be important, but the nature of dependence is not well understood.

The knowledge of the metabolic profile of a drug is essential in assessing whether a given compound may be excreted in bile. Conjugation reactions increase the polarity as well as the molecular weight of the drug. Glucuronide

conjugates, for instance, are strong acids with *pKa* values of 3 to 4, are nearly completely ionized at physiologic pH, and have molecular weights 176 daltons greater than the parent compound. Not surprisingly, therefore, many compounds are excreted in bile in their conjugated forms.

Many of these generalizations are illustrated by the data from two non-steroidal antiinflammatory agents, indomethacin and sulindac. Indomethacin is pharmacologically active whereas sulindac is a prodrug that is metabolized reversibly to the active moiety, the sulfide metabolite. Sulindac is also irreversibly biotransformed to the inactive sulfone metabolite. Structures and molecular weights are given in Fig. 10.11. Both are arylacetic acids with molecular weights of about 350 daltons. In addition, these compounds undergo glucuronidation so that the effective molecular weight is about 525. Biliary and renal clearances of each in various animal species, including man, have been determined from the total (free plus conjugate) amount of drug present in bile and urine. The data, summarized in Table 10.3, are expressed in terms of the renal to biliary clearance ratios as an index of the relative importance of the two routes of elimination.

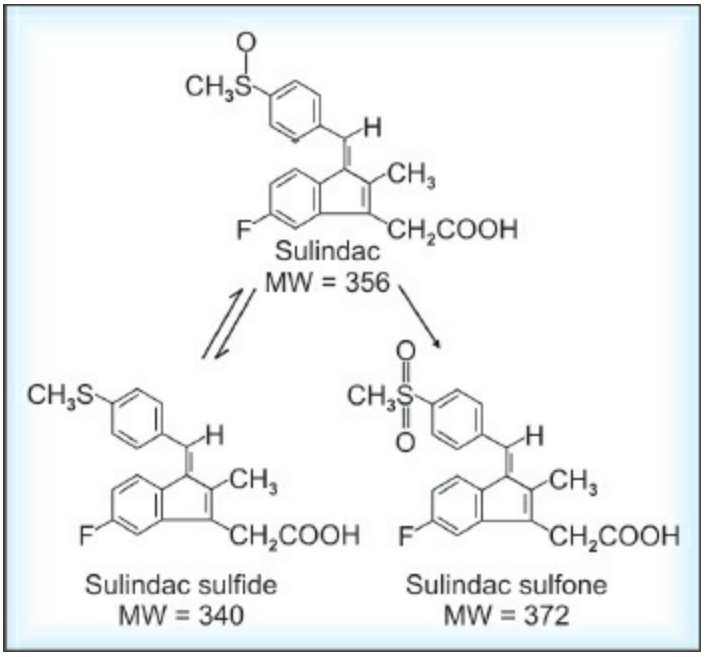



**Fig. 10.11:** Sulindac and metabolites

**Table 10.3:** Renal to biliary clearance ratios in various species

| Species | Renal to biliary clearance ratio | | | |
|---|---|---|---|---|
| | Indomethacin | Sulindac | Sulindac sulfone | Sulindac sulfide |
| Dog | <0.008 | <0.004 | — | — |
| Rat | 0.03 | <0.03 | — | — |
| Rhesus monkey | 1.36 | 0.11 | 0.11 | ~0 |
| Man | 0.73 | 0.12 | 0.10 | <0.03 |
| Rabbit | 2.73 | 1.65 | 2.5 | 2.1 |

A substantial inter-species difference is evident for both indomethacin and sulindac. Biliary clearance is, by far, the dominant route in the dog and rat, while renal clearance is slightly favoured in the rabbit. Based on molecular weight alone, a species difference is expected; the conjugated compounds exceed the “threshold” molecular weight by approximately 200 daltons in the rat but by approximately only 50 daltons in the rabbit.

The renal-to-biliary clearance ratios are of similar magnitude in the rhesus monkey and man, especially for sulindac and metabolites, and are intermediate between those of the rat and rabbit. Only minimal amounts of the sulfide metabolite are excreted in the urine.

Also, sulindac sulfide is an interesting example of structural dependence: while it differs from the parent drug and the sulfone metabolite only in the oxidation state of the sulfur atom, its clearance into bile (or urine) of man is marginal compared to either sulindac or its sulfone metabolite.

## *Biliary Recycling*

Following its secretion into bile, the drug is stored in the gallbladder. When the gallbladder contracts, the drug is released into the duodenum and may then be metabolized, reabsorbed, or excreted in feces. If reabsorbed back into the portal circulation, it once again is subject to biliary secretion in the liver and thus completes an “enterohepatic cycle” (Fig. 10.12). Biliary clearance is a route of drug elimination from the body only to the extent that drug is excreted in feces, bio-transformed, or otherwise degraded in the intestinal lumen (i.e. it undergoes irreversible clearance). If drug is reabsorbed, the hepatoportal system is simply an organ of drug distribution (i.e., reversible clearance). As indicated previously, compounds are often cleared in bile as their conjugates. Commonly, these are hydrolyzed within the lumen by the gut flora liberating the original drug, which is then free to be reabsorbed.

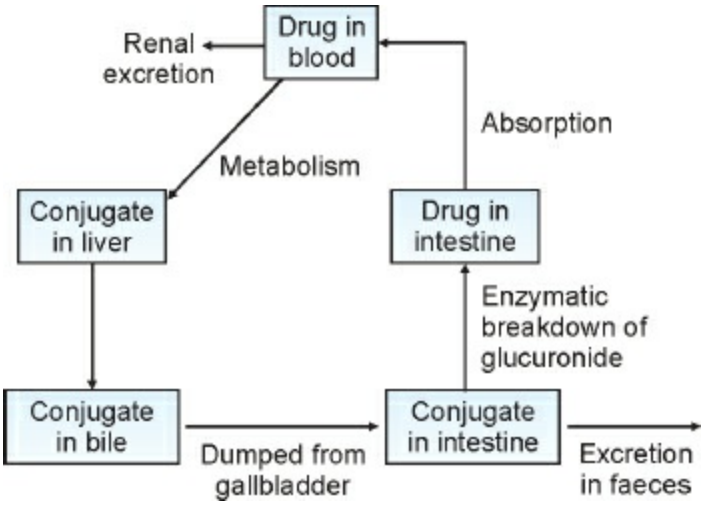



**Fig. 10.12:** Enterohepatic cycling

With the notable exception of the rat, all common laboratory animals and humans have a gallbladder. Contraction of the gallbladder occurs only intermittently, usually in response to food stimuli, so that considerable amounts of drug may accumulate in the gallbladder. Since gallbladder emptying is sporadic, different amounts of drug are released into the gut at uneven time intervals. If the drug is reabsorbed, gallbladder emptying serves as an additional source of drug input into the body. The overall input function is complex in that neither the amount nor the time course of reabsorption is known, and reabsorption may overlap with part of the dose being absorbed for the first time. Therefore, drugs that are recycled tend to exhibit unusual pharmacokinetic properties. The overall effect of enterohepatic recycling of drugs is to delay its elimination from the body and to prolong its pharmacological effect.

The effect of recycling on drug elimination may be described quantitatively. Again, sulindac is used to illustrate this point. The biliary clearance, $CL_b$, of sulindac and its sulfone and sulfide metabolites was studied in man by the use of a duodenal intubation technique and by direct collection of bile through a surgically implanted T-tube. Results of both techniques were similar, with CLb averaging about 200 ml/min for both sulindac and its sulfone, and about 14 ml/min for the sulfide. The total amount of drug secreted in bile, $CL_b$-AUC„, as a percentage of the administered dose averaged 135% for sulindac, 186% for the sulfone metabolite, and 16.2% for the sulfide metabolite; the total biliary drug flux

was 336% of the administered dose. Thus, on an average, a given dose of sulindac is recycled approximately 3.4 times.

Drug recycling may appear as secondary reentry peaks in the plasma concentration profiles. In patients who have fasted, these peaks are typically seen 8 to 12 h after administration of the dose coincident with gallbladder emptying following resumption of eating with the evening meal (Fig. 10.13).

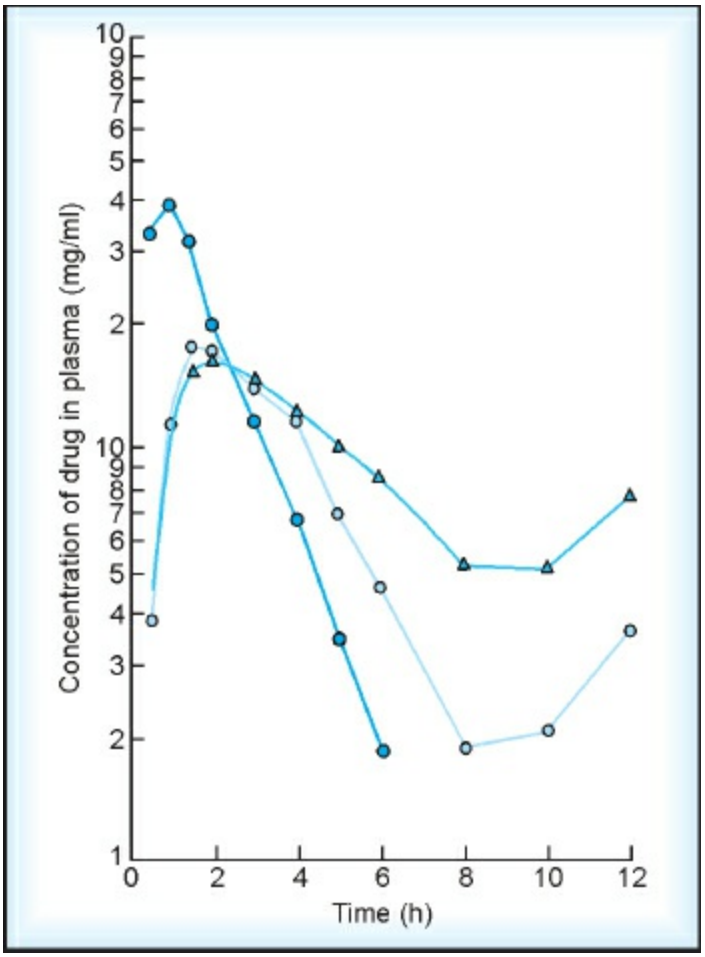



**Fig. 10.13:** Plasma-concentration profiles of sulindac ( ● ) and its sulfone ( ▲ ) and sulfide ( ● ) metabolites demonstrate secondary re-entry peaks due to biliary recycling in humans

## *Pulmonary Excretion*

The lung is the major organ of excretion for gaseous and volatile substances. In addition to gases, many volatile compounds such as alcohol and ethereal oils are also excreated via this route. The breath analyzer test is based on the

quantitative pulmonary excretion of ethanol.

## *Salivary Excretion*

This is not actually a method of drug excretion as the drug will usually be swallowed and reabsorbed, and thus a form of ‘salivary recycling’. Drug excretion into the saliva appears to be dependent on pH partition and protein binding. This mechanism appears attractive in terms of drug monitoring, that is, determining drug concentration to assist in drug-dosage adjustment. For some drugs, the saliva/free plasma ratio is fairly constant. The saliva/plasma free drug concentration ratio will be less than unity for weak acids and will exceed unity for weak bases. Therefore, drug concentrations in saliva could be a good indication of drug concentration in plasma. For some drugs, localized side effects may be due to salivary excretion of the drug. The use of saliva concentrations is an ideal method for therapeutic drug monitoring because it is noninvasive, a large number of samples can be obtained, and is particularly useful in ambulatory, pediatric and geriatric patients.

## *Mammary Excretion*

Studies during the past two decades have revealed that many drugs pass into the milk and even attain a higher concentration in milk than in plasma. The transport mechanism primarily is passive diffusion of the unionized moiety. Extent of drug transfer into the milk depends upon the drug’s *pKa*, pH difference between milk and plasma, amount of drug in maternal blood, and lipid solubility of the unionized drug. Since the pH of human milk is more acidic (pH 6.6) than the pH of plasma (pH 7.4), the milk-plasma ratio is higher for weak bases and lower for weak acids. Drugs appear in milk in much higher concentrations if the nursing mother has decreased renal excretion.

## *Excretion of Drugs into Sweat*

Excretion of drugs into sweat follows passive diffusion of the un-ionized moiety. Weakly acidic drugs such as sulphonamides, salicylic acid and benzoic acid; and weakly basic drugs such as thiamine; and metals such as iodine, bromine, mercury and iron have been found to be excreted into sweat.

## CLEARANCE (CL)

Clearance is a measure of the ability of the body to eliminate a drug, and as such is one of the most important pharmacokinetic parameters, as it gives a well-defined, physiologically relevant measurement of how drugs are eliminated by the organism as a whole, or by a particular organ. Clearance relates drug concentration to the elimination rate from the body Eq. (14):

$$CL = \text{Elimination rate}/C \quad \ldots (14)$$

or at steady state the average concentration $C_{ss}$ to the dosing rate because at steady state the input rate into the body will equal the output rate:

$$CL = \text{Dosing rate}/C_{ss} \quad \ldots (15)$$

The clearance concept is based on physiology, where it is used as a measure of the renal function (creatinine clearance). Creatinine is formed from muscle breakdown at a constant rate, and thus a constant creatinine concentration in plasma results. The magnitude of this concentration is dependent on the elimination rate of creatinine and the size of the muscle pool (formation rate). By measuring the plasma concentration and the renal excretion of creatinine, renal clearance can be estimated, and thereby kidney function indicated, as creatinine is mainly filtered into the urine Eq. (16):

$$CL_{creatinine} = \frac{\text{Urine volume} \times \text{Urine concentration}}{\text{Plasma concentration}} \quad \ldots (16)$$

Drugs are not only eliminated via the kidneys, but also eliminated in the bile by the liver and metabolized in the liver and elsewhere, which makes direct measurement of the elimination rate of a drug difficult. Indeed, other routes of elimination could include loss in expired air, saliva, sweat, partition into tissue stores, efflux from the blood into the gut lumen, and gut metabolism as well as other sites of metabolism such as the lung. The total clearance, *CL*, can be defined as:

$$CL = \frac{Dose_{IV}}{AUC} \quad \ldots (17)$$

where, $Dose_{IV}$ is the intravenous dose, and *AUC* is the resulting area under the plasma concentration.

Clearance is referenced to plasma (plasma clearance, $CL_p$), blood (blood clearance, $CL_b$) or plasma water (unbound clearance, $CL_u$), depending upon where the concentration is measured. Total clearance can be divided into the contributions of each of the eliminating organs, the most important being renal clearance, $CL_R$, and hepatic clearance, $CL_H$.

$$CL = CL_R + CL_H + CL_{gut} + CL_{other} \quad ...(18)$$

One method of quantitatively describing the renal excretion of drugs is by means of the renal clearance value of drug. Renal clearance can be calculated as part of the total body clearance for a particular drug. It can be used to investigate the mechanism of drug excretion. If the drug is filtered but not secreted or reabsorbed, the renal clearance will be about 120 ml/min in normal subjects. If the renal clearance is less than 120 ml/min, then we can assume that at least two processes are in operation, glomerular filtration and tubular re-absorption. If the renal clearance is greater than 120 ml/min then tubular secretion must be contributing to the elimination process. It is also possible that all the three processes are occurring simultaneously and the renal clearance is then defined as:

$$CL_{renal} = \frac{\text{Filtration rate + secretion rate − reabsorption rate}}{C_p} \quad ... (19)$$

Renal clearance values can range from 0 ml/min, the normal value for glucose which is usually completely reabsorbed, to a value equal to the renal plasma flow of about 650 ml/min for compounds like p-aminohippuric acid. If a drug is extensively excreted unchanged into urine, alteration of renal function will alter the drug elimination rate. Fortunately, creatinine clearance can be used as a measure of renal function. For most drugs which are excreted extensively unchanged, it has been found that there is a good correlation between creatinine clearance and drug clearance or observed elimination rate. Clearance is measured in units of volume per time (ml/min or L/h) and thus is defined using the same units as blood or plasma flow. By definition, clearance gives the volume of plasma (blood) from which a drug is completely removed per unit time.

## Intrinsic Clearance

Although the liver is the major metabolic organ in the body, significant metabolic activity may also exist in the lungs, kidneys, blood, or gut mucosa. A drug may be metabolized by competing pathways in the liver. The overall hepatic clearance of a drug is then given by the summation of clearances by each pathway. The ability of the liver to metabolize a drug depends in part on the intrinsic activity of the enzymatic system associated with each metabolic pathway. The overall intrinsic reaction velocity, $V_{int}$, may be expressed in terms of the familiar Michaelis-Menton equation applied to each of the $N$ enzymatic systems involved in hepatic drug removal:

$$V_{int} = \Sigma_{i=1}^{N} \frac{V_{m,i} C_{u,1}}{K_{m,i} + C_{u,1}} \qquad \text{... (20)}$$

where, $C_{u,1}$ is the unbound concentration of drug available to the liver enzymes, $V_m$ is the maximum reaction velocity, and $K_m$ is the Michaelis constant, which is equivalent to the $C_{u,1}$ at which $V_{int} = V_m/2$. Since the instantaneous rate of drug elimination is equal to the product of its clearance and concentration, the intrinsic hepatic clearance, $CL_{int}$ is given by:

$$CL_{int} = \frac{V_{int}}{C_{u,1}} = \Sigma_{i=1}^{N} \frac{V_{m,i}}{K_{m,i} + C_{u,1}} \qquad \text{... (21)}$$

According to Eq. (21), if $C_{u,1}$ is significant relative to the value of $K_m$, then $CL_{int}$ is not constant but varies with the dose of drug administered. The $CL_{int}$ of either enantiomer of propranolol, for instance, is reduced more than 50% when the dosage of the racemate is increased from 160 to 320 mg/day in humans. For most drugs, however, $C_{u,1}$ is small relative to $K_m$ over the therapeutic dosage range and Eq. (21) reduces to Eq. (22):

$$CL_{int} = \Sigma_{i=1}^{N} \frac{V_{m,i}}{K_{m,i}} (C_{u,1} \ll K_{m,i}) \qquad \text{... (22)}$$

where, $CL_{int}$ is constant and independent of the dose of drug administered.

In addition to the intrinsic metabolic activity of the liver, delivery of drug to the liver (i.e. perfusion) may be an important determinant of the liver's ability to metabolize the drug: hepatic drug clearance cannot exceed the rate at which the drug is delivered to the liver. Flow rate to the liver may limit full

expression of the intrinsic metabolic enzyme activity of the liver, and hepatic drug clearance is then lower than $CL_{int}$. A commonly used model that relates blood flow and $CL_{int}$ to hepatic drug clearance ($CL_h$) is the perfusion-limited, or well-stirred, compartment model of hepatic clearance.

$$CL_h = Q_h\left(\frac{f_u CL_{int}}{Q_h + f_u CL_{int}}\right) \quad \text{... (23)}$$

where, $f_u$ is the fraction of unbound drug in blood, $Q_h$ is the hepatic blood flow rate (normally ~1.5 L/min in man), and $CL_h$ is the hepatic drug clearance from blood. Eq. (23) is defined in terms of blood since this is the fluid that carries drug to the eliminating organ. If measurement of drug in blood is not experimentally feasible, the blood-to-plasma concentration ratio can be determined and used to convert plasma concentration to blood concentration.

The quantity in Eq. (23) equivalent to $CL_h/Q_h$ is referred to as the *hepatic extraction ratio, E,* of the drug. The effects of changes in $Q_h$ and $f_u\ CL_{int}$ on $CL_h$ are readily evident if compounds are classified on the basis of their extraction ratio. If a drug is highly extracted ($E \rightarrow 1$), then $f_u\ CL_{int} >> Q_h$, and Eq. (23) reduces to:

$$\lim_{E \rightarrow 1} CL_h = Q_h \text{ ... (24)}$$

Hepatic drug clearance is then insensitive to intrinsic enzymatic activity or binding and is approximated by hepatic blood flow rate, i.e. it is perfusion-limited. On the other hand, if the drug is poorly extracted ($E \rightarrow 0$), then $Q_h >> f_u CL_{int}$ and Eq. (23) reduces to Eq. (25).

$$\lim_{E \rightarrow 0} CL_h = f_u CL_{int} \text{ ... (25)}$$

Hepatic drug clearance is insensitive to changes in $Q_h$ but sensitive to changes in both drug binding to blood components and metabolic enzyme activity in the liver. Thus, induction of hepatic enzyme activity, for instance, would increase the hepatic clearance of a low extraction compound but have little effect on a high extraction compound. Finally, the hepatic clearance of a drug with an intermediate extraction ratio (i.e. $E$ ~ 0.2 to 0.8) is affected by changes in $Q_h$, $f_u$, or $CL_{int}$.

The above relationships are exemplified by the block diagram in Fig.

10.14. Shown is the effect on extraction ratio (block height) and hepatic drug clearance of a ± 0.5 L/min change from normal hepatic blood flow for a compound with $CL_{int}$ equal to 0.1, 1, or 10 times the normal hepatic blood flow. In all cases, the drug is not bound in blood, i.e. $fu$ = 1. When $CL_{int}$ is high (15 L/min) relative to blood flow, drug is highly extracted by the liver, and $E$ = 0.94 when $Q_h$ = 1 L/min. Substantial increase in $Q_h$ result in only a slight decrease in $E$. Hepatic clearance, on the other hand, increases almost proportionately with $Q_h$ and approaches $Q_h$ as $E$ approaches unity, as shown in Eq. (24). In contrast, when $CL_{int}$ is low (0.15 L/min) relative to blood flow, the drug is poorly extracted by the liver, i.e. $E$ = 0.13 when $Q_h$ = 1 L/min. In this case, as blood flow increases, there is a proportionate decrease in extraction ratio. Hepatic clearance is approximately equal to $CL_{int}$ and is independent of blood flow, as shown in Eq. (25). For the intermediate case where $CL_{int} > Q_h$ (1.5 L/min), there is a modest decrease in $E$ and increase in hepatic clearance as blood flow increases.

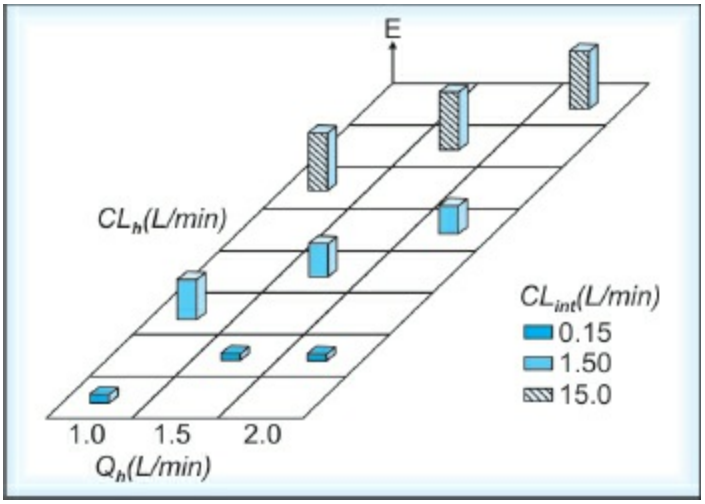



**Fig. 10.14:** Relationship between extraction ratio ($E$), blood flow ($Q_h$), and hepatic drug clearance ($CL_h$) for a compound with an intrinsic clearance ($CL_{int}$) equal to 0.1, 1, or 10 times normal hepatic blood flow.

## Capacity Limitations

Drug removal from the body may be excretory or metabolic. Either route may be further composed of multiple parallel pathways, each with its own inherent capacity to remove the drug. The overall elimination rate of the drug from the body is the sum of these individual contributions. The Michaelis constants, $K_m$, for most elimination processes are large relative to the concentrations of drug following therapeutic doses, as was shown by Eq. (22). Hence, the rate of drug elimination is usually proportional to the concentration. The proportionality constant is the plasma clearance ($CL$). Drugs that behave in this way are said to obey linear elimination kinetics. Occasionally encountered, however, are drug concentrations that may not be insignificant relative to the $K_m$ for one or more processes, whereby the elimination rate becomes disproportionate in relation to concentration, and increasingly so with increasing dose or concentration. Drugs that behave non-linearly within the therapeutic range include salicylic acid (SA), phenytoin, theophylline, diflunisal, acetaminophen, and 5-fluorouracil.

The disposition of (SA) in humans is schematically depicted in Fig. 10.15. Levy and co-workers have shown that capacity limitations exist in the formation of salicylurate (SU). Deviations from linear kinetics become evident even after the administration of a single aspirin tablet. With increasing dosage, the metabolic pathway leading to the formation of salicyl phenolic glucuronide (SPG) also becomes non-linear. On the other hand, the renal excretion of SA and biotransformation to salicyl acyl glucuronide (SAG) and gentisic acid (GA) appears linear throughout the therapeutic range. Therefore, the overall rate of salicylate elimination from the body can be represented by Eq. (26), which is based on the one-compartment, open model (discussed later in this chapter).

$$-V_l\frac{dC_l}{dt}=\left(CL_m^{SAG}+CL_m^{GA}+CL_r+\frac{V_m^{SU}}{K_m^{SU}+C_l}+\frac{V_m^{SPG}}{K_m^{SPG}+C_l}\right)C_l \quad \ldots (26)$$

where, $C_l$ $V_l$ and $CL_r$ refer to the plasma concentration volume of distribution, and renal clearance of SA while $CL_m^{SAG}$, $CL_m^{GA}$, $V_m^{SU}/(K_m^{SU}+C_l)$ and $V_m^{SPG}/(K_m^{SPG}+C_l)$ are the metabolic clearances of SA leading to the formation of SAG, GA, SU, and SPG, respectively.

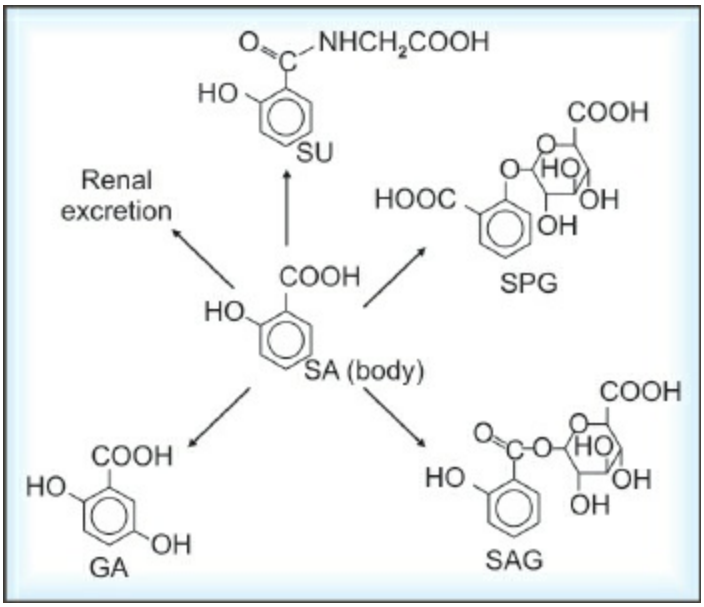



**Fig. 10.15:** Disposition of salicylic acid

## Effects of Diseases

The pharmacokinetics of a drug may be markedly altered in the presence of disease. As discussed earlier, the kidneys and the liver are the major drug-eliminating organs in the body. Any disease process that affects the functional integrity of an eliminating organ or delivery of drug thereto (e.g. blood perfusion) may decrease the rate at which the drug is cleared from the body. In addition, a decline in organ function occurs with advancing age. Cardiac output declines with age and results in a reduced renal and splanchnic blood flow. The effects of these and other age-related physiological changes on drug pharmacokinetics have been reviewed. A reduction in the usual recommended dose or a change in dosage regimen may be required in the diseased or elderly patients to avoid side effects due to excessive drug accumulation and yet achieve the desired therapeutic response. Immature renal and metabolic function must be considered as well when devising dosage regimens for neonates.

The effect of renal impairment on drug elimination depends on the fractional contribution of the kidneys to the total drug clearance. In general, if $f_r$ exceeds 0.3, a reduction in dosage is warranted in the renally-impaired patient. For some compounds, the first-order rate constant for renal excretion, $k_r$ (= $f_r$ $k_{10}$), is directly proportional to the creatinine clearance, $CL_{cr}$, a measure of renal function:

$$k_r = aCL_{cr} \quad \ldots (27)$$

The effect on the overall drug elimination rate constant, $k_{10}$ (= $k_r$ + $k_{nr}$) is then given by:

$$k_{10} = aCL_{cr} + k_{nr} \quad \ldots (28)$$

where, $k_{nr}$ is the first-order rate constant of extra-renal elimination, i.e. (1-$f_r$)$k_{10}$. This equation is used in the individualized dosage adjustment for the renally impaired patient.

For drugs with a narrow therapeutic margin, or when $k$10 at low $CL_{cr}$ cannot otherwise be measured, a first approximation of the relationship between $k_{10}$ and $CL_{cr}$, can be obtained by linear interpolation between the extremes ($CL_{cr}$ = 0 or 120 ml/min), with $f_r$ and $k_{10}$ in healthy subjects being

the only known values ($k_{nr} = (1 - f_r)\, k_{10}$). This method has been successful in predicting the elimination rate constants for a number of antibiotics in severely uremic patients.

In contrast to renal impairment, the effect of hepatic dysfunction on drug elimination is not well denned. This lack of definition is not surprising, considering the host of diseases that affect blood flow, drug-metabolizing enzymes, and biliary secretion, all of which may be involved in drug clearance by the liver. Generalizations, however, can be made based on Eq. (23), that is:

$$CL_h = Q_h E = Q_h \left[ \frac{f_u CL_{int}}{Q_h + f_u CL_{int}} \right] \qquad \text{... (29)}$$

Diseases that cause alterations in hepatic blood flow, protein binding, and intrinsic clearance have an effect on hepatic drug elimination. Drugs that are highly extracted ($E \rightarrow 1$) by the liver are particularly sensitive to change in hepatic blood flow ($Q_h$). In other words, Eq. (29) reduces to Eq. (24), namely, $CL_h \sim Q_h$. For example, because of a reduced cardiac output, the overall $CL$ of lidocaine ($E = 0.7$) in patients with heart failure is 60% of that in healthy subjects.

At the opposite extreme, drugs whose hepatic extraction ratio is low ($E < 0.2$) are insensitive to changes in blood flow. Their hepatic clearance is affected by drug binding to proteins and intrinsic metabolic activity. Tolbutamide, a low extraction drug cleared almost entirely by hepatic metabolism ($CL \sim CL_h$), is an interesting example of this type of dependence. During the acute phase of viral hepatitis, $CL$ for tolbutamide is increased, volume of distribution is unchanged, and plasma $t_{½}$ is decreased. However, the fraction of the drug unbound in plasma ($f_u$) was higher during the active phase of the disease and the intrinsic clearance of unbound drug ($CL_{int}$) was similar during and after recovery from the disease. Since pharmacological activity is related to free drug concentration, no change in dosage is indicated in this case, even though half-life was shorter and plasma concentrations were lower for total drug.

Binding may be altered because of changes in protein composition, concentration, and conformation. For example, changes in binding affinity have been shown to occur with changes in pH in patients with uremia and

cirrhosis of the liver. Dehydration and hypoproteinemia would have opposing effects on protein concentration, and therefore on the free fraction, $f_u$. Concentrations of otj-acid glycoprotein rise dramatically under stress, which is likely to affect the binding of basic drugs. The effect of binding is especially critical for drugs that are highly bound, because a small change in the fraction bound represents a major change in the fraction unbound. Thus, a change from 99% to 98% bound represents a two-fold increase in $f_u$.

# PHARMACOKINETICS

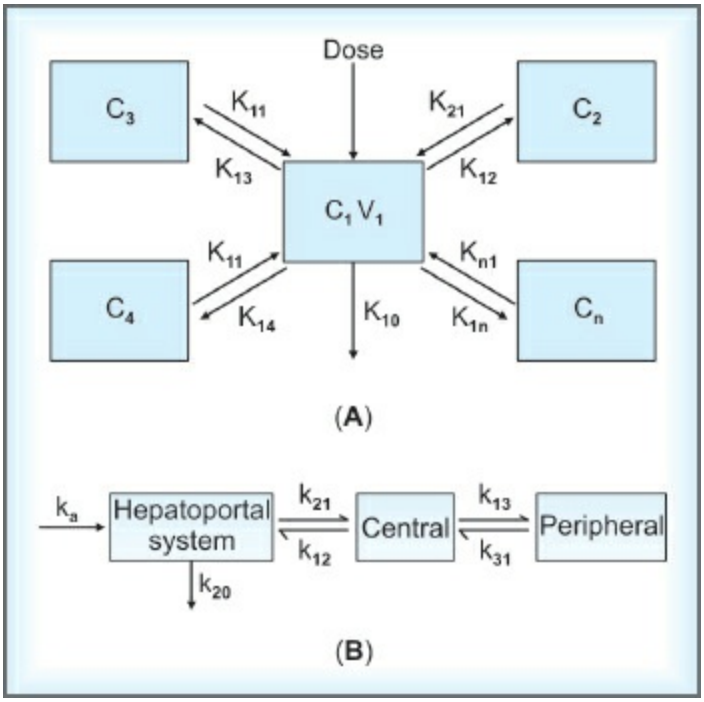
Pharmacokinetics, derived from the Greek words, pharmakon (drug) and kinesis (motion, rate of change), is a discipline that concerns the study and characterization of time course of changes in drug or metabolite concentrations in body fluids. The primary goals of pharmacokinetics are to quantify drug absorption, distribution, biotransformation, and excretion in the intact, living animal or man; and to use this information to predict the effect of alterations in the dose, dosage regimen, route of administration, and physiological state on drug accumulation and disposition. Pharmacokinetics is also used in the clinical settings for the safe and effective therapeutic management of the individual patient, where it is known as clinical pharmacokinetics. The drug concentration in plasma or urine at any time after the administration of a known dose or dosage regimen is the net result of its absorption, distribution, biotransformation (metabolism) and excretion. The task, therefore, is to resolve the observed kinetic profiles into their component parts. The contribution of absorption, distribution, biotransformation, and excretion can be individually isolated by appropriate experimental design and kinetic analysis of the data, often with the aid of models. Quantitation is then achieved by maintaining material balance at all times. There are three models or approaches by which pharmacokinetic analysis of experimental data is performed: (1) compartment modeling, (2) non-compartmental analysis and (3) physiological modeling.

## Compartment Modeling

The most commonly employed approach to the pharmacokinetic characterization of a drug is to represent the body as a system of compartments, even though these compartments usually have no physiological or anatomical reality, and to assume that the rate of transfer between compartments and the rate of drug elimination from compartments follow first-order or linear kinetics. A compartment is an entity which can be described by a definite volume and concentration of drug contained in the volume. A model in pharmacokinetics is a hypothetical structure which can be used to characterize with reproducibility the “fate” and behaviour of a drug in biological systems when given by a certain route of administration and in a particular dosage form. Usually, the behaviour of a drug in biological systems can be described by a one-compartment model or a two-compartment model. Sometimes, it is necessary to employ multi-compartment models. One should begin by determining whether experimental data can be fitted into a one-compartment model. And if only no fitting is obtained, one more sophisticated models can be tried.

Actually, the human body is a multimillion compartment model considering drug concentration in different organelles, cells or tissues. However, in the living body we have access to only two types of body fluids-blood (or plasma or serum) and urine. Compartment models in pharmacokinetics are used to fit experimental data from blood concentration versus time curves or urinary cumulative excretion versus time curves. A certain type of model is not necessarily specific for a particular drug. Often, a blood level versus time curve upon extra-vascular administration can be fitted to a simple one-compartment model, whereas the blood level versus time curve upon intravascular administration is best fitted to a two-compartment model. Two- or more-compartments may be linked together because a drug may move back from one compartment into another. The movement occurs at different rates and is described by distribution rate constants. Depending upon the arrangement of com-partments relative to each other they form two main types of models, i.e. mammillary model and caternary model. ***Mammillary model*** consist of one or more peripheral compartments (comprising of tissues with low vascularity and poor perfusion) connected parallely to the central compartment (comprising of plasma and highly-

perfused tissues such as lungs, liver and kidney) and is most commonly used compartmental model in pharmacokinetics. ***Caternary model*** consists of one or more compartments connected to each other in series and is a rarely used model, since various organs are directly linked to the blood compartment. Figures 10.16A and B depicts a scheme of both types of models.




**Figs 10.16A and B:** (A) Mammillary model representing polyexponential decay and (B) Caternary model

## *One-compartment Open Model*

The one-compartment model, the simplest model, depicts the body as a single, kinetically-homogeneous unit. This model is particularly useful for the pharmacokinetic analysis of drugs that distribute relatively rapidly throughout the body. In an open-one compartment model the drug is not necessarily (and indeed is rarely) confined to the circulatory system. The drug may occupy the entire extracellular fluid, soft tissue or the entire body; however, distribution occurs instantly and is not "pooled" in a specific area. Almost invariably, the plasma or serum is the anatomical reference compartment for the one-compartment model, but it is not assumed that the drug concentration in plasma is equal to the concentration of drug in other body fluids or in tissues. It is rather, assumed that the rate of change of drug

concentration in plasma reflects quantitatively the change in drug concentrations throughout the body. The term “open” in conjunction with a compartment model refers to the fact that we do not have a closed system but a unidirectional input and output, into and out of the system.

This model is generally used to describe plasma levels following administration of a single dose of a drug and depending upon the rate of input, various one compartment open models can be possible, i.e.

- *One-compartment open model:* Intravenous (IV) bolus administration
- *One-compartment open model:* Intravenous infusion
- *One-compartment open model:* First order absorption, and
- *One-compartment open model:* Zero order absorption

## *IV Bolus Administration*

Following an intravenous bolus dose, the drug takes about one to three minutes for complete circulation (therefore, the rate of absorption is neglected in calculations) and plasma concentrations of a drug often decline exponentially with time. Figure 10.17 depicts a one-compartment, open model, consistent with mono-exponential decay in plasma levels, in which $C_l$ is plasma concentration, $V_l$ is the apparent volume of distribution, $k_{10}$ is the overall elimination rate constant, $f_r$ is the fraction of the dose that is excreted in the urine unchanged, $f_m$ is the fraction metabolized, and $f_x$ is the fraction eliminated by all other routes, e.g. biliary.

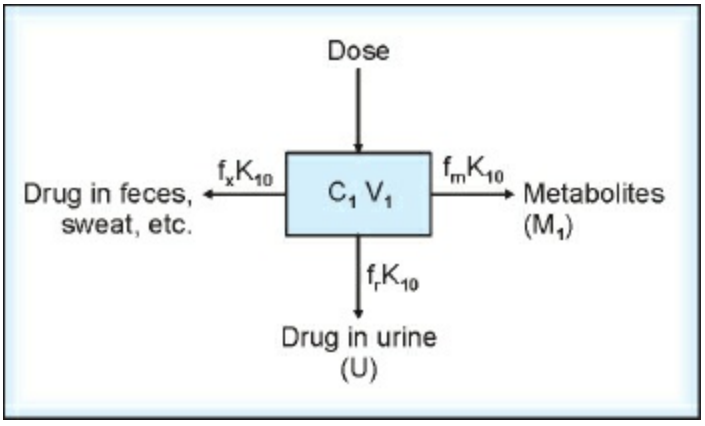



**Fig. 10.17:** One-compartment open model

The rate of change in $C_l$ is therefore:

$$\frac{dC_l}{dt} = k_{10} C_l \qquad \text{... (30)}$$

which upon integration yields:

$$C_l = C_l(0) e^{-k_{10}t} = \frac{D}{V_l} e^{-k_{10}t} \qquad \text{... (31)}$$

At time zero, i.e. $t = 0$, the plasma concentration $C_l$ $(0)$ is equal to the dose, $D$ divided by the volume of distribution, $V_l$.

Experimentally, a semilogarithmic plot of $C_i$ versus $t$ is a straight line, as shown in Fig. 10.18. Model parameters $k_{10}$ and $V_l$, can be calculated from its slope and intercept, respectively. Thus,

$$\log C_l = \log\left(\frac{D}{V_l}\right) - \left(\frac{k_{10}}{2.303}\right) t \qquad \text{... (32)}$$

The volume of distribution is useful in relating plasma concentration to the amount of drug in the body at a given time and, is expressed in units of volume. Thus, the product of $V_l$ and $C_l(t)$ is the amount of drug in the body at time $t$. The elimination rate constant has units of reciprocal time. As such, its numerical value denotes the fractional (or percentage) loss from the body per unit time. For example, $k_{10} = 0.25\ h^{-1}$ signifies that the instantaneous rate of elimination is 25% per hour.

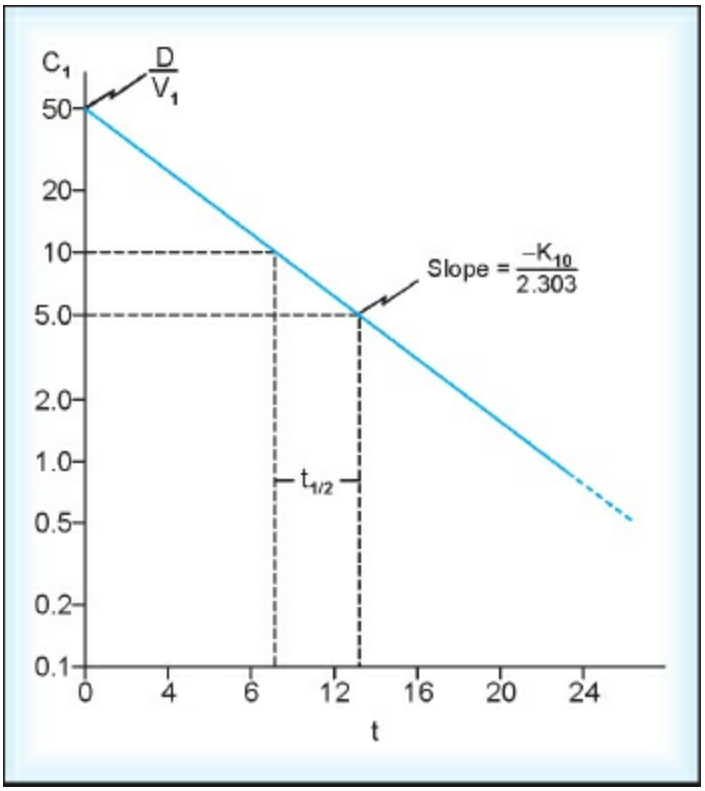



**Fig. 10.18:** A semilogarithmic plot of plasma concentration ($C_1$) vs. time ($t$) for a one-compartment model

The plasma half-life, $t_{½}$, is the time required for a given plasma concentration to be halved. For systems undergoing monoexponential decay, $t_{½}$ is a constant, independent of plasma concentration. Figure 10.18 shows a graphic solution for $t_{½}$. Alternatively, it can be evaluated by substituting $D/2V_l$ for $C_l$ in Eq. (32), whereby:

$$t_{1/2} = \frac{2.303 \log 2}{k_{10}} = \frac{0.693}{k_{10}} \qquad \text{... (33)}$$

The plasma half-life of a drug is inversely proportional to its elimination rate constant.

The rate expression of unchanged drug for ***urinary excretion***, according to the model in Fig. 10.17 is given by:

$$\frac{dU}{dt} = f_r k_{10} V_l C_l \qquad \text{... (34)}$$

where, U is the cumulative amount of drug excreted in urine to time $t$. The time course of urinary excretion is obtained by integrating Eq. (34) and

substituting $C_l$ from Eq. (31),

$$U(t) = f_r k_{10} V_l \int_0^t C_l dt = f_r D(1 - e^{-K_{10}t}) \quad \ldots (35)$$

The amount of drug ultimately recovered in the urine unchanged is obtained by evaluating Eq. (35) at $t = \infty$ as follows:

$$U(\infty) = f_r D \ldots (36)$$

The time course of urinary excretion is a mono-exponential that approaches U(∞) asymptotically. Its shape is, in fact, a mirror image of the plasma concentration profile. Eq. (35) can be rearranged to show that:

$$\log \frac{U_{(\infty)} - U_{(t)}}{U_{(\infty)}} = -\left(\frac{k_{10}}{2.303}\right)t \quad \ldots (37)$$

Thus, a semilogarithmic plot of the fractional amount remaining to be excreted as a function of time should be a straight line with a slope identical to that shown in Fig. 10.2 for plasma decay. Figure 10.19 is sometimes referred to as the *deficit plot*, or *sigma-minus plot* it or Eq. (8) can be used to estimate the ***overall elimination rate constant***, $k_{10}$, from urine data. The rate constant for urinary excretion is $f_r k_{10}$. Experimentally, $f_r$ can be obtained by the ratio of U(∞) to $D$.

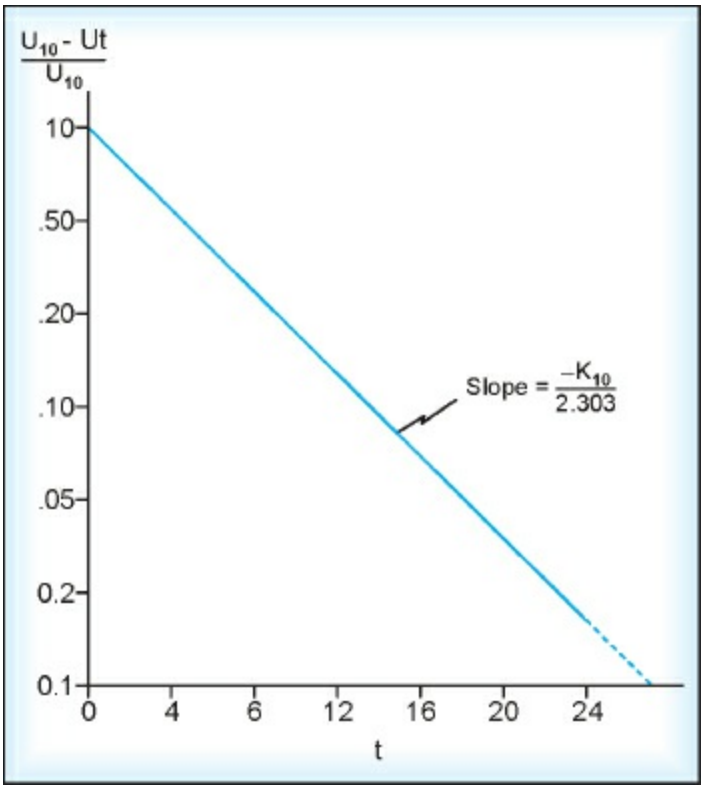



**Fig. 10.19:** Deficit plot, or sigma-minus plot, of the fraction of drug remaining to be excreted in the urine vs. time

According to Eq. (35), the cumulative amount excreted in the urine to time *t* is proportional to the area under the plasma concentration-time curve from 0 to *t*. This relationship holds for all times such that the amount excreted over any interval between $t_1$ and $t_2$ is also proportional to the corresponding area over the same interval, i.e.

$$U(t_1) = f_r k_{10} V_l \int_0^{t_1} C_l dt \qquad \text{... (38)}$$

$$U(t_2) = f_r k_{10} V_l \int_0^{t_2} C_l dt \qquad \text{... (39)}$$

$$U(t_2) - U(t_1) = f_r k_{10} V_l \int_{t_1}^{t_2} C_l dt \qquad \text{... (40)}$$

The proportionality constant relating plasma concentration and urinary excretion of a drug is known as its ***plasma renal clearance rate***, $CL_r$, or simply ***renal clearance***.

$$CL_r = f_r k_{10} V_l = \frac{U(t_2) - U(t_1)}{\int_{t_1}^{t_2} C_l dt} \qquad \text{... (41)}$$

Similarly, blood or serum renal clearance rates would apply to situations

in which drug concentrations are measured in blood or serum. Explicit reference to blood, plasma, or serum is usually unnecessary, except for emphasis or for contrasting results from different experiments.

The renal clearance of substance is the product of its urinary excretion rate constant, $frk_{10}$, and its volume of distribution, $V_1$. Dimensionally, $CL_r$ has units of volume per unit time, e.g. ml/min. Physiologically, the renal clearance rate represents that volume of plasma from which drug is completely eliminated per unit time as a result of passage through the kidneys. A schematic illustration of this phenomenon is shown in Fig. 10.20, in which the density of the shaded areas denotes plasma concentrations before and after passage through the kidneys (or any other clearing organ). The clear zone represents a volume that is completely devoid of drug. Experimentally, the renal clearance of a substance can be determined by dividing the amount of drug present in a timed urine sample by the corresponding-area-under-the plasma concentration-time curve, as shown in Eq. (41). This is depicted graphically in Fig. 10.21.

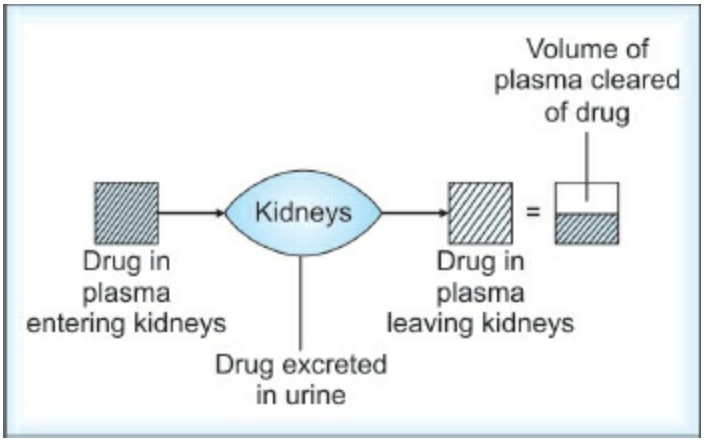



**Fig. 10.20:** Schematic illustration of renal clearance

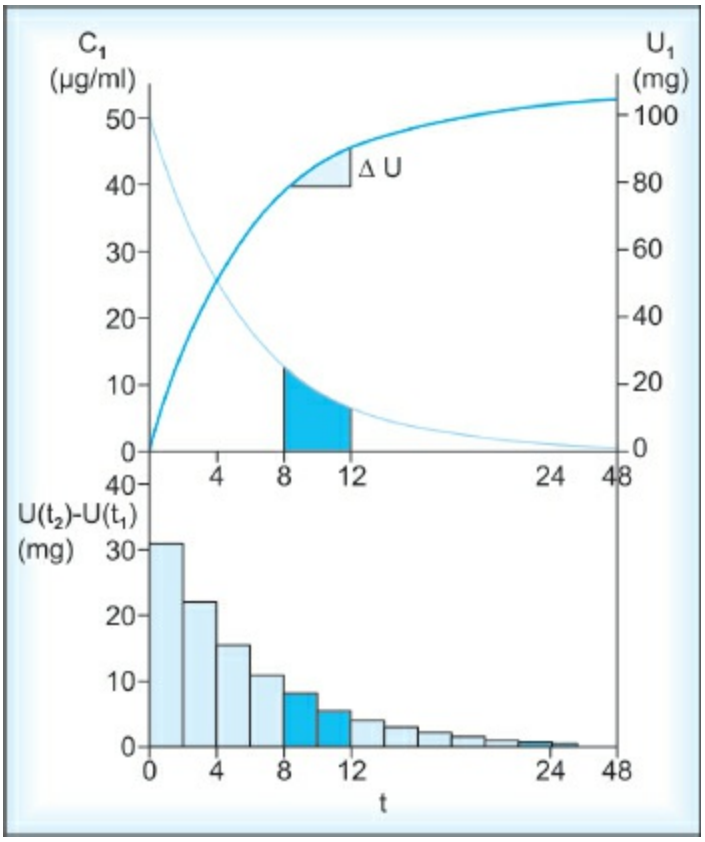



**Fig. 10.21:** Graphic representation of renal clearance calculation. The shaded areas represent a typical area under the plasma concentration curve and the corresponding amount of drug excreted in the urine

The numerical value of renal clearance can be applied in various ways. For example, a combination of Eqs (34) and (41) indicates that the rate of urinary excretion at any instance is equal to the product of renal clearance and the plasma concentration at that time:

$$\frac{dU}{dt} = CL_r C_l \qquad \ldots (42)$$

In the absence of total urinary collection, U(∞), and therefore $f_r$, can be estimated indirectly from the product of $CL_r$ and the total area-under-the plasma concentration-time curve, $AUC_\infty$.

$$U(\infty) = f_r k_{10} V_l \int_0^\infty C_l dt = CL_r AUC_\infty \qquad \ldots (43)$$

$AUC_\infty$ may be obtained as the integral of Eq. (31) evaluated from time zero to infinity, i.e. $\frac{D}{k_{10}V_l}$ or by suitable interpolation and extrapolation techniques.

The clearance concept applies to any body organ or tissue capable of eliminating a drug. For example, hepatic clearance refers to the rate of elimination resulting from passage through the liver, while biliary clearance refers to that part of hepatic clearance that is due to excretion in bile. Unlike measurement of renal clearance, however, their direct experimental determination in the intact animal or person is difficult, and in some cases, impossible. Nevertheless, insights concerning their contribution are often obtained through pharmacokinetic inferences.

The overall elimination of a drug from the body is composed of fractions that are excreted in urine, metabolized, and eliminated by all other means such that $f_r + f_m + f_x = 1.0$. Since renal clearance can be represented by $f_r k_{10} V_l$, the total clearance rate from plasma is $k_{10} V_l$. Multiplying both sides of Eq. (30) by $V_l$ we get:

$$-V_l \frac{dC_1}{dt} = k_{10} V_l C_l = CLC_l \qquad \text{... (44)}$$

where, *CL* is the total plasma clearance rate, or simply plasma clearance. Eq. (44) states that *CL* is the proportionality constant relating the overall elimination rate and plasma concen tration at any time. Integrating Eq. (44) with respect to time:

$$V_l C_l(0) - V_l C_l(t) = CL \int_0^t C_l dt \qquad \text{... (45)}$$

In as much as $V_l\ C_l\ (0)$ is the administered dose *D* and $V_l\ C_l\ (t)$ is the amount of drug in the body at time t, the left-hand side of Eq. (16) represents the amount of drug already eliminated at *t*. In other words, the product of plasma clearance and the cumulative area under the plasma concentration-time curve at *t* is equal to elimination by all routes to that time:

$$\text{Total amount eliminated at } t = CL \int_0^t C_l dt \qquad \text{... (46)}$$

When $t = \infty, C_l(\infty) = 0$; therefore;

$$D = CL.\ AUC_\infty \qquad \text{... (47)}$$

Ultimately, elimination by all routes accounts for the amount administered.

The experimental determination of plasma clearance can be accomplished in various ways following an intravenous dose. Given a reasonably complete set of plasma concentration data, *CL* can be estimated by the ratio of *D* to

$AUC_{\infty}$ or by the product of $k_{10}$ and $V_1$ obtained from the slope and intercept in Fig. 10.18. Given complete urine collection and an estimate of renal clearance, *CL* can also be calculated by:

$$CL = \frac{CL_r}{f_r} = \frac{CL_r D}{U(\infty)} \qquad \ldots (48)$$

In summary, clearance is a proportionality constant that relates plasma concentration of a substance to its elimination rate. It serves the same purpose in relating area-under-the plasma concentration-time curve to the amount eliminated. Plasma clearance can be readily determined from plasma concentration and urinary excretion data following an intravenous dose. Renal clearance can be estimated directly from any conveniently timed plasma and urine sample irrespective of the mode of administration.

The difference between plasma and renal clearance is clearance by other routes:

$$CL_{nr} = CL - CL_r \ldots (49)$$

The dominant component of extrarenal clearance, $CL_{nr}$, is usually biotransformation, although biliary excretion, expiration, or perspiration may also contribute.

The metabolic clearance of a drug, $CL_m$, is the sum of all processes that result in the formation of primary metabolites, *mf*. If extrarenal clearance is due solely to biotransformation, $CL_m$ can be estimated directly from Eq. (49). Otherwise, the contribution of each primary metabolic path must be examined independently. The procedure outlined below assumes a single metabolite but may be applied repeatedly for any specific precursor/successor pairs. Note, however, that secondary or tertiary metabolite formation has no impact on the metabolic clearance rate of the parent drug.

Figure 10.22 depicts a situation where drug and metabolite disposition can each be described by a one-compartment open model. Thus, following an intravenous dose, *M* of the metabolite, its plasma concentrations decline monoexponentially with time i.e.

$$C_l^m = \frac{M}{V_l^m} e - k_{10}^m t \qquad \ldots (50)$$

where, the superscript *m* on $C_l$ $V_l$ and $k_{10}$ denotes plasma concentration,

volume of distribution, and elimination rate constant of the metabolite. The total area-under-the metabolite concentration-time curve is therefore:

$$\int_0^\infty C_l^m dt = \frac{M}{K_{10}^m V_l^m} \qquad \ldots (51)$$

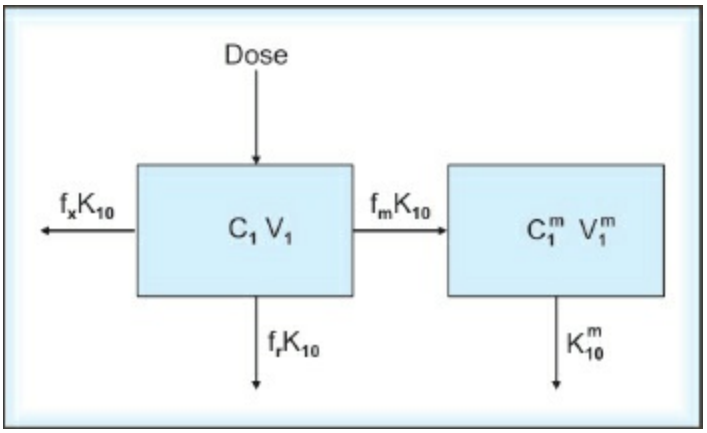



**Fig. 10.22:** Disposition of drug and metabolite according to a one-compartment model for each

The rate expression for metabolite concentrations in plasma following an intravenous dose, *D* of drug is:

$$\frac{dC_l^{m^*}}{dt} = f_m k_{10} C_l - k_{10}^m C_l^{m^*} \qquad \ldots (52)$$

The asterisk (*) distinguishes metabolite concentrations following drug administration from those after metabolite administration. The plasma time course of the metabolite following drug administration is obtained by substituting $C_l$ from Eq. (31) into Eq. (52) and integrating, namely:

$$C_l^{m^*} = \frac{f_m k_{10} D}{V_l^m (k_{10}^m - k_{10})} [e^{-k_{10}t} - e - k_{10}^m t] \qquad \ldots (53)$$

Eq. (24) can be further integrated and rearranged to yield:

$$f_m D = k_{10}^m V_l^m \int_0^\infty C_l^{m^*} dt \qquad \ldots (54)$$

The left-hand side of Eq. (54) is the amount of metabolite formed, whereas the right-hand side is the amount of metabolite ultimately eliminated from the body. The plasma clearance of metabolite, $k_{10}^m V_l^m$ can be obtained from Eq. (51) following an intravenous dose *M* of the metabolite.

An alternative view of the same situation is obtained by combining Eqs

(51) and (54). The fractional contribution of a specific metabolic pathway to the plasma clearance of a drug is given by the dose normalized ratio of total areas under the metabolite curve following the intravenous administration of drug *D* and metabolite *M*, namely:

$$f_m = \frac{AUC_{\infty}^{m^*} / D}{AUC_{\infty}^{m} / M} \qquad \text{... (55)}$$

A potential source of error occurs when the metabolite formed undergoes further biotransformation before reaching the general circulation. This phenomenon, known as *sequential metabolism*, causes an underestimation of the area term in Eq. (54), and hence the amount formed.

## *IV Infusion*

Suppose a drug is administered intravenously by infusion at a constant rate *R*0 (amount/[or:] time). Plasma concentration (and the amount of drug in the body) would increase with time and approach a plateau, which is called as steady state. At steady state, the rate of input to the body, $R_0$, equals the rate of drug elimination from the body, that is:

$$R_0 = CL.C_l^{ss} \qquad \text{... (57)}$$

where, $C_l^{ss}$ is the plasma concentration at steady state. The value of $C_l^{ss}$ to be attained is proportional to the infusion rate. Eq. (57) is applicable to all linear mammillary models of drug disposition. The rate of approach to steady state depends on the manifestations of drug distribution.

Where the one-compartment, open model (*see* Fig. 10.17) applies, the rate of change in drug levels is given by:

$$V_l \frac{dC_l}{dt} = R_0 - k_{10} V_l C_l \qquad \text{... (58)}$$

which, upon integration, yields:

$$C_l = \frac{R_0}{k_{10} V_l}(1 - e^{-k_{10}t}) \qquad \text{... (59)}$$

At $t = \infty$, $C_l$ $C_l^{ss}$. Thus, Eq. (59) reduces to Eq. (57). Subtracting Eq. (59) from Eq. (57), one obtains:

$$\frac{C_l^{ss} - C_l}{C_l^{ss}} = e^{-k_{10}t} \qquad \text{... (60)}$$

or:

$$\log\left(\frac{C_l^{ss} - C_l}{C_l^{ss}}\right) = -\left(\frac{k_{10}}{2.303}\right) \qquad \ldots (61)$$

Therefore, a plot of the left-hand side of Eq. (61) versus *t* should be a straight line with slope of $-k_{10}/2.303$. Eq. (61) also indicates that the rate of approach to steady state is directly related to the plasma half-life. For example, the time required to reach 50% of steady state can be determined by setting $C_l = C_l^{ss}/2$ = in Eq. (61) whereupon as explained below:

$$\log(½) = -\left(\frac{k_{10}}{2.303}\right) t_{50\%} \qquad \ldots (62)$$

or:

$$t_{50\%} = \frac{2.303 \log 2}{k_{10}} \qquad \ldots (63)$$

which is identical to the definition of plasma half-life given in Eq. (33). Similarly, it can be shown that two half-lives would be required to achieve 75% of steady state, 3.3 half-lives for 90%, 6.6 half-lives for 99%, and so on. This relationship between plasma half-life and approach to steady state is shown in Fig. 10.23. Note that the approach to steady state follows exactly the same time course as that for drug disappearance from plasma following an intravenous injection. For drugs that exhibit polyexponential decay, the approach to steady state would also be polyexponential.

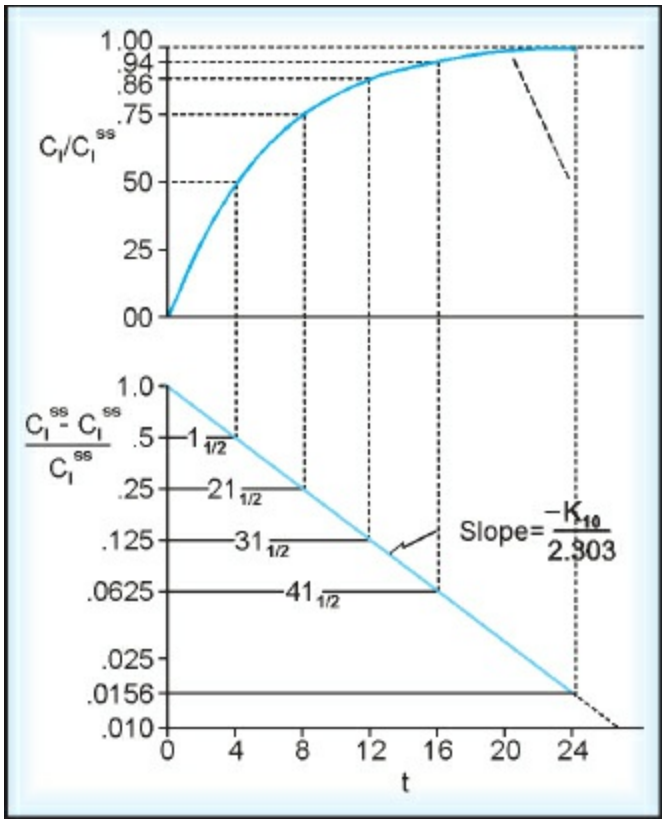



**Fig. 10.23:** Relationship between plasma half-life and approach to steady state for a one-compartment model drug

## *Repeated IV Administration*

A steady state in drug accumulation can also be achieved by dosing at regular intervals. In the simplest case, consider the events following repeated intravenous administration at intervals *t* of a drug that undergoes monoexponential decay. According to Eq. (31), plasma concentration after the first dose is represented by:

$$C_l^{(1)} = \frac{D}{V_l} e^{-k_{10}t} \qquad \text{... (64)}$$

At $t = \tau$, a second dose is administered. Plasma concentrations thereafter are then represented by:

$$C_l^{(2)} = \frac{D}{V_l}(1 + e^{-k_{10}\tau})e^{-k_{10}t'} \qquad \text{... (65)}$$

where, *t*' is the time from the most recent dose. Similarly, after the third and subsequent doses, one can calculate:

$$C_l^{(3)} = \frac{D}{V_l}(1 + e^{-k_{10}\tau} + e^{-2k_{10}\tau})e^{-k_{10}t'} \quad \ldots (66)$$

and;

$$C_l^{(n)} = \frac{D}{V_l}(1 + e^{-k_{10}t'} + e^{-2k_{10}\tau} + \mathrm{L} + e^{-(n-1)k_{10}\tau})e^{-k_{10}t'}$$

$$= \frac{D(1 - e^{-nk_{10}\tau})}{V_l(1 - e^{-k_{10}\tau})}e^{-k_{10}t'} \quad \ldots (67)$$

As *n* becomes large, the plasma concentration profiles from one interval to the next become indistinguishable. At steady state, $n = \infty$, the drug plasma-time course is represented by:

$$C_l^{(ss)} = \frac{D}{V_l}\frac{e^{-k_{10}t'}}{(1 - e^{-k_{10}\tau})} \quad \ldots (68)$$

The mean plasma concentration over a dosage interval at steady state, $\overline{C}_l^{(ss)}$ can be determined by dividing the integral of Eq. (68) by $\tau$ such that:

$$\overline{C}_l^{(ss)} = \frac{\int_0^\tau C_l^{(ss)} dt'}{\tau} = \frac{D}{\tau k_{10} V_l} \quad \ldots (69)$$

Mean plasma concentration at steady state, after intermittent dosage serves the same purpose as $\overline{C}_l^{(ss)}$ after a constant infusion. They would, in fact, be numerically identical if the rate of dosage, $D/\tau$, given discretely were the same as the infusion rate, $R_0$, given continuously. A graphic representation of drug accumulation after a constant infusion and intermittent oral dosage is shown in Fig. 10.24.

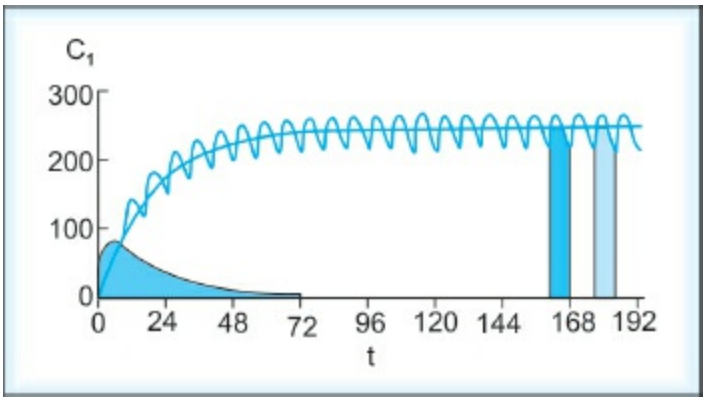



**Fig. 10.24:** Drug accumulation after a constant infusion and intermittent oral doses taken at equal intervals; shaded areas illustrate the equality between $AUC_\infty^{(1)}$ and $AUC_\tau^{(ss)}$

## *Extravascular Absorption*

When a drug is not administered directly into the vasculature, it must be transported to the general circulation before it can be counted. In pharmacokinetics, *absorption* is defined as the amount of drug that reaches the general circulation unchanged. Hence, that which is metabolized or chemically transformed at the site of application or in transit is, by definition, not absorbed.

Material balance dictates that total elimination must equal the amount absorbed, that is:

Amount absorbed = Amount eliminated

If $F$ is the fraction absorbed following a nonintravascular dose $D^x$, then,

$$FD^x = CL \times AUC_{\infty}^x \qquad \text{... (70)}$$

where, $AUC_{\infty}^x$ is the total area-under-the plasma concentration-time, curve following treatment $x$. The estimation of bioavailability requires knowledge of plasma clearance, which is obtained independently following an intravenous dose. *See* Eq. (47) for the determination of plasma clearance.

At any time $t$ following the administration of a dose, the amount absorbed, $A(t)$, is equal to the sum of that which is present in the body, $A_b(t)$, and that which is already eliminated from the body. In other words:

$$A(t) = A_b(t) + CL\int_0^t C_l dt \qquad \text{... (71)}$$

Estimates of $A_b(t)$ depend on how the drug is distributed in the body. When drug disposition can be adequately described by a one-compartment open model, $A_b\ (t) = V_1C_1(t)$ and the time course of absorption is estimated by:

$$A(t) = V_lC_l(t) + k_{10}V_l\int_0^t C_l dt \qquad \text{... (72)}$$

This is known as the Wagner-Nelson method of estimating absorption, $V_1$ and $k_{10}$ can be obtained following an IV dose (*see* Fig. 10.18).

## *Two-compartment Open Model*

If a drug entering the body (input) does not instantly distribute (equilibrate) between the blood and those other body fluids or tissues which it eventually reaches, its kinetics is supposed to follow two-compartment open model. The

distribution of drug in the blood and other soft tissues, on one hand, and into deep tissues, on the other hand, occurs at different rates (speed). Eventually, a steady state will be reached which terminates the “distribution” phase. From such data we cannot necessarily say to which specific tissues or organs a drug was slowly distributed. We may postulate from the *pKa* value and lipid/water partition coefficient, but a definitive answer may be obtained only by biopsy, animal experiments, use of radioactive materials and whole body scintillation. The body fluids or tissues which are in equilibrium with circulatory system comprise the central compartment which is accessible through blood sample. Those body fluids or tissues into which the drug distributes slowly comprise the peripheral compartment which are not accessible by blood sampling.

## *IV Bolus Administration*

After IV bolus of a drug that follows two-compartment open model kinetics, the decline in plasma concentration is biexponential indicating the presence of two disposition processes, i.e. distribution and elimination. A pharmacokinetic model consistent with biexponential decay is shown in Fig. 10.25. The body is perceived to be divided into two kinetically-distinct compartments. A central compartment 1, which includes plasma and from which elimination occurs, is linked to a peripheral compartment 2 by first-order processes having rate constants *k*12 and *k*21.

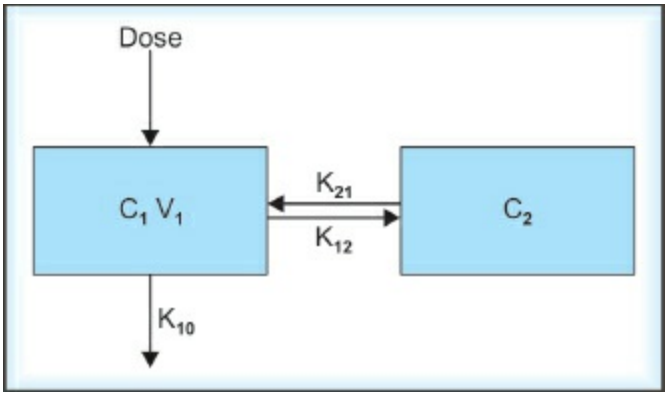



**Fig. 10.25:** Two-compartment open model

Rate expressions apropos to this model are:

$$\frac{dC_1}{dt} = -k_{12}C_1 + k_{21}C_2 - k_{10}C_1 \quad \text{... (73)}$$

$$\frac{dC_2}{dt} = k_{12}C_1 - k_{21}C_2 \quad \text{... (74)}$$

$$\frac{dU}{dt} = f_r k_{10} V_l C_l \quad \text{... (75)}$$

where, $C_2$ is a hypothetic concentration whose product with $V_l$ equals the amount of drug present in the peripheral compartment.

Figure 10.26 is a typical semilogarithmic plot of biexponential decay corresponding to Eq. (28).

$$C_1 = Ae^{-\alpha t} + \beta t \text{... (76)}$$

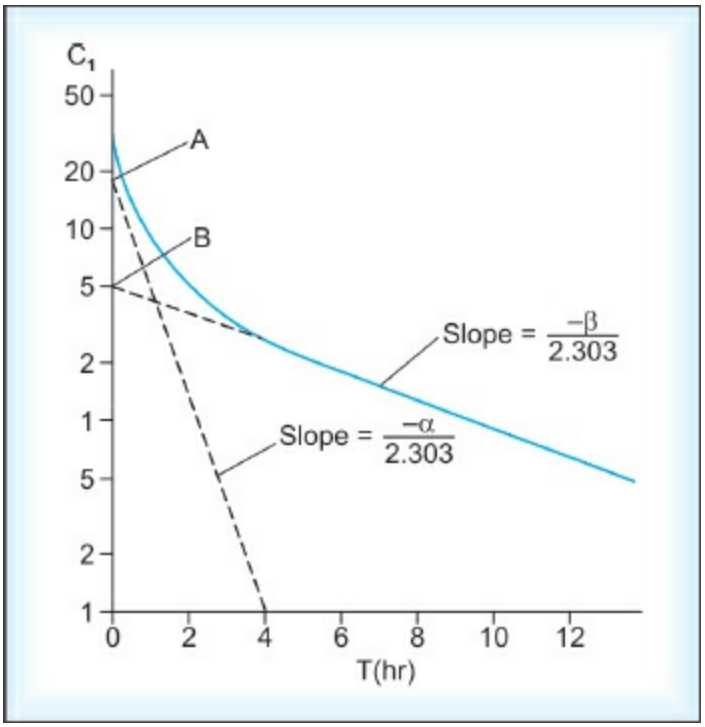



**Fig. 10.26:** Semilogarithmic plot of plasma concentration ($C_1$) vs. time for a two-compartment model

The profile shows an initial curvature that eventually becomes log-linear with a terminal slope, −β/2.303. The intercept B is obtained by extrapolation back to time zero. Taking the logarithm of the difference between plasma concentration $C_1$ and the value of $Be^{-\beta t}$ yields another straight line from which $A$ and α can be evaluated, that is:

$$\log[C_1 - Be^{-\beta t}] = \log A - \frac{\alpha t}{2.303} \quad \ldots (77)$$

This technique, known as *curve stripping*, can be repeated as often as necessary, and is generally useful in obtaining the eigenvalues (slopes) and eigenvectors (intercepts) of a polyexponential function.

The relationship between the model parameters and the experimentally-determined slopes ($\alpha$ and $\beta$) and intercepts ($A$ and $B$) is obtained by simultaneous solution of Eqs (73) and (74), whereby:

$$C_1 = \frac{D(k_{21} - \alpha)}{V_1(\beta - \alpha)} e^{-\alpha t} + \frac{D(k_{21} - \beta)}{V_1(\alpha - \beta)} e^{-\beta t} \quad \ldots (78)$$

such that:

$$A = \frac{D(k_{21} - \alpha)}{V_1(\beta - \alpha)} \quad \ldots (79)$$

$$B = \frac{D(k_{21} - \beta)}{V_1(\alpha - \beta)} \quad \ldots (80)$$

$$\alpha = \frac{(k_{10} + k_{21} + k_{12})}{2} + \frac{\sqrt{K_{10} + k_{21} + k_{12})^2 - 4k_{21}k_{10}}}{2} \quad \ldots (81)$$

$$\beta = \frac{(k_{10} + k_{21} + k_{12})}{2} - \frac{\sqrt{(k_{10} + k_{21} + k_{12})^2 - 4k_{21}k_{10}}}{2} \quad \ldots (82)$$

The corresponding time courses of change in the amount of drug in the peripheral compartment and in urine are, respectively:

$$V_1C_2 = \frac{k_{12}D}{(\beta - \alpha)} e^{-\alpha t} + \frac{k_{12}D}{(\alpha - \beta)} e^{-\beta t} \quad \ldots (83)$$

and:

$$U = f_r k_{10} V_1 \int_0^t C_1 dt$$

$$= f_r D \left[ 1 - \frac{(\beta - k_{10})}{(\beta - \alpha)} e^{-\alpha t} - \frac{(\alpha - k_{10})}{(\alpha - \beta)} e^{-\beta t} \right] \quad \ldots (84)$$

The model parameters can be evaluated by combining and rearranging Eqs (79) through (82).

$$V_1 = \frac{D}{A+B} \quad \text{... (85)}$$

$$k_{10} = \frac{A+B}{A/\alpha + \beta/\beta} \quad \text{... (86)}$$

$$k_{21} = \frac{\alpha\beta}{k_{10}} \quad \text{... (87)}$$

$$k_{12} = \alpha + \beta - k_{10} - k_{21} \quad \text{... (88)}$$

Obviously, the distribution process alters not only the plasma concentration and urinary excretion profiles but also the interpretation of many of the pharmacokinetic parameters. Notably, the terminal slope, $\beta$ or the corresponding $t_{½}$, $\beta$ does not reflect the rate of drug elimination from the body; it merely represents the slowest rate of drug disappearance from plasma. Also, the volume of distribution $V_1$ must be qualified by reference to the central compartment. The product of $V_1$ and $C_1$ is not sufficient to account for the quantity of drug in the body. The contribution from the peripheral compartment, i.e. $V_1C_2$, must be included.

The meaning of clearance remains unchanged, however. Eq. (78) can be integrated and rearranged to yield an expression for plasma clearance, namely:

$$CL = k_{10}V_1 = \frac{D}{AUC_\infty} \quad \text{... (89)}$$

which is identical to that obtained for the one-compartment open model. $AUC_\infty$ may be obtained as the integral of Eq. (76) evaluated from time zero to infinity, i.e. $A/\alpha$ + B/β, or by interpolation and extrapolation techniques. According to Eqs (75) and (84), the product of renal clearance and plasma concentration is the instantaneous rate of urinary excretion, while the product of $Cl_r$ and $AUC_\infty$ is the amount excreted in urine. Multiplying both sides of Eq. (73) by $V_1$, the rate of drug disappearance from plasma is:

$$-V_1\frac{dC_1}{dt} = k_{10}V_1C_1 + k_{12}V_1C_1 - k_{21}V_1C_2 \quad \text{... (90)}$$

The first term on the right-hand side of Eq. (80) represents the rate of drug loss from the body; the second and third terms indicate drug distribution. Again, this emphasizes the difference between drug elimination rate and plasma disappearance rate except for monoexponential decay.

## *IV Infusion*

Consider the situation whereby drug disposition can best be described by a two-compartment open model (*see* Fig. 10.25). Temporal changes in plasma concentration attendant to a constant infusion of a drug are given by:

$$C_1 = \frac{R_0}{k_{10}V_1}\left\{1 - \frac{(\beta - k_{10})}{(\beta - \alpha)}e^{-\alpha t} - \frac{(\alpha - k_{10})}{(\alpha - \beta)}e^{-\beta t}\right\} \quad \ldots (91)$$

At steady state, $t = \infty$, Eq. (91) also reduces to Eq. (57). Subtracting Eq. (91) from Eq. (57) we obtain:

$$\frac{C_1^{ss} - C_1}{C_1^{ss}} = \frac{(\beta - k_{10})}{(\beta - \alpha)}e^{-\alpha t} + \frac{(\alpha - k_{10})}{(\alpha - \beta)}e^{-\beta t} \quad \ldots (92)$$

In this case, the approach to steady state results from two independent exponential processes, $\alpha$ and $\beta$, both governed by their own half-lives, $t_{½,\beta}$ and $t_{½,\alpha}$,. The fractional contribution to the observed sum is given by the eigenvectors of Eq. (92), because of the following relationship:

$$\frac{(\beta - k_{10})}{(\beta - \alpha)} + \frac{(\alpha - k_{10})}{(\alpha - \beta)} = 1 \quad \ldots (93)$$

For each component, the times required to reach 50%, 90%, and 99% of the steady state are still 1.0, 3.3, and 6.6 half-lives, respectively. According to Eqs (81) and (82), however, $t_{½,\beta}$ is no longer than $t_{½,\beta}$ thus, steady state for the *a*process would be attained sooner. Figure 10.27 shows a biexponential approach to steady state. At all times, the observed plasma concentration represents the sum of two processes.

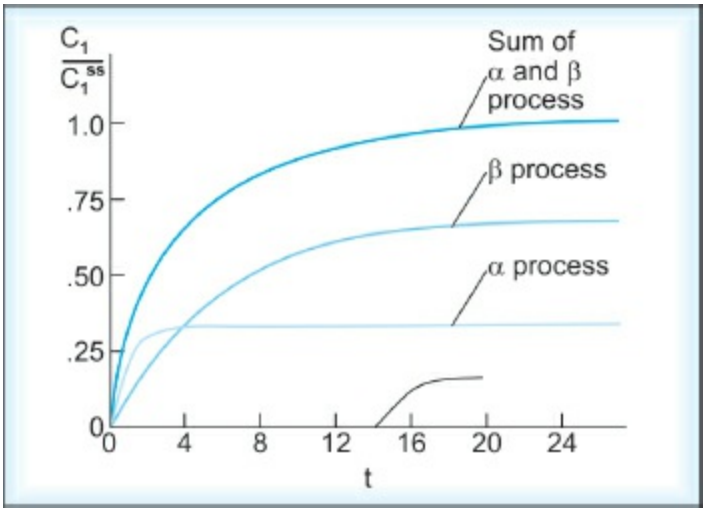


**Fig. 10.27:** Biexponential approach to steady state

## *Repeated IV Administration*

Eq. (69) would also apply to the repeated intravenous administration of drugs that undergo polyexponential decay, although the time courses of change in plasma concentration would differ. For example, with biexponential decay, plasma concentrations during the nth dosing interval are represented by:

$$C_1^{(n)} = \frac{D}{V_1}\left\{\frac{(k_{21}-\alpha)(1-e^{-\eta\alpha\tau})}{(\beta-\alpha)(1-e^{-\alpha\tau})}e^{-\alpha t'} + \frac{(k_{21}-\beta)(1-e^{-\eta\beta\tau})}{(\alpha-\beta)(1-e^{-\beta\tau})}e^{-\beta t'}\right\} \quad \ldots (94)$$

The rate of approach to steady state would, on average, emulate a constant infusion.

Two additional properties of repeated intermittent dosage should be considered. First, the area-under-the plasma concentrationtime curve over one dosage interval at steady state, $AUC_\tau^{(ss)}$, is equal to the total area under the plasma concentration-time curve after a single dose, that is:

$$AUC_\tau^{(ss)} = AUC_\infty^{(1)} \quad \ldots (95)$$

This identity is given by the shaded areas of Fig. 10.14. At steady state, drug input into the body (FD) is equal to drug elimination from the body. Mathematically, these relationships can be summarized by rearrangements of Eqs (89), (70) and (95) such that:

$$A(t) = V_1C_1(t) + V_1C_2(t) + k_{10}V_1\int_0^t C_1dt \quad \ldots (97)$$

Second, a steady state can also be attained by intermittent drug administration at unequal intervals provided that the dosage sequence recurs regularly. For example, daily intervals of dosage may be $\tau_1$, $\tau_2$, and $\tau_3$ such that $\tau_1 + \tau_2 + \tau_3 = 24$ h. If the same dosage schedule were maintained from day to day, the plasma concentration profiles would in time become indistinguishable from one day to the next. In general, the concept of a steady state applies whenever the dosage regimen consists of recurring cycles. Within each cycle, the dose and the dosage interval need not be uniform, as is shown in Fig. 10.28.

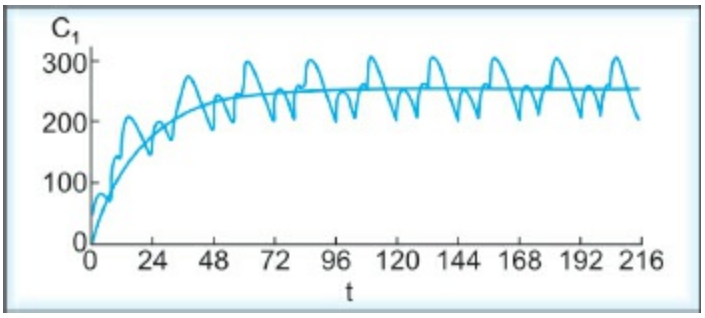



**Fig. 10.28:** Drug accumulation after intermittent oral doses taken at unequal intervals

## *Extravascular Absorption*

When a two-compartment open model applies, the absorption profile can be constructed from Eq. (49):

$$C_1(t_j) = C_1(t_{j-1}) + bT \quad \text{... (98)}$$

The amount of drug present in the body at any time is the sum of the amounts in the central and the peripheral compartment. To evaluate Eq. (97), some method of estimating $C_2$ is needed. As a first approximation, Loo and Riegelman proposed that $C_1$ varies linearly with time between any two adjacent data points, such that:

$$C_1(t_j) = C_1\,(t_{j-1}) + bT \text{ ... (98)}$$

where, $b$ is the slope and $T$ is the time between $t_j$ and $t_{j-1}$ as shown in Fig. 10.29. Integrating Eq. (98) into Eq. (74) yields:

$$\frac{dC_2}{dT} + k_{21}C_2 = k_{12}C_1(t_{j-1}) + bk_{12}T \quad \text{... (99)}$$

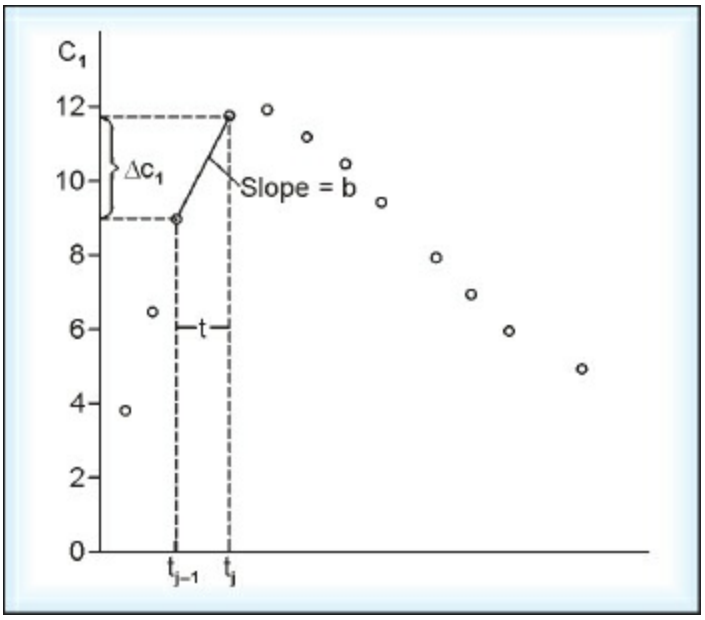



**Fig. 10.29:** Relationship between two adjacent plasma concentration data points proposed by Loo and Riegelmans

Integrating Eq. (99) with respect to $T$ and noting that at $T = 0$, $C_2 = C_2(t_{j-1})$, one obtains:

$$C_2 = C_2(t_{j-1})e^{-k_{21}T} + \frac{k_{12}}{k_{21}}C_1(t_{j-1})[1-e^{-k_{21}T}] + \frac{bk_{12}}{(k_{21})^2}[e^{-k_{21}T} + k_{21}T - 1] \quad \text{... (100)}$$

Now define the relationships:

$$\Delta C_1 = C_1(t_j) - C_1(t_{j-1}) \quad \text{... (101)}$$

and

$$\Delta t = t_j - t_{j-1} \quad \text{... (102)}$$

At T $= \Delta t$,

$$C_2 = C_2(t_j) \text{ and } b = \frac{\Delta C_1}{\Delta t} \quad \text{... (103)}$$

The substitution of Eqs (101), (102), and (103) into Eq. (100) results in:

$$C_2(t_j) = C_2(t_{j-1})e^{-k_{21}\Delta t} + \frac{k_{12}}{k_{21}}C_1(t_{j-1})[1-e^{-k_{21}\Delta t}] + \frac{k_{12}\Delta c_1}{(k_{21})^2\Delta t}[e^{-k_{21}\Delta t} + k_{21}\Delta t - 1] \quad \text{... (104)}$$

Starting at $t = 0$ when $C_2(t_{j-1}) = C_1(t_{j-1}) = 0$, the value of $C_2(t_j)$ can be

estimated sequentially by the repeated application of Eq. (104). At $t = \infty$, no more drug remains in the body, Eq. (97) collapses to Eq. (70), and $A(\infty) = FD$ as expected. Figure 10.30 is a typical absorption profile constructed according to Eq. (97) using the Loo-Riegelman approximation of $C_2$ given in Eq. (104).

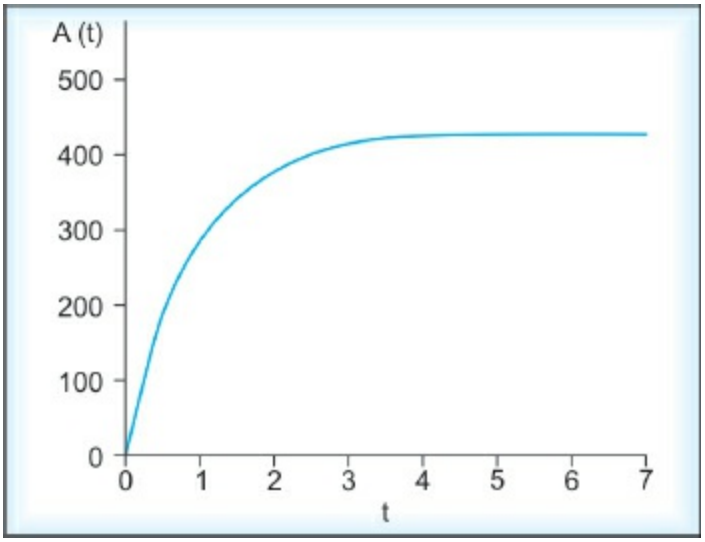



**Fig. 10.30:** A typical Loo-Riegelman absorption profile

In summary, to estimate the bioavailability of a drug administered by a nonintravascular route, knowledge of its plasma clearance is required. The plasma clearance in an animal or a person can be calculated following an intravenous dose. In essence:

$$F = \frac{AUC_{\infty}^{x} / D^{x}}{AUC_{\infty} / D} \qquad \ldots (105)$$

Since $AUC_{\infty}^{x}$ and $AUC_{\infty}$ can both be determined directly by interpolation and extrapolation of the terminal slope of the plasma concentration curve to infinity, the estimation of $F$ is model-independent. In contrast, determinations of the absorption time course based on compartmental analysis depend not only on plasma clearance but also on estimates of the volume of distribution and the rate constants for drug distribution. These values are obtained by interpreting the kinetics of drug disposition following an intravenous dose in reference to specific models.

## Non-compartment or Model Independent Analysis

The compartment model concept is imperative for understanding the principles of pharmacokinetics. However, in recent years the compartment model free or compartment independent analysis has gained increasing attention and application. The reason may be found in the fact that the human body in reality is a multi-million compartment model, but the most sophisticated kinetic multicompartment model may have only very few compartments, in which the microconstants are mathematical artifacts. Another reason is the realization that for clinical application of pharmacokinetics, it is neither possible to obtain a large number of blood samples to properly characterize a multi-compartmental concentration time course nor is it necessary, because dosage regimen design and dosage regimen adjustment require only a few parameters.

Non-compartmental analysis does not require the assumption of a specific compartmental model for either drug or metabolite and can be applied to virtually any compartmental model, provided that we can assume linear pharmacokinetics. These methods are usually based on the statistical moment theory, which involves collection of experimental data following a single dose of drug. According to this theory, the time course of drug concentration in plasma can usually be regarded as a statistical distribution curve, irrespective of the route of administration.

$$MRT = \frac{\int_0^\infty tC\,dt}{\int_0^\infty C\,dt} = \frac{AUMC}{AUC} \quad \ldots (106)$$

$$AUC = \int_0^\infty C\,dt \quad \ldots (107)$$

where, MRT is the mean residence time of a drug in the body. Like half-life, MRT is a function of both distribution and elimination, and can be defined as the average amount of time spent by the drug in the body before being eliminated. *AUC* is the area-under-the zero moment curve, i.e. the area-under-the curve of a plot of the concentration versus time from zero time to infinity and *AUMC* is the area under the first moment curve, i.e. the area under the curve of a plot of the product of concentration and time versus time from zero time to infinity. The moments defined above can be calculated by numerical integration using the ***trapezoidal rule***, from concentration-time data following drug administration. This method involves dividing the curve by a series of vertical lines into a number of trapezoids, calculating separately the

area of each trapezoid and adding them together (Fig. 10.31).

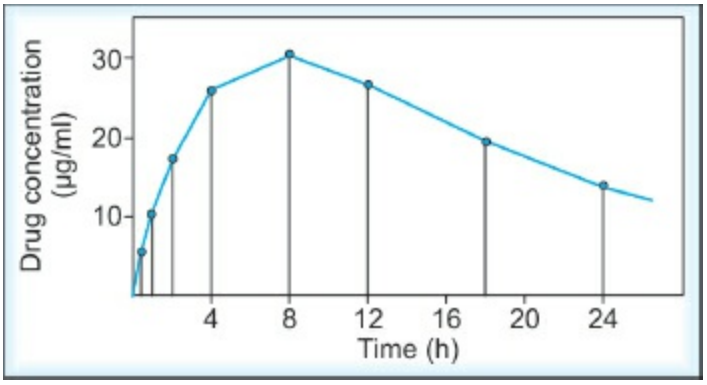



**Fig. 10.31:** Representation of drug concentration in plasma-time profile after oral administration for the application of linear trapezoidal method to estimate areas

Non-compartmental analysis has been used to estimate bioavailability, apparent volume of distribution, clearance and the fraction of a dose of a drug that is converted to a specific metabolite, based on data following single doses of drug and metabolite. These methods have also been used to predict the average steady-state concentration of a drug or its metabolites, based on data following a single dose of the drug, and the time required to reach a given fraction of the steady state concentration when a fixed dose of a drug is given at regular intervals. However, these methods have some limitations as nonlinear events are not adequately treated by statistical moment theory and only limited information is provided regarding the time course of drug concentrations. More often, it deals with averages.

## Physiological Modeling

The pharmacokinetic models described above are inherently limited in the amount of information they provide because, in the usual case, the compartments and the parameters have no obvious relationship to the anatomical structure or physiological function of the species under study. In recent years, efforts have been directed towards the development of physiologically-realistic pharmacokinetic models. These detailed models are elaborated on the basis of the known anatomy and physiology of humans or other animals and incorporate physiological, anatomical, and physiochemical data.

A physiological pharmacokinetic model is composed of a series of lumped compartments (body regions) representing organs or tissue spaces whose drug concentrations are assumed to be uniform. The compartments are arranged in a flow diagram as illustrated by the general example in Fig. 10.32. The first step in the development of a physiological pharmacokinetic model is the selection of compartments to be included. The number of compartments to be added in the model depends upon the disposition characteristics of the drug. Tissues or organs that have no drug penetration (such as bone) are excluded. Once the selection has been made, the kinds of information required by the model can be classified as (1) anatomical (e.g. organ and tissue volumes), (2) physiological (e.g. blood flow rates and enzyme reaction parameters), (3) thermodynamic (e.g. drug-protein binding isotherms), and (4) transport (e.g. membrane permeabilities).

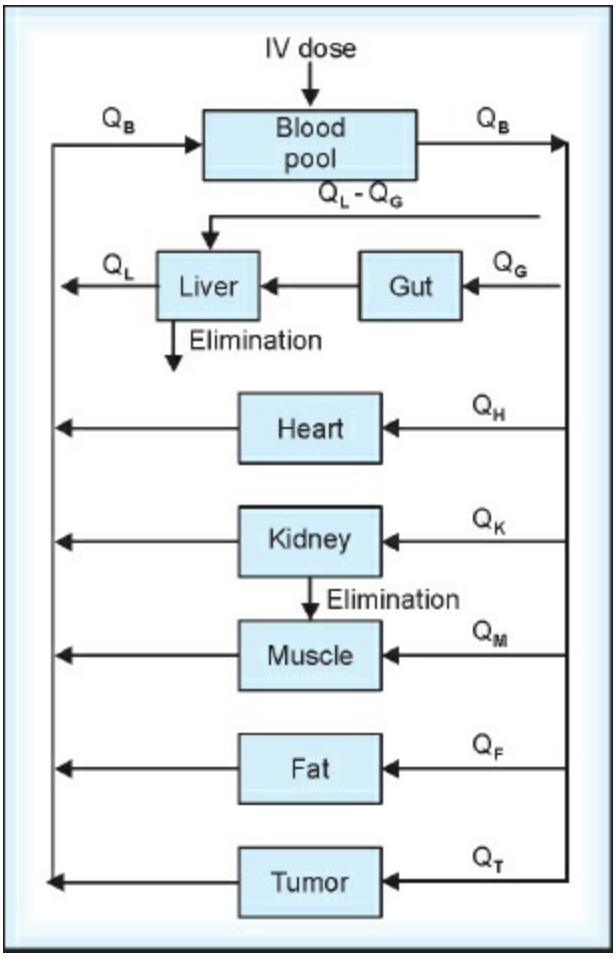



**Fig. 10.32:** Schematic representation of a physiological pharmacokinetic model. The term *Q* denotes blood flow rate to a region. Subscripts are as follows: B, blood; L, liver; G, gut; H, heart; K, kidney; M, muscle; F, fat; T, tumour

Most physiological pharmacokinetic models developed to date are based on the assumption that the drug movement within a body region is much more rapid than its rate of delivery to that region by the perfusing blood and therefore, the model is said to be ***perfusion rate-limited***. The assumption of perfusion-limited transport is applicable to relatively low molecular weight, weakly ionized, lipid-soluble drugs for which diffusion and movement across lipoidal membranes should be relatively rapid. On the other hand, for very polar, highly ionized, or charged drugs the model is referred to as ***membrane permeation rate-limited***.

These comprehensive models are superior to classical compartment models in several respects as they provide an exact description of the time

course of drug concentration in any organ or tissue, and are therefore able to provide greater insight to drug distribution in the body. Also, since the parameters of these models correspond to actual physiological and anatomical measures, such as organ blood flows and volumes, changes in the disposition kinetics of drug, because of physiological or pathological alterations in body, function, may be predicted by perturbation of the appropriate parameter(s). Finally, these models introduce the possibility of animal scale-up which would provide a rational basis for the correlation of drug data among animal species. The only disadvantage of these models is obtaining experimental data which is very exhaustive.

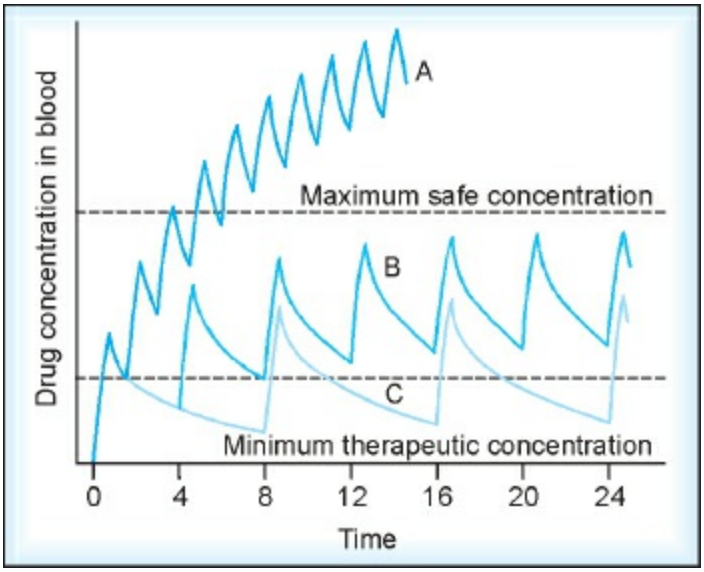
## Softwares for Pharmacokinetic Analysis

The most common software packages to fit pharmacokinetic data based on compartmental models are WinNonLin (Pharsight Corp., Mountain View, CA; www.pharsight.com), SAAM II (SAAM Institute, Seattle, WA; www.saam.com), and Adapt II (Biomedical Simulations Resource at the University of Southern California, bmsrs.usc.edu). All are designed to handle individual pharmacokinetic data, although Adapt II and SAAM II can handle multiple input-output experiments easier than WinNonLin. On the other hand, WinNonLin does batch processing of multiple subjects, whereas SAAM II and Adapt II do not. Adapt II has the advantage that it is free, but has poor graphic capabilities compared to WinNonLin and SAAM II. Of all the packages, SAAM II has the easiest and best simulation capability. There are many other subtle differences between the three pharmacokinetic packages and generally, the one a person chooses is a matter of personality, economics, and regulatory requirements.

## Application of Biopharmaceutics and Pharmacokinetics

### *Dosage Regimen*

The goal in the design of dosage regimens is to achieve and maintain drug concentrations in plasma or at the site of action that are both safe and effective. Maximum safe concentration and minimum therapeutic concentration are schematically illustrated in Fig. 10.33. Toxicity would result if doses were administered too frequently, whereas, effectiveness would wane if the dosage rate were too infrequent. The optimal regimens are combinations of dose and dosage frequency that would result in steady-state concentrations within the chosen limits.




**Figs 10.33A to C:** Effect on plasma concentration of too frequent (A), proper (B), and inadequate (C) frequencies of drug administrations

Acceptable plasma concentration-time profiles at steady state can be devised with the aid of pharmacokinetic parameters derived from single-dose experiments. The important parameters are plasma clearance, half-life, and bioavailability. Suppose that the desired mean plasma concentration, $\overline{C}_1^{(ss)}$, for a drug is 2 µg/ml, and its plasma clearance is 125 ml/min. According to Eqs (57), or (69), the dosage rate by the intravenous route should be 250/µg/min, or 360 mg/day. The target concentration can be attained by either a constant infusion or intermittent boluses (*see* Fig. 10.24). The dose and dosage intervals need not be equally divided; a steady state may be obtained as long

as the dosing pattern recurs in regular cycles (*see* Fig. 10.28). If the same drug is to be given by a non-intravascular route with a bioavailability of 50%, the same target concentration can be achieved with a dosage rate of 720 mg/day.

The peak and trough concentrations within a dosage cycle as well as the rate of approach to steady state depend on the half-life. The shorter the half-life is, the more the daily dose must be subdivided to maintain peak and trough concentrations within the chosen limits. Furthermore, approximately 7 halflives in time must elapse to achieve a steady state. To hasten the approach to steady state, a loading dose of the drug may be given. An appropriate loading dose can be calculated based on the mode of administration by using the following equations.

For a continuous intravenous infusion, an initial loading bolus dose $D^*$ is administered:

$$D^* = C_1^{ss} V_1 \qquad \ldots (108)$$

This initial dose followed immediately by an infusion at a dosage rate prescribed by Eq. (57) results in the immediate attainment of $C_1^{ss}$ for a drug whose disposition is described by a one-compartment pharmacokinetic model (*see* Fig. 10.17). For drugs whose plasma concentration decays polyexponentially (*see* Fig. 10.9), a similar approach causes plasma concentrations to fall for a period of time before returning to $C_1^{ss}$. To compensate for this initial fall in $C_1^{ss}$, a loading bolus dose $D^*$ equal to or greater than that given by Eq. (109) may be considered:

$$D^* \frac{C_1^{ss} CL}{\omega} \qquad \ldots (109)$$

where, $\omega$ is the eigenvalue associated with the terminal slope. This dose is followed immediately by a continuous infusion at a rate prescribed by Eq. (57), however, this strategy results in initial plasma concentrations higher than $C_1^{ss}$ and may be inappropriate for drugs having a narrow margin of safety.

For intermittent dosage at regular intervals, a loading dose ($D^*$) could be chosen such that the mean plasma concentration following the first dosage interval, $\bar{C}_1^{(ss)}$, would equal that at steady state, $\bar{C}_1^{(ss)}$. Combining Eqs (69) and (96) yields:

$$D^* = \frac{\tau \overline{C}_1^{(ss)} CL}{F} \qquad \text{... (110)}$$

where, $F$ represents the bioavailability of the drug. Administration of $D^*$ followed by the maintenance dose given every $\tau$ hours immediately establishes steady state plasma concentrations.

Special consideration must be given to the design of dosage regimens when the disease state of the patient might affect drug disposition. In patients with less than normal cardiovascular function, tissue perfusion is decreased. If hepatic or renal function is compromised, drug elimination is decreased because of a decrease in metabolic or renal clearance, respectively. In each case, plasma clearance decreases and half-life is prolonged. These altered parameters should be used in Eqs (57), (69), (113), (109), and (110) to calculate the doses and dosage rates needed to achieve the desired $\overline{C}_1^{(ss)}$ in patients.

## *Controlled Delivery*

The goals of controlled drug delivery are to conserve dose, maintain effective drug concentrations, eliminate nighttime dosage, improve compliance, and decrease side effects, thus optimizing drug therapy. Pharmacokinetic information is essential in determining the feasibility and design of a controlled-delivery dosage form. Drugs with plasma half-lives of 6 h or less, inactive metabolites, well-defined minimum therapeutic blood levels, and rapid absorption are the most likely candidates for controlled delivery. Dosages for drugs with longer halflives can be calculated conventionally so that therapeutic blood levels are established and then self-sustained, allowing for twice-daily dosage or less. A narrow margin of safety complicates this approach, as does the fact that well-defined minimum therapeutic drug levels are difficult to establish even in the absence of active metabolites.

A common approach has been to combine a rapid-release dose fraction with a fraction having pseudo-first-order release characteristics. When absorption is not rate-limiting, the ideal approach to this situation is zero-order delivery of drug to the absorption site. The amount of drug reaching the absorption site changes with time for first-order drug delivery, thus precluding the desired constant blood level profile. Following administration of a zero-order-input delivery system, steady-state blood levels ( $C_1^{ss}$ ) of the drug are obtained:

$$C_1^{ss} = \frac{k_0}{V_1 k_{10}} \quad \text{... (111)}$$

where, $k_0$ is the zero-order input rate. The dose $D$ is a function of $k_0$, the dosage interval τ, and the bioavailability:

$$k_0 \tau = FD \text{ ...(112)}$$

As indicated in Eq. (111), the $C_1^{ss}$ attained with a zero-order input delivery system is a function of both the input rate and the *CL* of the drug. The time required to reach $C_1^{ss}$ after the dose is administered, however, is governed primarily by the plasma half-life of the drug. Approximately 7 half-lives are required to approach $C_1^{ss}$. If more rapid attainment of the steady-state level is desired, then an immediate-release drug fraction (loading dose) can be administered with the zero-order delivery system. Design strategies accompanying this approach, including the effects following multiple-dose administration, have been reviewed.

While advantages of controlled-delivery dosage forms are readily apparent, limitations must also be addressed prior to and throughout the development process. A convenient once-a-day dosage form is precluded, for instance, for a drug that is bulky or where daily dosage requirements are high. The length of time that an oral controlled-delivery dosage unit remains functional within the GI tract must be considered—functional in terms of transit time as well as the release of drug at the site of absorption. Certain drugs manifest “absorption windows”, whereby absorption is limited to a specific region of the GI tract. Others may be absorbed, albeit nonuniformly, along the entire intestinal length. Once the dosage unit is embedded in faecal matter, drug diffusion to the gut wall may become difficult.

Controlled drug delivery systems for routes of administration other than oral are becoming increasingly popular. Topical dosage forms such as ophthalmic delivery devices (e.g. Ocusert) and dermal drug delivery patches are usually intended for local drug effects. Drugs with lipophilic characteristics may also be delivered transdermally for their systemic effect. Once the skin becomes saturated, the drug should enter the systemic circulation at a constant rate as long as drug activity at the skin surface remains unity. The rate of drug entry into the systemic circulation is governed by the inherent flux of drug through the skin or by the combined effect of a rate-controlling membrane on the patch itself and the skin flux. With the

further advantage of being able to keep a transdermal device in contact with the skin for days, this method may be an ideal approach to the goals of controlled drug delivery. Hence, complete pharmacokinetic characterization of the drug and the delivery device becomes even more critical.

## *Combination Therapy*

Rational combination therapy may be divided into three general categories (1) those combinations whereby the individual components are independently required in the treatment of a specific illness or symptom complex (e.g. an analgesic may be combined with an antipyretic or an antihistamine to alleviate the symptoms of influenza or allergy, while a $\beta$-blocker may be combined with a diuretic to control hypertension), (2) therapy whereby a second agent is needed to ameliorate an unwanted pharmacodynamic effect of the primary agent (e.g. anticholinergic/narcotic antidiarrhoeals reduce potential for narcotic abuse, while some diuretic combinations minimize potassium loss) and (3) combinations designed to effect an improvement in pharmacokinetic properties, (e.g. the co-administration of levodopa and a decarboxylase inhibitor greatly reduces dopamine formation in the GI lumen and outside the central nervous system, the net result being a reduction in the dose and dosing frequency of levodopa, and in side effects associated with peripheral dopamine).

Fixed combinations are dosage forms designed to effect a planned interaction between the components or to provide convenience to the patient. Since each component is required for rational therapy, fixed combinations ensure compliance. Where the components interact, by design or otherwise, fixed combinations define the intensity and duration of the net effect. In any event, clinical proof of safety and efficacy is needed to support the usefulness of a proposed combination.

For convenience dosage forms, the primary biopharmaceutical concern is to ensure that the bioavailability of each component in the fixed combination is equivalent to that following concomitant administration of the single entities at the same dose by the same route. As long as there is proof of safety and efficacy, incidental pharmacokinetic interactions between the components should not be a concern, but may in fact be a strong argument for the fixed combination.

With planned pharmacokinetic interactions, biopharmaceutical

considerations are to optimize the dose of the individual components, their ratio, and the dosage regimen of the combination. More often than not, the desired therapeutic activity resides with one of the components whose absorption or disposition characteristics are improved by a second agent. For example, carbidopa, a dopa decarboxylase inhibitor, is given in combination to maximize the availability of L-dopa for transport to the central nervous system and to minimize dopamine formation in the periphery. In designing this particular combination, it became evident that there is a minimum daily dose of carbidopa above which peripheral dopa decarboxylase activity is maximally inhibited. Above this threshold, a relatively large range in the ratio of L-dopa to carbidopa can be accommodated. Finally, the inhibitory effect of carbidopa requires one or two days of intermittent dosing to be fully manifested. The requirement is not a concern, however, since the combination is intended for chronic therapy.

In contrast, there are physical and pharmacokinetic arguments against fixed combination dosage forms of a $\beta$-lactam antibiotic and probenecid. Like carbidopa, a threshold dose of probenecid is required for maximal blockade of the renal secretory component of antibiotic elimination, and mainly because probenecid is usually given orally, some time must elapse before its effect is fully manifested. Since the optimal route for the antibiotic is often parenteral, however, and the effective dose of the components often exceeds 500 mg each, their combination in a single dosage form is not usually considered. Secondarily, pharmacokinetic considerations suggest pretreatment with probenecid to ensure maximal conservation of the antibiotic. This pretreatment is particularly relevant when only a single dose of the antibiotic is indicated, such as in the treatment of gonorrhea.

Divergent pharmacokinetic properties among individual components in fixed combinations of convenience are unimportant, even if incidental interactions exist. With planned interactions, on the other hand, the adjuvant (e.g. carbidopa and probenecid) should ideally be long-lived relative to the therapeutic moiety. This ensures a stable platform on which the pharmacokinetic behavior of the therapeutic moiety can be reproduced from one dose to the next. Without this stable platform, the kinetic behavior of the compound of primary interest may be difficult to assess.

## *Bioavailability*

Bioavailability is defined as the extent and rate at which the active ingredient is delivered to the general circulation from the dosage form. Thus, by definition, intravenously administered drugs are completely bioavailable. The bioavailability, $F$ following a nonintravenous dose $D^x$, is given by Eq. (70), which can also be expressed as:

$$FD^x = \frac{CL}{CL_r} U_\infty^x = \frac{U_\infty^x}{f_r} \qquad \text{... (113)}$$

Estimation of the product $FD^X$ requires knowledge of $CL$ (or $f_r$), which is obtained in a separate treatment following an intravenous dose of the drug (i.e. $F = 1$). When intravenous use of the drug is precluded, an oral solution of the drug may serve as the reference standard for a solid oral dosage form. In principle, the bioavailability of the drug from an oral solution is the maximum to be expected from a solid oral dosage form. Ideally, a solid dosage form is compared to both an intravenous and oral solution dose of the drug. In this manner, the effect of formulation as well as the absolute bioavailability can be determined. If bioavailability from the solid dosage form is lower than that from an oral solution, then the performance of the dosage form may be improved by reformulation. Low bioavailability from the oral solution, on the other hand, indicates that the drug is intrinsically poorly absorbed or is subject to significant first-pass metabolism and is not likely to be improved by formulation.

In all cases, bioavailability of a test dosage form X is obtained by comparison with a reference standard s, which may be an intravenous or oral solution, or in bioequivalence studies, another formulation of the same drug. Based on Eq. (70), the bioavailability of dosage form X relative to that of s is defined as:

$$\frac{F^x}{F^s} = \frac{D^s}{D^x}\frac{CL^x}{CL^s}\frac{AUC_\infty^x}{AUC_\infty^s} = \frac{D^s}{D^x}\frac{CL^x}{CL^s}\frac{CL_r^s}{CL_r^x}\frac{U_\infty^x}{U_\infty^s} \qquad \text{... (114)}$$

If the standard is an intravenous dose, then $F^s = 1$ and the absolute bioavailability of X is determined; otherwise, a relative bioavailability is obtained. In any event, an assumption must be made regarding body clearances between treatments before bioavailability can be assessed. Depending on the assumption, different procedures may be employed.

If the assumption of $CL^X = CL^S$ is adopted and substituted into Eq. (114), the relative bioavailability is given by the dose-corrected ratio of *AUC* values:

$$F^x / F^s = \frac{D^s}{D^x}\frac{AUC^x_\infty}{AUC^s_\infty} \qquad \text{... (115)}$$

This method requires data defining the entire area under the plasma concentration time curve for both treatments.

On the other hand, for drugs excreted unchanged in the urine, bioavailability may be obtained assuming that the renal to plasma clearance ratio remains constant between treatments. Eq. (114) is reduced to:

$$F^x / F^s = \frac{D^s}{D^x}\frac{U^x_\infty}{U^s_\infty} \qquad \text{... (116)}$$

Thus, the bioavailability ratio equals the ratio of dose-corrected total urinary recoveries of the unchanged drug.

If both urine and *AUC* data are available, a third method may be employed. This method assumes that the change in plasma clearance between two treatments is caused by that of the renal clearance only, and that the nonrenal clearance remains unchanged. The ratio of the two plasma clearances becomes:

$$\frac{CL^x}{CL^s} = \left(\frac{U^x_\infty}{AUC^x_\infty} + \frac{F^sD^s - U^s_\infty}{AUC^s_\infty}\right) \Big/ \left(\frac{F^sD^s}{AUC^s_\infty}\right) \qquad \text{... (117)}$$

Substitution of Eq. (117) into Eq. (114) yields:

$$F^x / F^s = \frac{D^s}{D^x}\frac{AUC^x_\infty}{AUC^s_\infty} + \frac{1}{F^sD^x}\left(U^x_\infty - U^s_\infty\frac{AUC^x_\infty}{AUC^s_x}\right) \qquad \text{... (118)}$$

Eq. (118) differs from Eq. (115) in that it incorporates, in addition to the dose-corrected *AUC* ratio, a correction term to account for the assumed change in plasma clearance. When s is an intravenous dose, $F^S = 1$ and Eq. (118) reduces to:

$$F^x = (CL^s - CL^s_r + CL^x_r)\frac{AUC^x_\infty}{D^x} \qquad \text{... (119)}$$

Thus, $CL^x$ is calculated from the sum of $CL^x_r$ and the nonrenal clearance observed in the IV treatment, that is, $CL^S - CL^s_r$. If the bioavailability of the

reference standard is unknown, an approximate solution to Eq. (118) may be obtained by setting $F^S D^x$ in the second term equal to $D^x$. The nature of this approximation has been defined, and an optimal solution suggested.

A unique approach that obviates the need for assumptions concerning *CL* between treatments is to administer simultaneously an IV dose of the drug labeled with a stable radioisotope and the oral dose of the unlabeled drug.

The foregoing methods of estimating the extent of bioavailability apply only to drugs that obey linear pharmacokinetics. For drugs following nonlinear pharmacokinetics, methodology should be developed on a case-by-case basis. For example, in the absence of first-pass metabolism, total urinary recovery is a measure of bioavailability if the unchanged drug and its metabolites are quantitatively excreted in the urine.

After a nonintravenous dose of the drug is administered, the time to reach the maximum plasma concentration, $t_{max}$, may be regarded as a qualitative measure of the rate of absorption, with the recognition that $t_{max}$ is a function of not only absorption but also elimination of the drug. The preferred method is to use an intravenous dose as a reference standard and to isolate the absorption profile for the non-intravenous formulation by the use of deconvolution or model-dependent means such as the Wagner-Nelson or Loo-Riegelman methods given in Eqs (72) through (104).

## *Example*

Suppose a drug is to be marketed in tablet formulations containing 250, 500, and 1000 mg of the drug. The bioavailability of the drug and, therefore whether the amount of drug absorbed is proportional to the administered dose (dose-proportionality), need to be determined. If it is already established that the drug obeys linear disposition kinetics, both objectives can be met by comparing the three formulations with a single IV reference dose of the drug in a single-dose, randomized, four-way crossover study.

Simulated results for drug assay in plasma and urine after the 250-, 500-, and 1000-mg oral dose treatments are given in Table 10.4, and the resulting pharmacokinetic parameters are as given in Table 10.5. Plasma drug profiles after the oral doses are compared in Fig. 10.34. The drug is apparently rapidly absorbed, peak plasma concentrations are observed within an hour of each dose and increase in proportion to the administered dose. After ~6 h,

plasma concentrations decline log-linearly with a terminal disposition half-life that is similar among all treatments, including the IV dose (Table 10.5).

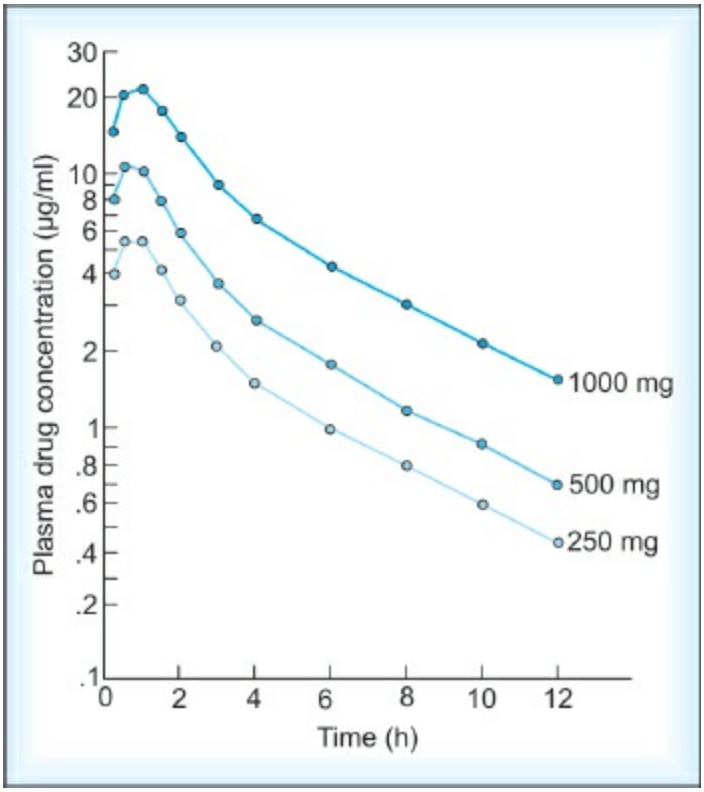



**Fig. 10.34:** Typical plasma concentration profiles after oral administration of single doses of a drug

**Table 10.4:** Analytic results for a subject in a typical bioavailability study comparing single 250, 500, and 1o0o mg oral tablet doses of a drug

| Plasma drug concentration (μg/mL) | | | | | Urinary drug excretion (mg) | | |
|---|---|---|---|---|---|---|---|
| Time(h) | 250 mg | 500 mg | 1000 mg | Time (h) | 250 mg | 500 mg | 1000 mg |
| 0 | 0 | 0 | 0 | –2 to 0 (predrug) | 0 | 0 | 0 |
| 0.25 | 4.0 | 7.9 | 14.7 | 0–2 | 66.2 | 136.3 | 256.1 |
| 0.5 | 5.4 | 10.7 | 20.8 | 2–4 | 32.7 | 63.2 | 137.3 |
| 1.0 | 5.4 | 10.1 | 21.7 | 4-6 | 18.5 | 35.5 | 77.6 |
| 1.5 | 4.2 | 7.8 | 17.9 | 6–8 | 12.7 | 24.1 | 53.4 |
| 2.0 | 3.2 | 5.9 | 14.1 | 8–12 | 15.5 | 28.8 | 65.4 |
| 3.0 | 2.1 | 3.7 | 9.1 | 12–24 | 13.8 | 24.7 | 59.7 |
| 4.0 | 1.5 | 2.7 | 6.7 | 24–36 | 1.8 | 3.0 | 8.1 |
| 6.0 | 1.0 | 1.8 | 4.4 | 36–48 | 0.27 | 0.30 | 1.1 |
| 8.0 | 0.71 | 1.2 | 3.1 | | | | |
| 10.0 | 0.51 | 0.87 | 2.2 | Total | 161.5 | 315.9 | 658.7 |
| 12.0 | 0.36 | 0.61 | 1.6 | | | | |

**Table 10.5:** Summary of results for a subject in a dose-proportionality/bioavailability study comparing 250, 500, and 1000 mg oral tablet formulations with a 500 mg intravenous dose as the standard

| | | $D^{po}$ | | | $D^{IV}$ | Source or method of calculation |
|---|---|---|---|---|---|---|
| Parameter | Units | 250 mg | 500 mg | 1000 mg | 500 mg | |
| $C_{max}$ | μg/ml | 5.4 | 10.7 | 21.7 | — | Table 10.4 |
| $t_{max}$ | h | 0.5 | 0.5 | 1.0 | — | Table 10.4 |
| $t_{½}$ | h | 4.08 | 3.88 | 4.10 | 3.89 | – 0.693/terminal slope |
| $CL_r$ | mL/min | 125 | 135 | 120 | 137 | Eq. (41): $t_1 = 0, t_2 = 12$ |
| $AUC_\infty$ | μg · h/ml | 21.5 | 39.0 | 91.5 | 43.7 | Eq. (43): $U_\infty CL_r$ |
| $U_\infty$ | % Dose | 64.6 | 63.2 | 65.9 | 72.1 | Table 10.4 |
| $F$ | % Dose | 98.4 | 89.2 | 105 | | Eq. (115) |
| | | 89.6 | 87.7 | 91.4 | | Eq. (116) |
| | | 91.8 | 88.0 | 95.0 | | Eq. (119) |

The renal clearance of the drug is calculated by Eq. (41), where $t_1$ and $t_2$ are times corresponding to the beginning and the end of a urine collection interval; the area-under-the plasma concentration-time curve over the same time interval, $\int_{t_1}^{t_2} C_1$, is evaluated by the trapezoidal method or by other suitable methods of interpolation. The average $CL_r$ observed in each oral dose treatment (i.e. 0 to 12 h) is given in Table 10.5. Note that by design, urine collections are segmented into five discrete intervals over the time plasma is

being sampled (Table 10.4), and plasma samples are included at the beginning and end of each urine collection period. Eq. (41) can then be used to calculate $CL_r$ for each of these intervals to compare these "incremental" values of $CL_r$ with the average value. In the present example, $CL_r$ is constant within a treatment (calculations not shown), and thereby is also independent of the plasma concentration of the drug. Typically, there is some variation in mean $CL_r$ between treatments (Table 10.4).

Since $CL_r^{po}$ is constant within a treatment, the total area under the curve, $AUC_\infty^{po}$, can be obtained from the ratio of total urinary recovery, $U_\infty^{po}$ and $CL_r^{po}$, as given in Eq. (43). This method is model-independent and does not require extrapolation based on $t_{½}$ of the plasma profile after the last data point (12 h) to obtain the total area under the curve. Dose-proportionality is observed with the three oral formulations because $AUC_\infty^{po}$ aproximately doubles when the administered dose is doubled (Table 10.5).

Bioavailability estimates of the drug calculated under the three alternative assumptions expressed in Eqs (115), (116), and (119), regarding *CL* between treatments, are given in Table 10.5. Without evidence to the contrary, estimates based on Eq. (119) may be preferred on the basis that plasma clearance in the oral dose treatments is adjusted only for the portion that is experimentally observable, the renal clearance. The bioavailability of the drug, F, is independent of the dose, and the drug is nearly quantitatively absorbed over the 250 to 1000 mg dosage range.

To obtain the rate of absorption, the intravenous dose data (*see* Table 10.5) are used as the reference standard to define the disposition of drug, and the absorption profile is constructed based on the Loo-Riegelman method, as given in Eqs (97) through (104). The absorption profiles for the oral dose treatments are given in Table 10.6. For this calculation, $AUC_\infty^{po}$ values were estimated by the trapezoidal method, and *CL* was adjusted for the observed change in $CL_r$ between the IV and p.o. routes of administration consistent with the estimate of *F*. Absorption of the drug is rapid and is nearly over by 6 h after the dose has been administered, by 12 h, the total amount absorbed is nearly identical to that calculated by the model-independent methods (*see* Table 10.5).

**Table 10.6:** Absorption profiles for drug after single oral doses of 250, 500,

and 1000 mg

| Time (h) | Cumulative amount of drug absorbed, A(t) | | | | | |
|---|---|---|---|---|---|---|
| | 250 mg | | 500 mg | | 1000 mg | |
| | mg | % Dose | mg | % Dose | mg | % Dose |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0.25 | 96.5 | 38.6 | 191.2 | 38.2 | 354.1 | 35.4 |
| 0.5 | 150.6 | 60.2 | 300.2 | 60.0 | 574.5 | 57.4 |
| 1.0 | 202.2 | 80.9 | 389.6 | 77.9 | 794.4 | 79.4 |
| 1.5 | 217.4 | 87.0 | 418.8 | 83.8 | 881.0 | 88.1 |
| 2.0 | 222.8 | 89.1 | 429.2 | 85.8 | 917.0 | 91.7 |
| 3.0 | 228.6 | 91.4 | 436.3 | 87.3 | 941.2 | 94.1 |
| 4.0 | 228.5 | 91.4 | 437.5 | 87.5 | 947.7 | 94.8 |
| 6.0 | 229.3 | 91.7 | 440.4 | 88.1 | 952.8 | 95.3 |
| 8.0 | 229.5 | 91.8 | 438.8 | 87.8 | 952.1 | 95.2 |
| 10.0 | 229.7 | 91.9 | 439.1 | 87.8 | 950.8 | 95.1 |
| 12.0 | 229.7 | 91.9 | 439.2 | 87.8 | 951.2 | 95.1 |

# 11:
# Pharmaceutical Statistics and Optimization

Statisticians have become familiar faces in the research laboratories of industrial pharmaceutical companies. Part of this recent upsurge has been due to the recognition by research scientists that the application of statistics can be a useful tool in the design, analysis, and interpretation of experiments. Also, governmental agencies, principally the Food and Drug Administration (FDA) have promulgated rules and regulations that virtually necessitate the application of statistical techniques to fulfill the law or its recommendations.

The “Good Manufacturing Practices” (GMPs) and the “Good Laboratory Practices” (GLPs) are more recent examples of FDA regulations that recommend statistical usage, formally or implied, as part of the routine of careful implementation and record-keeping of research and manufacturing operations.

Statistical input in sampling and testing for quality control, stability testing, process validation, design of preclinical protocols, including statistical methods and appropriate statistical analysis of the resulting data, are applications routinely applied by the pharmaceutical industry to satisfy both internal requirements and FDA recommendations. Both statistical treatment and interpretation of clinical data are requirements of all FDA submissions on new drugs or new dosage forms.

Applications of statistical techniques to pharmaceutical research are beginning to be more appreciated. Some statisticians and pharmaceutical researchers, alone or in collaboration, have had the foresight to recognize applications in this area, and recent publications in the areas of optimization and experimental design indicate the activity in formulation research.

Since this book deals with pharmaceutical technology, a chapter on statistics cannot be expected to be in any way complete. Statisticians should be consulted for all but the simplest problems. Data from real experiments almost always have unexpected wrinkles that need special consideration. The interaction should allow both the statistician and the pharmaceutical scientist

to learn and grow.

There will always be more than one way to examine, analyze, and interpret data. Statistics is not an exact science; given a set of data, two statisticians may analyze and present the results differently. Data that comes from a *well-designed experiment*, however, should lead the experimenter to the same conclusion, independent of the analyses. In fact, if this is not the case, something is a miss, and the data and/or analysis should be carefully reexamined.

Computations in the statistical analysis of data are not difficult, but can be complicated, especially for the novice. Statistical packages are readily available, and if the data are compatible with software programs, they should be used. Often, however, the analysis that should be used is not obvious except to the expert and frequently, there are situations in which only a custom-made program can fit the data. In addition, the statistician must sometimes improvise, using creativity and ingenuity for a situation for which an analysis has yet to be devised. Nevertheless, anyone who uses statistics should practice some of the computations. This allows a “hands-on” appreciation of what is involved, helping the user to learn more about what he is doing.

# INTRODUCTORY STATISTICAL CONCEPTS

Statistics encompasses many areas of human endeavor, and within each area is a scientific field in itself. The statistical techniques most applicable to pharmaceutical experimentation can be categorized broadly as follows:

1. Descriptive statistics: Presentation of data including tabular and graphic representation
2. Hypothesis testing
3. Estimation
4. Experimental design

1. Although this chapter is mainly concerned ‘with the latter three categories, *descriptive statistics* are important. Professional statisticians have expertise in this area, and collaboration between pharmaceutical scientist and statistician should be complementary, using to full advantage the knowledge and experience of both parties. One should carefully consider various alternatives, presenting “pictures” of the data in the most convincing manner.
2. The possible overuse of *hypothesis testing* in statistics has been debated, but at present it is certainly a popular standard statistical approach to decision making. Government agencies, the FDA in particular, rely on significance levels (5%) for drug-related claims.
3. *Estimation* of unknown parameters such as a population mean or the slope of a line is often the objective of a statistical investigation. This concept is illustrated and discussed in some detail in this chapter.
4. Statistical *experimental design* pervades all areas, and some designs commonly used in pharmaceutical experiments are presented. Poor or inadequate design can invalidate an otherwise carefully conducted experiment.

Statistical analysis does not yield unequivocal answers, nor can the analysis redeem a poor experiment. Good experimental design is essential. The statistics presented here are concerned with probability. Observations by their very nature are variable. By “variable,” we mean that if an observation (an experiment) is repeated, the new outcome cannot be exactly predicted. Variability and unpredictability characterize most of our experience; here, the

uncertainty is harnessed and put to work.

## PROBABILITY DISTRIBUTIONS

Probability can be conveniently thought of as the proportion of times a value or range of values is observed after many observations. Observations can be categorized as discrete or continuous. A discrete observation is one of a countable finite number, such as (1) a tablet categorized as either out of limits or within limits or (2) the number of tablets in a bottle whose label indicates 100 tablets. A continuous observation is one that can be measured more and more precisely according to the sensitivity of the measuring instrument. Weights of tablets and blood pressure are continuous measurements. Most of the statistical problems in pharmaceutical research can be dealt with using the normal (continuous) and the binominal (discrete) *probability distributions*. The data that at least approximate these distributions are not mysterious as suggested by the foregoing examples. These are the ordinary data that are observed in real-life experiments: weight, blood pressure, intact drug in a formulation, dissolution, blood level of a drug, proportion of tablets out of specification, tablet weights, or number of defective bottles. By defining distributions of hypothetical data, inferences can be made about samples of real data that are observed in the laboratory, the production area, or the clinic.

A probability distribution can be visualized as a frequency distribution constructed from a large number of observations. For example, the weights of 200 tablets can be summarized in a frequency distribution, as shown in Table 11.1. The number of tablets that fall into a given interval is the frequency for that interval. The frequency distribution for the 200 tablets approximates the true distribution. The weights of an entire batch of 1,000,000 tablets can be thought of as the *universe*, or true underlying distribution, from which the 200 tablets were selected. In Table 11.1, the tablets are placed in a discrete number (12) of class intervals. If the intervals were made small enough, resulting in a large number of intervals of equal width, a plot of the proportion of tablets in the batch falling into each interval (as in a histogram) would result in a smooth curve as the distinction between intervals disappears. Such a curve might look like one of the normal curves shown in Figs 11.1 or 11.2.

**Table 11.1:** Frequency distribution of weights of 200 tablets

| **Interval*** | **Frequency (number** | **Proportion of tablets** |
|---|---|---|

| | of tablets) | in interval |
|---|---|---|
| 176–180 | 3 | 0.015 |
| 180–184 | 6 | 0.03 |
| 184–188 | 12 | 0.06 |
| 188–192 | 17 | 0.085 |
| 192–196 | 26 | 0.13 |
| 196–200 | 33 | 0.165 |
| 200–204 | 36 | 0.18 |
| 204–208 | 23 | 0.115 |
| 208–212 | 22 | 0.11 |
| 212–216 | 11 | 0.055 |
| 216–220 | 7 | 0.035 |
| 220–224 | 4 | 0.02 |
| **TOTAL** | **200** | **1.000** |

* Tablets included from the lower weight in interval up to, but not including, the higher weight in interval.

## Measures of Centrality and Spread

The arithmetic average, or *mean*, is the most common measure of the "center" of a distribution and is equal to $\Sigma X/N$. Another often used measure of centrality is the *median*. The median, also known as the $50^{th}$ percentile, is the value that splits the data in half, i.e., half of the observations are greater and half are less than the median value. With an even number of observations, the median may be considered to be the average of the two observations in the middle of the data set, the numbers having been ordered from lowest to highest. For example, if $N = 10$, the median is the average of the fifth and sixth values. If a distribution of data is perfectly symmetric (as in the normal curve), the median equals the mean. The *median* is often used to describe asymmetric distributions, e.g. the *median* income.

The *spread*, or dispersion of data, is commonly expressed as the *standard deviation*, $S = \sqrt{\Sigma(X_i - \overline{X})^2/(N-1)}$. The larger the value of S, the greater is the spread of the data.

Another common way of expressing the spread is the *range*, which is equal to the difference between the highest and lowest values in the data set.

The *coefficient of variation (C.V.)* is a measure of relative variation and is equal to the standard deviation divided by the mean, $\sigma/\mu$ or $S/\overline{X}$, the sample estimate.

The sum of the squared deviations of each value from the mean $\Sigma(X_i - \overline{X})^2$ divided by $(N-1)$ is called the *variance* ($S^2$). The square root of the variance is the standard deviation. The number of observations minus one $(N-1)$, is known as the *degrees of freedom*.

## Normal Distribution—A Continuous Distribution

Examples of two normal distributions are shown in Fig. 11.1. The reader may find it helpful to visualize these distributions as frequency distributions of tablet weights, as has just been described. The two distributions in Fig. 11.1 have certain similarities and differences. Both curves are symmetric about a central value designated as $\mu$, the mean, and both are bell-shaped. This shape indicates that most of the values in the distribution are near the mean, and as values are further from the mean, they are less prevalent. Although theoretically the data comprising a normal distribution can take on values between $\infty$ and $+\infty$, values sufficiently far from the mean have little chance of being observed. For example, if the mean weight of a batch of tablets were 200 mg, the chances of having a 100 or 300 mg tablet in a typical batch would be small.

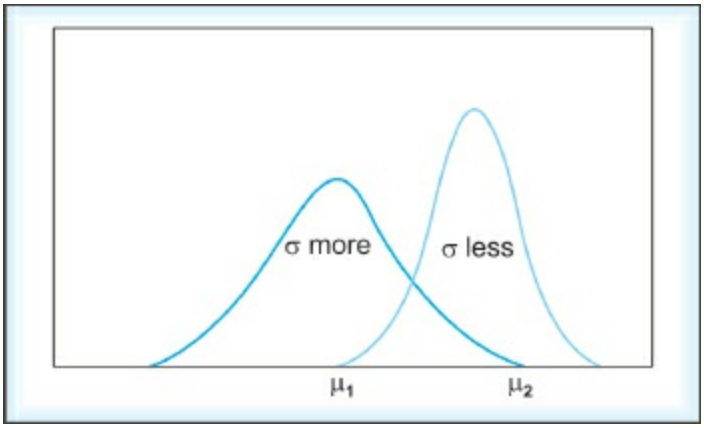



**Fig. 11.1:** Examples of two normal curves with different means and variances. The curve on the right has a smaller variance and a larger mean

In addition to normal curves being distinguished by the mean or central value, they differ in their "spread." The curve on the left-hand side of Fig. 11.1 is more spread out than the one on the right, a consequence of its larger standard deviation(s). Normal curves are exactly defined by two parameters: the mean, a measure of location, and the standard deviation, a measure of spread.

The Greek letters $\mu$ and $\sigma$ refer respectively to the mean and standard deviation of the universe or population. The population is the totality of data from which sample data are derived. In the example of 200 tablet weights, the

200 tablets are samples taken from a population of a batch of tablets that may consist of 1,000,000 tablets or more. In general, $\mu$ and $\sigma$ are unknown and are estimated from the data derived from the sample. The sample mean, or average, $\bar{X}$ is an *unbiased* estimate of the true population mean $\mu$. That is, although the average obtained from any single experiment cannot be expected to equal $\mu$, if an experiment is repeated many times, and all the $\bar{X}$'s are averaged, this overall average would equal $\mu$. The average is calculated as $\Sigma X_i$,/N, where $\Sigma X_i$, is the sum of $N$ data points. The average of the tablet weights in Table 11.1 is 200.32 mg. The average weight of the entire batch of tablets is probably not equal to 200.32 mg. The sample mean, $\bar{X}$ however, is the best estimate of the population mean, $\mu$.

Many statistical calculations involve $\Sigma(X_i - \bar{X})^2$ and this expression appears often in this chapter, with examples of the calculations. However, future use of summation notation in this chapter does not include the subscript *i*, although its use is implied.

*Example:* As a simple illustration of the calculation of the standard deviation, consider three data points with values 2, 4 and 6. The mean is 12/3 = 4. The value for $\Sigma(X_i - \bar{X})^2$ is calculated as follows:

| $X$ | $\bar{X}$ | $(X-\bar{X})$ | $(X-\bar{X})^2$ |
|---|---|---|---|
| 2 | 4 | –2 | 4 |
| 4 | 4 | 0 | 0 |
| 6 | 4 | 2 | 4 |
| | | | $\Sigma(X-\bar{X})^2 = 8$ |

The standard deviation of the numbers 2, 4 and 6 is $\sqrt{8/2} = 2$ . The variance $S^2$ is the square of the standard deviation and is equal to $2^2$ = 4 in this example. The sample variance, $S^2$ is an unbiased estimate of the population variance $\sigma^2$. The use of ($N$–1) in the denominator of the expression for $S^2$ ensures that the estimate is unbiased. A shortcut and more accurate computing formula for $\Sigma(X - \bar{X})^2$ is $\Sigma X^2 - (\Sigma X)^2/N$ equal $2^2 + 4^2 + 6^2 - \frac{12^2}{3} = 8$ in the above example.

Another property of the normal distribution is that the area under the normal curve as shown in Fig. 11.1, for example, is exactly one (1)

irrespective of the values, $\mu$ and $\sigma$. *The area between any two points* (Fig. 11.2), *e.g. 190 and 210, represents the probability of observing a value between 190 and 210*. Because the theoretic normal curve comprises an infinite number of values, the probability of observing any single value is zero. However, in many statistical procedures, there is a need to compute the probability of observing values in some interval. For example, suppose that the distribution of tablet weights approximates a normal distribution as shown in Fig. 11.2. The mean weight, y, of this batch of tablets is 200 mg, and the standard deviation, $\sigma$, is 10. The proportion of tablets weighing between 190 and 210 mg can be considered a measure of homogeneity of the batch and is equivalent to the probability of choosing a tablet at random that weighs between 190 and 210 mg.

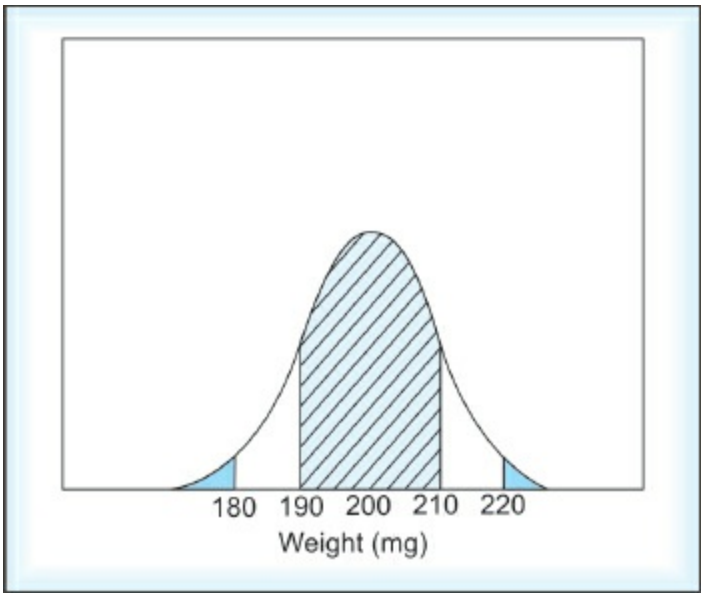



**Fig. 11.2:** Distribution of tablet weights with mean = 200, and standard deviation = 10. The small shaded areas in the tails each represent a probability of approximately 0.025.

This problem can be solved by referring to a table of "areas under the standard normal curve." Table 11.2 is a short version of such a table that gives the area between $-\infty$ and Z, where, *Z* is a transformation that changes all normal curves into the *standard normal curve*, which has a *mean* of 0 and a *standard deviation* of 1. This transformation allows the use of a single table, which can be used to calculate the area between any two points for any normal curve. The transformation is:

$$Z = (X - \mu)/\sigma$$

To calculate the area between $-\infty$ and $X$, compute $Z$ and find the area in Table 11.2. The reader may find it convenient to refer to Figs 11.2 and 11.3 in the following discussion, which describes the calculation for finding the area between 190 and 210.

1. Area between $-\infty$ and 210: $Z = (X-\mu)/\sigma = (210–200)/10 = 1$. Area between $-\infty$ and 210 = 0.84 from Table 11.2.
2. Area between $-\infty$ and 190: $Z = (190–200)/10 = -1$. Area between $-\infty$ and 190 = 0.16 from Table 11.2.
3. Area between 190 and 210 = (Area between $-\infty$ and 210) – (Area between $-\infty$ and 190) = 0.84–0.16 = 0.68.

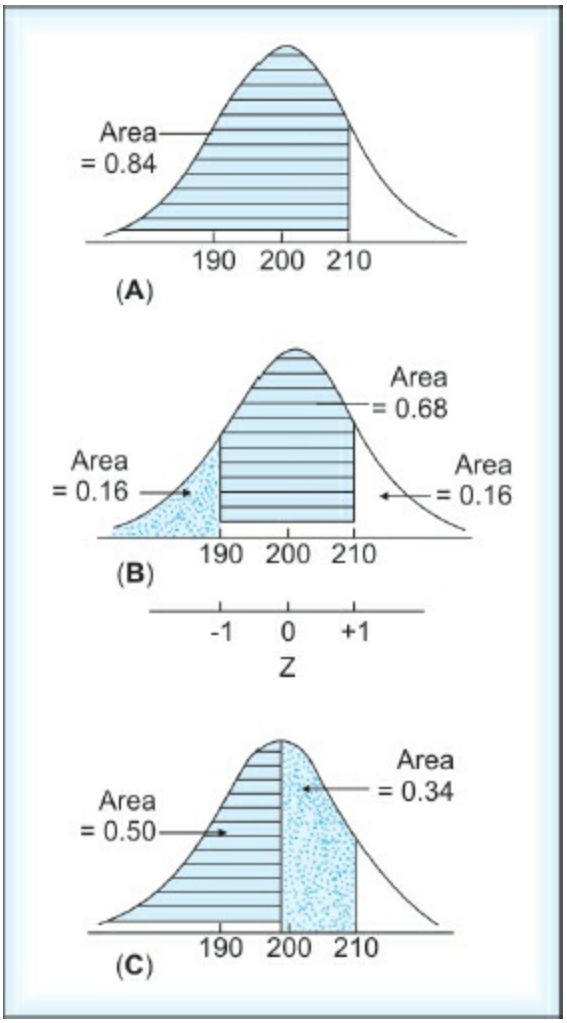
**Table 11.2:** Short table of cumulative areas under the standard normal curve

| $Z = (X–\mu)/\sigma$ | Area from $-\infty$ to $Z$ |
|---|---|
| -2.576* | 0.005 |
| -2.326 | 0.01 |
| -1.96* | 0.025 |
| -1.645* | 0.05 |
| -1.58 | 0.057 |
| -1.28 | 0.10 |
| -1.00 | 0.16 |
| -0.50 | 0.31 |
| 0 | 0.50 |
| 0.50 | 0.69 |
| 1.00 | 0.84 |
| 1.28 | 0.90 |
| 1.58 | 0.943 |
| 1.645* | 0.95 |
| 1.96* | 0.975 |
| 2.326 | 0.99 |
| 2.576* | 0.995 |

* Z = ±2.576, ±1.96 and ± 1.645 are the cutoff points used for a two-sided test at the 1%, 5% and 10% levels respectively. $Z = 1.645$ is the cutoff point for a one-sided test at the 5% level where, $H_0$ is $\mu \leq \mu_0$ and $H_A$ is $\mu > \mu_0$. $Z = -1.645$ is the cutoff point at the 5% level for a one-sided test where $H_0$: $\mu \geq \mu_0$ and $H_A$: $\mu < \mu_0$.

This computation is illustrated in Figs 11.3A to C.

The calculations may also be made based on the symmetry of the normal curve; the portion of the curve below the mean is the mirror image of that above the mean, and therefore the area above and below the mean each equal 0.5 by definition. In this case, the area between 210 and 200 (μ) is 0.84–0.5 = 0.34. Therefore, the area between 190 and 210 is 2 × 0.34 = 0.68 (see Fig. 11.3C). Thus, 68% of the tablets weigh between 190 and 210 mg, or equivalently, the probability of choosing a single tablet at random that weighs between 190 and 210 mg is 0.68 (chances are 68/100 that a tablet randomly taken from the batch will weigh between 190 and 210 mg).




**Figs 11.3A to C:** Calculation of areas under a normal curve with $\mu$ = 200 and $\sigma$ = 10 from table of areas under the standard normal curve (*see* Table 11.2)

What are the chances of finding a tablet chosen at random that weighs 10% more or less than the mean weight? This question is equivalent to asking, "What is the probability that a tablet will weigh less than 180 mg or more than 220 mg?" As before, calculate $Z$ for $X$ = 220 and $Z$ for $X$ = 180. $Z_{220}$ = (220–200)/10 = 2; the area is approximately 0.975. $Z_{180}$ = (180–200)/10 = −2; the area is approximately 0.025. The area between 180 and 220 is 0.975–0.025 = 0.95. Therefore, 5% of the tablets will weigh above 220 or less than 180 mg. By using the area between 180 and −∞ (0.025), and the symmetric properties of the normal curve, the same result is obtained (*see* Fig. 11.2).

Another, slightly different question is also germane: "What is the

probability that the *mean of 10 randomly chosen tablets* is between 195 and 205 mg?" An important result in statistical theory that relates to the distribution of *means*, the ***central limit theorem***, must be considered before this question can be answered. This theorem states that for any probability distribution (with finite variance), the distribution of means of $N$ randomly selected samples will tend to be normal as $N$ becomes large. Thus, if a new distribution is formed, based not on the original data but on means of size $N$ drawn from the original data, the distribution of means will tend to be normal. This consequence of the central limit theorem allows a comfortable application of normal curve theory when dealing with means. A natural question is, "What is meant by large N?" If the original data are normal, then means of any size $N$ will be normal. The sample sizes needed to make means from nonnormal distributions close to being normally distributed vary depending to a great extent on how aberrant the distribution of the original data is from a normal distribution. A good guess is that the means of sample sizes of 30 or more for the kind of data that are usually encountered will closely approximate a normal distribution.

The distribution of means of $N$ observations will have the same mean as the original distribution, but the variance will be smaller, equal to $\sigma^2/N$ where $\sigma^2$ is the variance of the original distribution. The means of many samples of size 10, for example, will tend to cluster more closely together than the individual data because extreme values will be compensated by the other sample values composing the mean. Therefore, the mean of samples of size 10 from the batch of tablets will have a normal distribution with a mean weight, $\mu$, of 200 mg and a standard deviation, $\sigma$, equal to $\sqrt{\sigma^2/N} = \sqrt{100/10} = 3.16$. The value $(\text{or } S_{\bar{x}} = S/\sqrt{N}$, the sample estimate) is commonly known as the ***standard error of the mean***. Figure 11.4 shows how the distribution of means ($N$ = 10) compares with the original distribution.

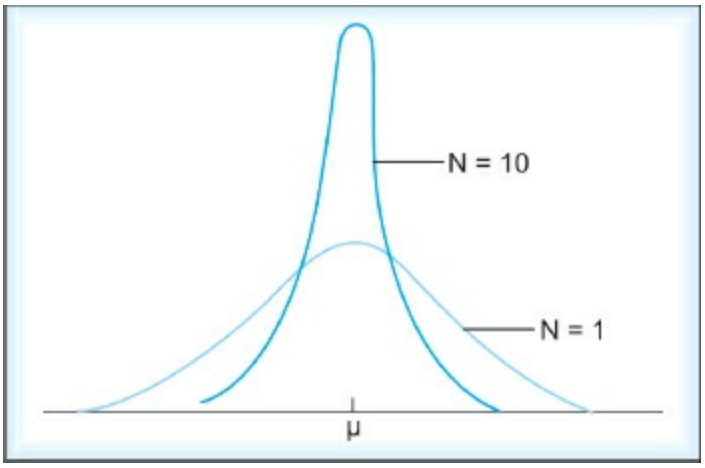



**Fig. 11.4:** Normal distribution of single observations ($N$ = 1) and means of size 10 ($N$ = 10)

To answer the foregoing question, the probability of the mean being less than 195 is calculated, as before, with $\mu$ = 200 and with $\sigma$ = 3.16. Thus, Z = (195–200)/3.16 = −1.58. From Table 11.2, the probability of $X$ being between −∞ and 195 is 0.057. Using the symmetry of the normal curve, the probability of an average weight of 10 tablets being above 205 is also 0.057 and the probability of being below 195 or more than 205 is 0.057 + 0.057 = 0.114. Therefore, the probability that the mean of ten tablets will be between 195 and 205 mg is (1–0.114) = 0.886.

The distribution of tablet weights cannot be identically equivalent to a normal distribution because there is some upper limit on tablet weight, and the lower limit must be greater than zero. Usually, real-life data does not conform exactly to theoretic distributions, and an exact analogy of the normal curve does not exist in real examples. This does not mean that we cannot make practical use of this well-known bell-shaped symmetric distribution. Much of the raw or transformed data encountered in pharmaceutical sciences are close enough to normal distributions to allow adequate treatment, as if such data were normal. Usually, insufficient data are available to define the probability distribution exactly since only a sample, a small part of the totality of data, is available. There are quantitative methods of assessing if a sample of data is likely to belong to some known distribution, e.g. the normal distribution; however, in this presentation, the true distribution underlying the relative small data sets that are often subjected to analysis will be assumed to be known, as a result of experience or other available information.

Data plotted in various ways, pictures of the data, are revealing, not only to help visualization of the possible underlying distribution from which the data are drawn, but also to obtain some insight into the meaning and interpretation of the data. A histogram, or bar graph, is a useful way of displaying data. Figure 11.5 shows an example of a histogram of assays of a sample of 100 tablets from a pilot batch. The number of tablets per mg interval is plotted such that the area of each bar is proportional to the number of tablets in the interval. Although the distribution of these tablets is slightly skewed to the right, to say that the distribution of tablets in this batch is approximately normal would not be unreasonable.

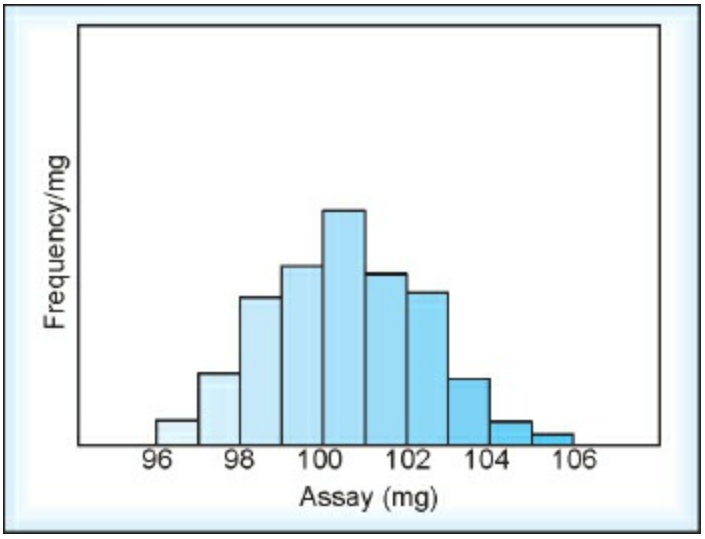



**Fig. 11.5:** Histogram of assays of a sample of 100 tablets

## Binomial Distribution-A Discrete Distribution

The normal distribution is an example of a continuous probability distribution. Although never exactly fitting this distribution, experimental data often approximate normality and inferences based on this assumption are reasonably accurate. Often, however, data are clearly not continuous, and other distributions must be found to accommodate these situations. The binomial distribution is an example of a discrete probability distribution that can be used to describe the outcome of experiments common in the pharmaceutical sciences. It consists of dichotomous data, data that can have one of two possible outcomes. Examples of experiments with a binomial outcome are (1) preference for one of two formulations, (2) life or death (used to compute the $LD_{50}$), (3) acceptance or rejection of dosage units in quality control and (4) improvement or worsening after treatment with a drug.

The binomial is a two-parameter distribution (1) $p$, the probability of one of the two possible outcomes and (2) $N$, the number of trials or observations. If 100 tablets sampled from a batch have 5 rejects, $p$, the probability of rejection, is estimated as 0.05, and $N = 100$. Note that in this example, probability is equated with the proportion of rejects; the best estimate of the unknown probability is the sample proportion. If the entire batch were inspected, the *true probability* of a reject would equal the proportion of rejects in the batch.

An example of a binomial probability distribution for $N = 18$ and $p = 0.3$ is shown in Fig. 11.6 and Table 11.3. The probability of observing *exactly* 5 successes in 18 trials, for example, is 0.202. A "success" is one of the two possible outcomes of a single binomial trial with probability $p$, and a "failure" is the other outcome (probability of a failure = $1 - p = q$). In this example, $p = 0.3$ and $q = 0.7$. The sum of the probabilities of all $N + 1$ possible results from $N$ binomial trials is one. There is a discrete number of possible results, $(N + 1)$ in $N$ trials. In 18 trials, 0, 1, 2... or 18 successes are possible, but the probability of more than 10 successes is extremely small if $p = 0.3$. If a batch of tablets were assumed to have 30% defects ($p = 0.3$), the probability of observing 10 or more defective tablets (successes) in a random sample of 18 ($N$) tablets would be so small that such an observation would probably lead to the conclusion that the batch really had more than 30%

defects. The distribution shown in Fig. 11.8 looks some-what symmetric. In fact, if $Np$ and $N(1 - p) = Nq$ are both greater than or equal to 5, the distribution is close enough to normal to allow use of normal curve probabilities to approximate the probability of binomial results.

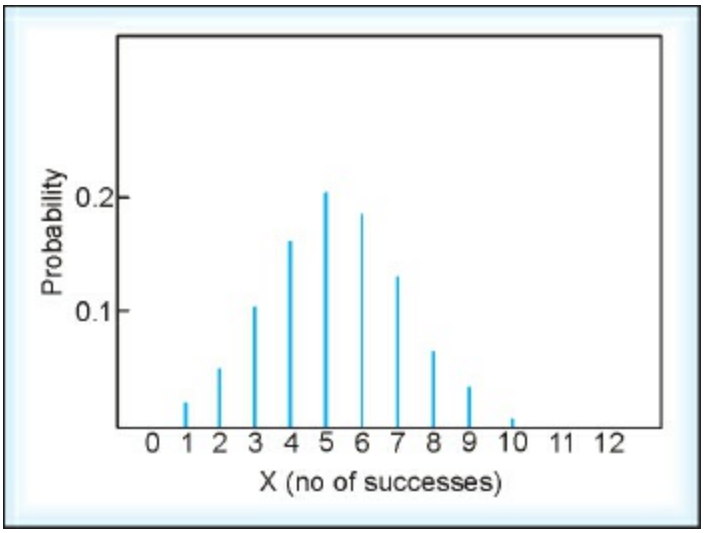



**Fig. 11.6:** Binomial distribution with $N = 18$ and $p$ 0.3

**Table 11.3:** Binomial probabilities for $N = 18$ and $p = 0.3$

| Number of successes | Probability |
|---|---|
| 0 | 0.002 |
| 1 | 0.013 |
| 2 | 0.046 |
| 3 | 0.105 |
| 4 | 0.168 |
| 5 | 0.202 |
| 6 | 0.187 |
| 7 | 0.138 |
| 8 | 0.081 |
| 9 | 0.039 |
| 10 | 0.015 |
| 11 | 0.005 |
| 12 | 0.001 |

As is the case for the normal distribution, the binomial can be parsimoniously presented in terms of its mean and standard deviation. The mean corresponds to the true probability of success, $p$, and the standard deviation (a function of $p$ and $N$) is, $\sigma = \sqrt{pq/N}$ where, $q = (1 - p)$. To compute probabilities of events using the binomial, one can refer to binomial tables, using the normal approximation when appropriate (*see* next section, "Normal Approximation to the Binomial Distribution"), or calculate probabilities using the binomial formula. The binomial formula gives the probability of $X$ successes in $N$ trials.

$$P(X) = \binom{N}{X} P^X q^{N-X} \qquad \ldots (1)$$

where, $P(X)$ is the probability of observing $X$ successes in $N$ trials and $\binom{N}{X} = N!/(N-X)!(X!)$.

*Example:* As an example of the use of the binomial formula, consider the following. According to the USP weight variation test for tablets weighing less than 130 mg, no single tablet out of 20 tablets weighed should differ by more than ±20% from the average weight. For a batch of 100 mg tablets, suppose 3% weigh less than 80 mg or more than 120 mg. What is the probability of finding at least one aberrant tablet in a random sample of 20 tablets? If the average weight is 100 mg, the probability of finding one or more bad tablets equals (1 – probability of finding 0 bad tablets out of 20). This probability equals:

$$\left[1 - \binom{20}{0}(0.03)^0(0.97)^{20}\right] = 0.46$$

(Note that $\binom{20}{0} = 20!/20!0!$ = 1 because 0! equals 1 by definition)

Thus, there is a 46% chance that one would find at least one bad tablet in this test if 3% of the batch were outside the limits.

## *Normal Approximation to the Binomial Distribution*

If $Np$ and $Nq$ are both equal to or greater than 5, cumulative binomial

probabilities can be closely approximated by areas under the standard normal curve. The following example illustrates this concept.

*Example:* It is known from past experience that the incidence of a specific malignant tumor is 10% during the lifetime of a certain strain of normal rats. A drug is administered to 50 rats, and the tumor occurs in 9 (18%) of the animals. Is this an unlikely event if the drug is not carcinogenic? The probability of observing 9 or more afflicted animals can be computed if indeed the normal rate of 10% has not changed. The normal approximation can be used since both Np and Nq are equal to or greater than 5 ($N$ = 50, $p$ = 0.10, $q$ = 0.90). Note that $p_0$, the hypothetical population value of $p$, is used for this calculation. The probability is computed using a normal curve with mean 0.10 and $\sigma = \sqrt{(0.1)(0.9)/50}$, calculating the area for values greater than 0.18.

$Z = (\overline{X} - \mu)/\sigma_{\overline{X}} = (p - p_0)/\sqrt{p_0 q_0 / N}$ (where $p$ is the observed proportion) = (0.18 – 0.10)/ $\sqrt{(0.10)(0.90)/50} = 1.89.$

The interpretation depends on whether a one-or two-sided test is used. The one-sided test would be appropriate if the drug cannot truly decrease the proportion of cancerous events below 10%; that is, an observation of a tumor incidence of less than 10% in the 50 rats would only be due to chance. A two-sided (two-tailed) test allows for the possibility of results both greater and smaller than 10%, suggesting that the drug might improve the carcinogenic profile. Since the value of Z is 1. 89, less than 1.96, a two-sided test would fail to reach significance at the 5% level (*see* Table 11.2). Since Z is greater than 1.65, a one-sided test would be significant at the 5% level; the drug increases the tumor incidence. This example shows clearly that the choice of either one- or two-sided tests should be seriously considered and justified a priori.

The calculation of $Z$ as described above approximates cumulative binomial probabilities, estimating the probability of 9 or more events in 50 animals if p = 0.10. A “continuity” correction suggested by Yates improves the normal approximation resulting in less significance. The correction consists of subtracting 1/2N from the absolute difference of the numerator of $Z$. For a two-sided test, Conover describes an improved correction. If the fractional part of $| N_p - N_{p0}|$ is greater than 0.5 but less than 1.0, the fractional part is replaced by 0.5. If the fractional part is greater than 0 but less than or

equal to 0.5, the fractional part is deleted. If the value is an integer (i.e. the fractional part is 0), the value is reduced by 0.5, which is equivalent to the Yates correction. The adjusted value of $|N_p - N_{p0}|$ is divided by $N$ for the numerator of the $Z$ ratio.

In the example, $|N_p - N_{p0}|$ is equal to |0.18 × 50 – 0.10 × 50I= 4.00. According to the foregoing rule, the value is decreased by 0.50: 400 minus 0.50 = 3.50. The numerator of $Z$ is 3.50/50 = 0.07.

$$Z = 0.07 / \sqrt{p_0 q_0 / N} = 0.07 / \sqrt{(0.1)(0.9)/50} = 1.65$$

Now, a one-sided test is just significant at the 0.05 level.

## Other Common Probability Distributions

### *The t Distribution (Student's t Test)*

The *t* distribution is an extremely important probability distribution. This distribution can be constructed by repeatedly taking samples of size *N* from a normal distribution and computing the statistic. The ratio

$$t = (\overline{X} - \mu)/(S/\sqrt{N})$$

has a student's *t* distribution with $N - 1$ degrees of freedom (df), where *N* is the sample size. The value $\overline{X}$ is the mean of the *N* samples selected; μ is the mean of the underlying population distribution from which the samples were selected; and *S* is the sample standard deviation, $\sqrt{\Sigma(X - \overline{X})^2/(N-1)}$.

This formula for *t* is the same as that for *Z* for the normal distribution, where $Z = (\overline{X} - \mu)/(\sigma/\sqrt{N})$, except that for *t*, the sample standard deviation, *S*, is used in the denominator.

Practical examples of experiments in which data are derived from populations with a normal distribution are common place. The examples in which probabilities have been calculated thus far assume a prior knowledge of the parameters of the distribution, that is, the mean and variance are known. In the great majority of cases, population parameters are unknown. In fact, often the purpose of the statistical analysis is to estimate these unknown parameters based on the sample statistics. The sample mean and variance, $\bar{X}$ and $S^2$, are unbiased estimates of these parameters, and if the underlying distribution (or population) is normal, probabilities of events based on these estimates can be obtained from the t distribution in a manner similar to that described previously for the normal distribution.

The *t* distribution is similar in shape to the normal distribution (Fig. 11.7), but is more spread out with more area in the tails, the extremities of the curve comprising the smaller and larger values. There are any number of t distributions, each defined by the degrees of freedom. The mean of the t distribution is zero and the spread depends on the degrees of freedom. With 1 df ($N = 2$), the curve is spread out, but as the degrees of freedom increase, the t distribution becomes tighter, with less spread, approximating more closely the standard normal distribution. When df $=\infty$ (i.e. the standard deviation is known), the t distribution is identical to the standard normal distribution. The

*t* distribution is used to calculate probabilities when the standard deviation is unknown and estimated from the sample. The use of the *t* distribution is illustrated below in the testing of hypotheses involving continuous data derived from a normal distribution.

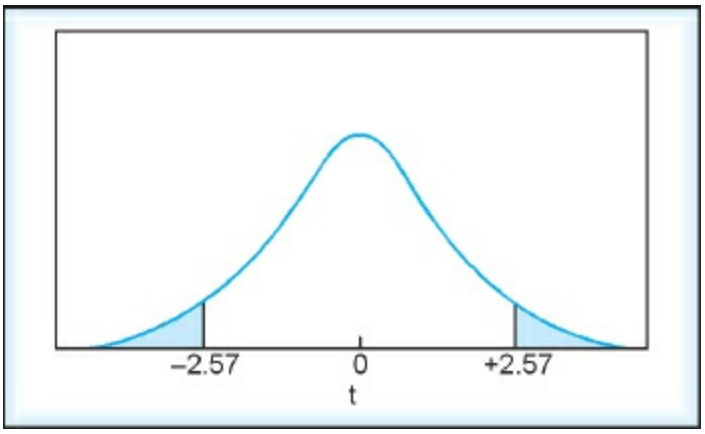



**Fig. 11.7:** Illustration of "*t*" distribution with 5 degrees of freedom. The area in each of the shaded portions at $t = \pm 2.57$ is 0.025

## *The $x^2$ (Chi-square) Distribution*

Another important probability distribution in statistics is the chi-square distribution. The chi-square distribution may be derived from normally distributed variables, defined as the sum of squares of independent normal variables, each of which has mean 0 and standard deviation 1. Thus, if $Z$ is normal with $\mu = 0$ and $\sigma = 1$,

$$\chi^2 = \sum Z_i^2$$

The chi-square distribution is appropriate where the normal approximation to the distribution of discrete variables can be applied. In particular, when comparing two proportions, the chi-square distribution with 1 d.f. can be used to approximate probabilities. The chi-square distribution is frequently used in statistical tests involving the comparison of proportions and counts.

*Example:* An example of testing for significance of the difference of two proportions is by means of the $X^2$ (chi-square) distribution is illustrated by a preclinical study performed to determine the carcinogenic potential of a new drug in which 100 control animals were compared to a group of 100 animals given the drug. At the end of the experiment, the animals were examined for

tumors. Ten animals in the control group and eight in the drug group died of nondrug related causes before the experiment was completed and these animals were not included in the final count. The results are summarized in Table 11.4.

**Table 11.4:** Actual and expected* number of animals with tumors after drug treatment and placebo

| | Number with tumors | Number without tumors | Total |
|---|---|---|---|
| Drug | 18(16.18) | 74 (75.82) | 92 |
| Placebo | 14 (15.82) | 76 (74.18) | 90 |
| **Total** | **32** | **150** | **182** |

* Values in parentheses are expected values

The first step in the statistical analysis is to compute the numbers of animals that would be expected to be observed in each of the four "cells" of the table. This is accomplished by multiplying the marginal totals for each cell and dividing by the grand total. For example, in the upper left cell (animals on drug with tumors) the expected number is (32 × 92)/182 = 16.18. Theoretically, this means that if the treatments were identical and no variation occurred, "16.18" of 92 animals would develop tumors in the drug group. For the fourfold table, only one calculation is needed to obtain the expected values for each cell, since the cell totals must sum to the marginal totals as shown in Table 11.4. The expected numbers in row 1 must add to 92; therefore, the value in the second column must be 75.82. Similarly, the value in the second row, first column, must be 15.82 to ensure that the column total is 32, and so forth.

*The chi-square statistic*, which is used to assess significance, is calculated as $\Sigma(O–E)^2/(E)$, where, $O$ is the observed count and $E$ is the expected count. In this example, chi-square is calculated as follows:

$$(1.82)^2/16.18 + (1.82)^2/75.82 + (1.82)^2/15.82 + (1.82)^2/74.18 = 0.50$$

The chi-square distribution is a probability distribution defined by a single parameter, degrees of freedom. The chi-square test can also be used for experiments other than that described by the 2 × 2 table. If three drugs are

being compared rather than two, a 3 × 2 table would describe the results for a dichotomous response. In an $R \times C$ (rows × columns) table, degrees of freedom equal $(R–1) \times (C–1)$. For the 2 × 2 table illustrated here, df = 1. The cutoff point for significance for chi-square with one df is equal to $Z^2$, where $Z$ is the standard normal deviate previously discussed (*see* Table 11.2). At the 5% level, for example, $X_1^2$ must exceed $(1.96)^2 = 3.84$ for significance, since 1.96 is the cutoff point for a two-sided test at the 5% level. Therefore, in the previous example, the difference between drug and placebo is not significant. (Cutoff points for $X^2$ tests with more than one d.f. can be found in most standard texts.) As is the case for the normal approximation to the binomial, the chi-square test is also approximate. The expected count ($E$) in each cell should be equal to or greater than 5 for the approximation to be valid. In this example, all four expected values are considerably greater than 5, as can be seen in Table 11.4.

As in the previously described binomial tests, a continuity correction can be used to improve the approximation. If the fractional portion of $|O–E|$ is greater than 0 but less than or equal to 0.5, delete the fractional portion (e.g. if $| O\text{-}E | = 5.3$, replace 5.3 by 5.0). If the fractional portion is greater than 0.5 but less than 1.00, replace the fractional portion by 0.5 (e.g. if $| O – E | = 6.98$, replace 6.98 by 6.5). If the difference $| O–E |$ is an integer, decrease the value by 0.5. In the above example, $| O–E |$ is 1.82. Therefore, use 1.5 as the value of $| O – E |$, rather than 1.82. The recalculated value of $X^2$ is 0.34.

In this experiment, a comparison of the total number of tumors found in the two groups may also be of interest when tumors may be found in different organs, i.e. when a single animal may have more than one tumor. The analyses discussed here, however, would be inappropriate because of the lack of independence. Suppose in the above experiment that 18 animals on drug had at least one tumor with a total of 33 tumors, and that 14 animals on placebo had at least one tumor with a total of 16 tumors. Clearly, one would have to look carefully at both statistics (1) the proportion of animals with tumors and (2) the total number, type, and location of tumors. However, the comparison of the total number of tumors, 16 versus 33, in the two groups must somehow take into account the number of animals with and without tumors.

## *The F Distribution*

After the normal distribution, the $F$ distribution is probably the most important probability distribution used in statistics. This distribution results from the sampling distribution of the ratio of two independent variance estimates obtained from the same normal distribution. Thus the first sample consists of $N_1$ observations and the second sample consists of $N_2$ observations:

$$F = \frac{S_1^2}{S_2^2}$$

The $F$ distribution depends on two parameters, the degrees of freedom in the numerator ($N_1$-1) and the degrees of freedom in the denominator ($N_2$-i). This distribution is used to test for differences of means (analysis of variance) as well as to test for the equality of two variances.

# Statistical Inference: Hypothesis Testing and Estimation

## *Hypothesis Testing (For Statistical Significance)*

Testing of hypotheses is a traditional use of statistical methodology. Because many people have been trained to understand the meaning of these tests, hypothesis testing serves to communicate clearly certain experimental conclusions from a statistical point of view. This statistical procedure is important for assessing differences in treatment effects for government submissions involved with new drugs and dosage forms, including results from clinical trials and bioavailability studies.

The process involves a hypothesis about one or more parameters of a statistical model. An example of a hypothesis is that the mean, $\mu$, of a population, e.g. a batch of tablets, is some value, $\mu_0$, perhaps 200 mg. A sample is chosen, and the estimate of the parameter $\bar{X}$ is calculated. If the sample estimate is close to the hypothesized value, the hypothesis is accepted. If the sample estimate or statistic is sufficiently far from the hypothesized value, the hypothesis is rejected. Rejection implies that the sample estimate of the parameter is evidence that the population parameter is different from that hypothesized. The conclusion in this case is that a statistically significant difference exists between the hypothesized parameter and the true parameter estimated from the sample.

To illustrate this concept, suppose that a hypothesis states that the average 90% dissolution time of a tablet batch is 30 min or less. A dissolution test using 12 tablets from a new batch shows an average result of 33 min. Can one conclude that the average dissolution time of the batch is greater than 30 min? If the average result of the 12 tablets were 50 min, the conclusion might be reached more easily. Depending on the individual tablet results, the decision would be more or less equivocal. Such decisions as whether to accept or reject the hypothesis, based on the sample data, can be made using probabilities derived from the *t* distribution. The procedure followed in making such decisions, as well as sample computations, are given here for some simpler problems; and the reader is encouraged to try as many computations as possible to gain insight into the statistical process.

The procedure of hypothesis testing is exemplified by a simple test that compares a sample mean with a hypothesized mean.

1. Initially, the *hypothetical mean* against which the mean of the sample data points is to be compared is defined. This is the *null hypothesis*. Statistically, the question being asked is, "Does the mean being estimated from the sample come from a distribution described by the hypothetical mean?" The null hypothesis is traditionally written as:

$$H_0: \mu = \mu_0$$

where, $\mu_0$ is the hypothetical mean. Some examples of hypotheses follow, (a) The target weight of a batch of tablets is 325 mg (the specification weight). ($H_0$: μ = 325 mg.) (b) An antihypertensive agent is hypothesized to lower the blood pressure by 10 mmHg on the average. ($H_0$: Δ = 10 mm Hg, where A is the blood pressure reduction.) (c) The specification for disintegration time of tablets is not more than 15 min. ($H_0$: $\mu \leq 15$ min.)

2. The *alternative* to the null hypothesis should be stated, i.e. the domain of answers if $H_0$ is not true. This may be written as $H_A$: $\mu \neq \mu_0$. In the previous example (a), the alternative could be written as $H_A$: μ ≠ 325 mg, i.e. the mean weight of the batch is either less than or greater than 325 mg. This is a two-sided alternative. A one-sided alternative, $H_A$: $\mu < 325$ mg, suggests that if the true mean weight is not 325 mg ($H_0$), only values of the mean less than 325 mg are relevant or possible.
3. Having defined the hypothesis and alternative, a ***level of significance*** **(the *α*, or Type I, error)** is chosen, the probability of erroneous rejection of the null hypothesis. The statement that a result is statistically significant, e.g. $P < 0.05$ or "significant at the 5% level," refers to the *α* error. A hypothesis tested at the 5% level means that if the test shows significance ($H_0$ is rejected), there is a 5% chance that the decision to reject $H_0$ is incorrect. In quality control plans, the *α* error can be thought of as the manufacturer's or producer's risk. For example, at $\alpha = 0.05$, a good batch of material is erroneously rejected 5% of the time.
4. An appropriate *sample* is chosen, and the mean and standard deviation are calculated. This is a simple concept, yet more complex than it seems at first glance. What is an *appropriate* sample? *How many* observations should one make? (How many times should the experiment be replicated?) How are samples chosen? Are the observations independent? If a sample of 20 tablets is chosen from a batch, the mean weight

represents the mean of the entire batch, perhaps a million or more tablets. A large burden is placed on these 20 tablets. In this situation, independence means that the selection and/or weighing of any one tablet is not influenced by or will not influence the weighing of the other tablets. Tablets should be selected at random or in a known "designed" way so that the statistical analysis will have a valid interpretation. One method of sampling might be to take the 20 tablets at regular intervals during the run. (Are there situations in which this method could lead to bias?) Selection of twenty consecutive tablets from any one part of a run would not be a good procedure. Why should 20 tablets be sampled rather than 10 or 30, for example? The sample size may be dictated by an official method, or by cost, or to obtain sufficient *power*. In the case of tablet weights, "power" might refer to the ability of the statistical test to correctly find a significant difference if the tablets are truly out of specification.

5. The next step in the hypothesis testing procedure is to *compute the t statistic*, which leads to a *decision* of whether to accept or reject the null hypothesis.

$$t = |\bar{X} - \mu| / (S\sqrt{N})$$

The value of t determines whether or not the null hypothesis will be rejected, rejection leading to a declaration of *significance*. If *t* is small, the hypothesis is accepted. If *t* is large, that is, if $|\bar{X} - \mu|$ is large compared to the standard error of $\bar{X}(S/\sqrt{N})$, then $\bar{X}$ is said to be significantly different from $\mu$. To make this decision, the computed value of *t* is compared to the value in a "*t*" table under $N - 1$ d.f. at the stated $\alpha$ level, usually 5% (Table 11.5). If the absolute value of *t* is greater than the tabled value, the difference $|\bar{X} - \mu|$ is statistically significant. Significance means that if the null hypothesis is true, the chances of observing a *t* value equal to or greater than that observed is less than *a*, or 5% in this case.

*Example:* The following example illustrates the hypothesis testing procedure just described. Table 11.6 shows the weights of 20 randomly selected tablets arranged in ascending order. The null hypothesis is $H_0$: $\mu = 325$ mg, and the alternative hypothesis is $H_A$: $\mu \neq 325$ mg. Since the sample size is 20, the degrees of freedom are 19 (20–1). Calculations of the mean, standard deviation, and t statistic follow. The shortcut formula for the

variance or standard deviation, previously illustrated, should always be used for speed and accuracy.

Also, when computing by hand, it is important to retain as many decimal places as possible.

$$H_0: \mu = 325 \text{ mg}; \quad H_A: \mu \neq 325 \text{ mg}$$

$$\overline{X} = (300 + 306 + \text{K } 336)/20 = 319.75 \text{ mg}$$

$$S^2 = \Sigma(X - \overline{X})^2/(N-1)$$

$$= [(300 - 319.75)^2 + \text{K } (336 - 319.75)^2]/19$$

$$= 67.2$$

*Shortcut formula for*

$$S^2 = [N(\Sigma X^2) - (\Sigma X)^2]/[N(N-1)]$$

$$= [20(300^2 + 306^2 + \ldots 336^2) - (300 + 306 + \ldots 336)^2]/(20)\ (19)$$

$$= [20(2046079) - (6395)^2]/380 = 67.2$$

$$S = \sqrt{S^2} = \sqrt{67.2} = 8.20$$

$$t_{19} = |\overline{X} - \mu|/(S/\sqrt{N})$$

$$= |319.75 - 325|(8.20/\sqrt{20})$$

$$= 2.86$$

From Table 11.5, for significance at $\alpha$= 0.05, $t_{19}$ must be greater than or equal to 2.09. Since the observed $t$ is 2.86, the decision is to reject $H_0$. Large values of $t$ lead to rejection of the null hypothesis; the test results in a decision of “significance”.

**Table 11.5:** Short table of absolute values of $t$ corresponding to various probability levels in a two-sided test

| Degrees of freedom (DF) | P = 0.01 | P = 0.02 | P = 0.05 | P = 0.10 |
|---|---|---|---|---|
| 1 | 63.66 | 31.82 | 12.71 | 6.31 |
| 2 | 9.93 | 6.97 | 4.30 | 2.92 |
| 3 | 5.84 | 4.54 | 3.18 | 2.35 |
| 4 | 4.60 | 3.75 | 2.78 | 2.13 |
| 5 | 4.03 | 3.37 | 2.57 | 2.02 |
| 7 | 3.50 | 3.00 | 2.36 | 1.89 |
| 10 | 3.17 | 2.76 | 2.23 | 1.81 |
| 11 | 3.11 | 2.72 | 2.20 | 1.80 |
| 12 | 3.06 | 2.68 | 2.18 | 1.78 |
| 13 | 3.01 | 2.65 | 2.16 | 1.77 |
| 15 | 2.95 | 2.60 | 2.13 | 1.75 |
| 19 | 2.86 | 2.54 | 2.09 | 1.73 |
| 25 | 2.79 | 2.49 | 2.06 | 1.71 |
| 40 | 2.70 | 2.42 | 2.02 | 1.68 |
| ∞ | 2.58 | 2.33 | 1.96 | 1.65 |

**Table 11.6:** Tablet weights (mg) of 20 randomly selected tablets

| | |
|---|---|
| 300 | 321 |
| 306 | 321 |
| 310 | 322 |
| 315 | 323 |
| 316 | 325 |
| 316 | 325 |
| 317 | 325 |
| 319 | 327 |
| 320 | 331 |
| 320 | 336 |

*Interpretation:* The conclusion based on this sample of 20 tablets is that the average batch weight is approximately 319.75 mg, which is sufficiently far from the target weight of 325 mg to declare a significant difference. In such an experiment, one can never be certain that this conclusion is correct. The statement that the difference is significant at the 5% level means that if, in fact, the batch mean were 325 mg, the probability would be small ($P < 0.05$, or less than 1 chance in 20) that a result as small as 319.75 or less (or

greater than 330.25) would be observed. The decision that the batch mean is not 325 mg may be incorrect, but the chance that this error will occur is 5% or less.

No matter how small or large the sample size is, a declaration of significance rings true. There is a small, but known, probability of erring in coming to the conclusion that a difference exists. On the other hand, a verdict of nonsignificance is not conclusive. A decision of nonsignificance can be virtually ensured by choosing a sufficiently small sample size, resulting in weak power, i.e. a weak ability to obtain a statistically significant difference. One should also be aware of the distinction between statistical and practical significance. Just as a small sample size can result in a large difference being statistically nonsignificant, a large sample size can result in a small difference being statistically significant, because the large sample size effectively reduces the variance ( $(S_{\bar{X}}^2 = S^2/N)$ ). In this example, the difference of 5.25 mg (325 – 319.75 mg) from the target weight is statistically significant. The important question is, “Is the difference apt to cause a problem from a therapeutic or regulatory point of view?” If the difference is not sufficiently large to reject the batch, the next or other future batches should be closely monitored. The use of the 5% significance level has no basis other than tradition and the fact that the risk associated with the 5% level seems reasonable; *one time in twenty the null hypothesis will be erroneously rejected*.

The assumptions implicit in the t test, both in this and in other examples are (1) that the data comes from a normal distribution, (2) that the observations are independent and (3) that the variance of the observations is equal (weighing error in this case). In general, independence and equality of variance are more stringent assumptions than normality. The central limit theorem helps to overcome nonnormality of data when using averages.

**Comparison of means of two independent samples:** An important class of problems in statistics is the comparison of the effects of two treatments or conditions on experimental outcomes. Examples of such experiments are (1) the comparison of the effects of two therapeutic treatments on blood pressure, (2) the comparison of two disintegrants on drug dissolution of tablets and (3) the comparison of two analytical methods. In these problems, the variance is usually unknown and is estimated from the data. If the experimental units are not related, i.e., are independent, the statistical test is known as an

independent two-sample *t* test.

In the case of a clinical trial in which two drugs are being compared, this statistical test would be applicable if all of the patients taking the two drugs are different. The null hypothesis and alternative hypothesis are:

$$H_0: \mu_1 = \mu_2; H_A: H_A: \mu_1 \neq \mu_2$$

The population means of the two treatments are assumed to be equal, and the null hypothesis is rejected only if the observed sample means are sufficiently different, based on the magnitude of the *t* statistic, which is computed as follows:

$$t = |\overline{X_1} - \overline{X_2}| / (S_p \sqrt{1/N_1 + 1/N_2})$$

where, $N_1$ and $N_2$ are the sample sizes associated with $\overline{X_1}$ and $\overline{X_2}$ respectively, and *t* has $(N_1 + N_2 - 2)$ degrees of freedom. The hypothesis of equal means is tested by comparing the value of the calculated *t* to tabled *t* values with appropriate df at the stated $\alpha$ level (Table 11.5).

To estimate $\sigma^2$, the common variance of the two groups, the variance of each sample is calculated, and the results are pooled, with the assumption that the true variances of the two groups are equal. (If the variances of the two groups are different, other statistical techniques are available). In general, to obtain an unbiased estimate of $\sigma^2$ when independent estimates of the common variance are available, a weighted average of the variance estimates is calculated with df $(N-1)$ as the weights. In the case of two samples, the weighted average is:

$$S_p^2 = \frac{(N_1 - 1)S_1^2 + (N_2 - 1)S_2^2}{(N_1 - 1) + (N_2 - 1)}$$

which equals:

$$\frac{[-\Sigma_1(X - \overline{X})^2 + \Sigma_2(X - \overline{X})^2]}{(N_1 + N_2 - 2)}$$

where, the subscripts 1 and 2 refer to treatments 1 and 1 respectively. Note the following:

$$(N-1)S^2 = (N-1)[\Sigma(X - \overline{X})^2(N-1)] = \Sigma(X - \overline{X})^2$$

*Example:* The following example illustrates the independent two-sample *t* test. Two batches of tablets were prepared using disintegrating agents A or B.

Dissolution was determined on randomly selected tablets with the following results: Disintegrant A: 45, 50, 47, 43, 41, 49, 35; $\bar{X} = 44.29$ and $S_2 = 26.9$. Disintegrant B: 38, 47, 42, 39, 32, 36, 41; $\bar{X} = 39.29$ and $S^2 = 22.57$. The variances are similar and are pooled to obtain an estimate of the common variance.

$$S_p^2 = [6(26.90) + 6(22.57)] / 12 = 24.74$$

$$t = |44.29 - 39.29| / \sqrt{24.74\,(1/7 + 1/7)} = 1.88$$

The tabulated $t$ at the 5% level with 12 df ($N_1 + N_2 - 2 = 12$) is 2.18 (Table 11.5). Therefore, the difference is *not significant at the 5% level*. One may conclude that although the data are insufficient to prove an effect due to disintegrants, the results are equivocal ($P < 0.10$), i.e. if the observed difference can be considered of practical significance, further testing is indicated. (Note that the tabulated $t$ for significance at the 10% level with 19 df is 1.73).

The sample sizes of the two groups need not be equal in this design, although maximum efficiency for discriminating two means is obtained by use of equal sample sizes, given a fixed total number of observations. Thus, $1/N_1 + 1/N_2$ is minimized, and the value of $t$ is thereby maximized.

**Comparison of means of paired samples:** Observations from two groups to be compared can often be paired in some natural way. The more alike the pairs are, the more precise the test is, and such pairing can be considered an example of a choice of experimental design. Rather than different experimental units being chosen at random for the two groups, pairs are chosen to be as alike as possible. Examples of pairing are (1) using the same individual for more than one treatment, (2) using twins or litter-mates for two or more treatments and (3) providing similar samples of granulation for comparison of assay methods. If a difference between two groups exists and the pairs are chosen judiciously, the difference will be detected more readily than if distinct individuals are randomly chosen for both treatments (given the same number of observations). Calculations for the statistical test comparing means of paired data consist of first taking differences of the pairs and then performing a one-sample t test on the differences. The null hypothesis is:

$$H_0: \mu_1 = \mu_2 \text{ or } \mu_1 - \mu_2 = \Delta = 0$$

A comparison of two analytic methods was made using five batches of material, and analysis was performed on each batch by each method. As shown in Table 11.7, the differences between methods for each batch is first computed, and the standard deviation of the differences is calculated, equal to 1.42. *N* is equal to 5 (there are 5 pairs). The calculated value of *t* is:

$$t = |0.50 - 0| / (1.42\sqrt{1/5}) = 0.79$$

Since the tabulated value of *t* with *4* df at the 5% level is 2.73, the difference is not significant at the 0.05 level.

**Table 11.7:** Comparison of two analytic methods

| | Method | | |
|---|---|---|---|
| | *mg activelg* | | |
| Batch | A | B | Δ A–B |
| 1 | 100.8 | 100.1 | 0.7 |
| 2 | 101.9 | 99.3 | 2.6 |
| 3 | 98.7 | 100.0 | –1.3 |
| 4 | 101.3 | 101.4 | –0.1 |
| 5 | 102.5 | 101.9 | 0.6 |
| Average | 101.04 | 100.54 | 0.50 |
| *S* | 1.46 | 1.07 | 1.42 |

**One-sided tests:** The tests described thus far have been two-sided tests. Values of the observed mean, or difference of means, that are either too small or too large lead to rejection of the null hypothesis. One-sided tests are used when the null hypothesis is rejected only for unidirectional differences of the observed mean(s). Care should be taken in using one-sided tests since this choice implies that the alternative (values either too high or too low) is not important or is of no interest. For example, in testing a new drug, $H_0$ might be rejected only if a positive effect is observed; however, if a negative response (opposite of the expected effect) is possible and of interest, a two-sided test might be more appropriate.

**Comparison of a sample proportion to a known proportion:** Quality control (QC) data gathered from many batches showed that 4% of tablets manufactured with a target weight of 200 mg weighed more than 220 mg or less than 180 mg, the upper and lower QC limits. Examination of a new batch

shows that 32 of 500 tablets (6.4%) are out of specifications. Is this result unexpected based on the previous history of the batch (4%, or 20 tablets, are expected to be out of limits)? As in the t test, the null hypothesis and the alternative hypothesis are stated. A "Z" ratio is then formed, using the continuity correction to compare the observed and hypothetical proportions:

$$H_0: p_0 = 0.04 \qquad H_A: p_0 \neq 0.04$$

$$Z = [|p - p_0| - 1/(2N)] / \sqrt{p_0 q_0 / N}$$

$$Z = (|0.064 - 0.040| - 1/1000) \div \sqrt{(0.04)(0.96)/500}$$
$$= 2.62$$

The new batch has significantly more tablets out of limits than are normally observed (Table 11.2, $P < 0.01$). A 95% confidence interval on the true proportion of out-of-limit tablets in this batch is:

$$0.064 \pm 1.96\sqrt{(0.64)(0.936)/500} = 0.064 \pm 0.021$$

* For 95% confidence limits, substitute 1.96 for 2.58.

**Comparison of two proportions:** When comparing the proportion of successes in two groups, the data are often presented in the form of a "fourfold" table as shown in Table 11.8. Diseased animals were treated with either placebo (control) or drug. Sixty-one of 75 of the control animals survived, whereas 69 of 75 animals given the drug survived. Is the drug more effective than the control in preventing death? The null and alternate hypotheses are stated as follows.

$$H_0: P_{drug} = P_{placebo}; H_A: P_{drug} \neq P_{placebo}$$

where, $P_{drug}$ is the probability that an animal will survive the drug treatment, and $P_{placebo}$ is the probability that an animal will survive placebo treatment.

**Table 11.8:** Four-fold table showing number of animals alive and dead after three months

| | Alive | Dead | Total |
|---|---|---|---|
| Control | 61 | 14 | 75 |
| Drug | 69 | 6 | 75 |
| **Total** | **130** | **20** | **150** |

This is the binomial analog of the independent groups two-sample *t* test.

In the *t* test, the variances were pooled under the assumption of equal variability in the two groups. Here, *all the data are pooled to estimate a common p*, the best estimate of the true probability under the null hypothesis, which states that the two populations have the same proportion of survivors. The pooled $p = p_0 = 130/150 = 0.867$ (130 of 150 animals survived and the overall proportion of survivors is 0.867). When the sample size (*N*) is equal in the two groups, the continuity correction in the test of significance consists of subtracting $1/N$ from the absolute value of the numerator. The *Z* ratio is computed as follows:

$$Z = [\,|p_1 - p_2| - 1/N]/\sqrt{p_0 q_0 (1/N_1 + 1/N_2)}$$

$$Z = [\,|0.92 - 0.813| - 1/75]$$

$$\div \sqrt{(0.867)(0.133)(1/75 + 1/75)}$$

$$= 1.68$$

where, $N_1 = N_2 = N$. As in the *t* test, the *Z* ratio is a "difference" divided by the variability of the difference, expressed as the standard deviation. In this case, the difference is significant at the 10% level, not significant at the usual 5% level. Does this mean that the drug and control do not differ? In all statistical tests, a nonsignificant difference does not imply sameness. The experiment may not have had sufficient sensitivity to pick up a difference that may have occurred with a larger sample size. A significant difference, no matter how small or large the experiment, means that a real difference probably exists (with odds of 19–1 at the 5% level). The practical meaning of any difference must be interpreted in context, with an understanding of the implications of decisions based on experimental results. In this example, if increase in survival due to the drug is small compared to comparable marketed drugs, there might be little interest no matter what the degree of significance. But if there is no drug on the market effective in this disease, these results might be a stimulus for further work.

**Comparison of variances:** In some experimental situations, a test of the equality of two independent variance estimates, $s_1^2$ and $s_2^2$ is of interest. In a one-sided test where $\sigma_1^2$ is hypothesized to be equal to or less than $\sigma_2^2 (H_0 : \sigma_1^2 \le \sigma_2^2; H_A : \sigma_1^2 > \sigma_2^2)$, the statistical test consists of forming the ratio $s_1^2/s_2^2$. If $s_1^2$ is smaller than $s_2^2$, $H_0$ is accepted as true. If $s_1^2$ is larger than $S_2^2$, a test of significance is performed using the *F* distribution. The ratio $S_1^2/S_2^2$ is

compared to tabulated values of the $F$ distribution with $\nu_1$ and $\nu_2$ degrees of freedom, where $\nu_1 = N_1 - 1$ and $\nu_2 = N_2 - 1$ ($N_1$ and $N_2$ are the sample sizes associated with $S_1^2$ and $S_2^2$, respectively). If the ratio exceeds the tabulated value at the specified $\alpha$ level, $\sigma_1^2$ is considered to be significantly greater than $\sigma_2^2$ The $F$ distribution is a probability distribution that is completely described by degrees of freedom in the numerator and denominator of the $F$ ratio (Fig. 11.8). The values tabulated in Table 11.9 are the upper cutoff points and are appropriate for a one-sided test $(H_0 : \sigma_1^2 \leq \sigma_2^2; H_A : \sigma_1^2 > \sigma_2^2)$ .

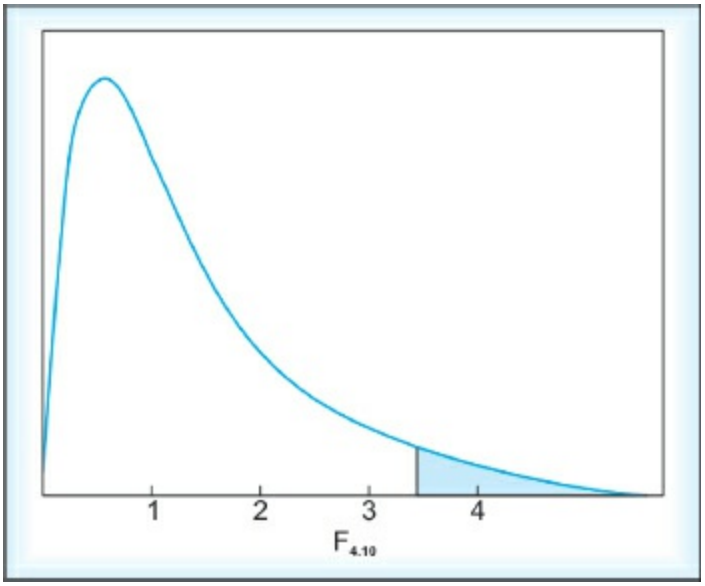



**Fig. 11.8:** Illustration of "*F*" distribution with 4 and 10 degrees of freedom. The shaded portion represents values from 3.48 to ∞ and equals 5% of the area

**Table 11.9:** Short table of upper points of the F distribution at 5% level of significance

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 3 | 10.10 | 9.55 | 9.28 | 9.12 | 8.94 | 8.89 | 8.79 |
| 4 | 7.71 | 6.94 | 6.59 | 6.39 | 6.16 | 6.09 | 5.96 |
| 5 | 6.61 | 5.79 | 5.41 | 5.19 | 4.95 | 4.88 | 4.74 |
| 7 | 5.59 | 4.74 | 4.35 | 4.12 | 3.87 | 3.79 | 3.64 |
| 8 | 5.32 | 4.46 | 4.07 | 3.84 | 3.58 | 3.50 | 3.35 |
| 10 | 4.96 | 4.10 | 3.71 | 3.48 | 3.22 | 3.14 | 2.98 |
| 13 | 4.67 | 3.81 | 3.41 | 3.18 | 2.92 | 2.83 | 2.67 |
| 20 | 4.35 | 3.49 | 3.10 | 2.87 | 2.60 | 2.51 | 2.35 |
| 40 | 4.08 | 3.23 | 2.84 | 2.61 | 2.34 | 2.25 | 2.08 |
| 120 | 3.92 | 3.07 | 2.68 | 2.45 | 2.18 | 2.09 | 1.91 |
| ∞ | 3.84 | 3.00 | 2.60 | 2.37 | 2.10 | 2.01 | 1.83 |

For a two-sided test $(H_0:\sigma_1^2=\sigma_2^2; H_A:\sigma_1^2\neq\sigma_2^2)$ form the ratio $S_1^2/S_2^2$ *If* $S_1^2$ is greater than $S_2^2$ if $S_2^2$ is the larger variance, form the ratio. That is, the ratio is formed with the larger variance in the numerator and is always greater than 1. The significance level using the cutoff points in Table 11.6 is 10% in this test because the ratio has been deliberately constructed with the larger variance in the numerator. In a twosided test, if the ratio $S_1^2/S_1^2$ is formed whether or not Sf is greater than Sf tables with the lower cutoff points in addition to the upper points are necessary. When the ratio is intentionally formed with the larger variance in the numerator, the *F* tables used to assess significance are simplified.

*Example:* An example follows to clarify the test procedure. The variances obtained from a formulation mixed in two different mixers will be compared. Five samples were analyzed from Mixer A with, a variance of 2; eight samples from Mixer B had a variance of 13. Does the use of the two mixers result in formulations of different homogeneity? This is a two-sided test, because a priori, it is not known which mixer will result in more variability, if a difference in variability exists.

$$H_0:\sigma_A^2=\sigma_B^2$$
$$H_A:\sigma_A^2\neq\sigma_B^2$$
$$\alpha=0.10$$

In a two-sided test, $F=S_B^2/S_A^2=6.5$ (where $S_B^2$ is the larger variance estimate). Since 6.5 is greater than the tabulated value ($F_{7,4,0.05}$ = 6.09), reject $H_0$ at the 10% level. If the variance from mixer B is known a priori to be no less than that in mixer A $(H_A:\sigma_B^2>\sigma_A^2)$, a one-sided test would be appropriate, and the difference would be significant at the 5% level.

## *Statistical Estimation (Confidence Intervals)*

A confidence interval can be formed for a mean, $\mu$, or the difference between two means, with confidence, 1-α, as follows:

$$D \pm (t_{\text{d.f.},\alpha})(S_{\bar{D}})$$

where, $D$ is a mean value or mean difference; $t_{\text{d.f.},\alpha}$ is the tabulated $t$ value with appropriate degrees of freedom (Table 11.5) at the $\alpha$ level of significance; and $S_{\bar{D}}$ is the standard error of the mean or mean difference.

This procedure is related to hypothesis testing in that *if the confidence interval covers zero,* the test of significance of the difference between two means is *not* significant at the *a* level (two-sided test). In the previous example, a 95% confidence interval for the mean difference between the two analytical methods is $0.50 \pm 2.78(1.42)(\sqrt{1/5})$ or −1.26 to 2.26. A confidence interval of $1 - \alpha$ (e.g. 95% if $\alpha = 0.05$) means that the probability is 95% that such intervals cover the true mean value, or difference of means. *There is no guarantee that in any single experiment the true value will be covered;* however, on the average, the true value will be included in the confidence interval of 19 out of every 20 such experiments. A lower degree of confidence is associated with a smaller interval. The smaller the interval is, the less is the confidence that the true value is contained within the interval. The confidence interval is a useful way of defining the region that contains the true mean or difference; it is a statement having a probability associated with it.

*Example:* In another example, a paired t test was used to analyze the data from a bioavailability study in which 12 volunteers were used to compare two dosage forms. The average areas under the time versus blood level curves were as follows: Formula A = 524; Formula B = 486. The variance computed from the 12 differences (A – B) was 94. A 95% confidence interval is calculated using the value 2.20 for *t* (*see* Table 11.5):

$$(524 - 486 \pm t_{11,0.05}\left(S\sqrt{1/N}\right)$$
$$= 38 \pm 2.20\left(\sqrt{94/12}\right) = 3.18 \text{ to } 44.2$$
$$\text{A 90\% interval is } 38 \pm t_{11,0.1}\left(S\sqrt{1/N}\right)$$
$$= 38 \pm 1.80\left(\sqrt{94/12}\right) = 33.0 \text{ to } 43.0$$

Some statisticians feel that in certain situations, such as bioavailability studies, confidence intervals are a better way of expressing results than stating only that one formula is or is not significantly different from the other.

By providing upper and lower limits for the comparison, the confidence interval allows the user or prescriber of a product to make a judgment about the practical equivalency of two products.

**Confidence limits in the binomial case:** Confidence limits can be constructed for binomial data in a manner similar to that for the normal distribution, as follows:

$$p \pm Z\sqrt{pq/N}$$

The value of $Z$ (*see* Table 11.2) depends on the degree of confidence. Suppose that of 1000 tablets inspected, 25 were found to be defective. The proportion of good tablets in the sample is 0.975 (975/1000). A 99% confidence interval for the true proportion of good tablets is:

$$0.975 \pm 2.58\sqrt{(0.025)(0.975)/1000} = 0.975 \pm 0.013$$

The width of the confidence interval is dependent on $p$ and $N$, the number of observations and is independent of the size of the batch, provided the number of observations is small relative to the batch size. A proportion (or any parameter for that matter) can be estimated with any desired precision by appropriately increasing the sample size. Realistically, time, expense and accuracy of observations are limiting factors.

# Sampling in Statistics

## *Choosing Samples*

*Samples* are usually a relatively small number of observations taken from a relatively large population or universe. An important part of the statistical process is the characterization of a population by estimating its parameters (any measurable characteristics of the population). Much of the work in statistics involves examining relatively small samples and then making inferences about the population from which the sample came. When designing an experiment, the population should be clearly defined so that samples chosen are representative of the population.

There are many reasons why the totality of data cannot always be observed or tested. For instance, 100% sampling may be precluded because of practical time and cost considerations. Situations in which 100% sampling cannot be accomplished practically are (1) destructive assays mat may occur in analytical procedures and (2) cases in which the population definition precludes 100% sampling, as occurs in clinical studies in which the population may include *all* patients with a particular disease.

The proper selection of samples is an essential part of a good experiment and is a consequence of the experimental design.

There are various methods of choosing samples may be roughly divided into *probability sampling* and *non-probability sampling* (sometimes called authoritative sampling).

Probability sampling methods can be of four types i.e. *random, stratified, systematic* and *cluster sampling*. Objects chosen to be included in probability samples have a known probability of being included in the sample and are chosen by some random device.

A *random sample* is one in which each of all possible experimental units have an *equal chance* of being included in the experiment or sample. Data derived from random samples yield fair, unbiased estimates of population parameters such as the mean. To sample tablets randomly from a batch for visual inspection or assay, each tablet should have an equal chance of being chosen. To obtain a random sample, a number can be conceptually assigned to each potential candidate, and then a table of random numbers can be used

to select those units or individuals to be included in the experiment or to be assigned to a particular treatment group. Alternatively, all possible experimental units can be thoroughly mixed (literally or figuratively), and samples can be chosen at random as in a lottery.

One must be aware that bias can easily be introduced into the selection of samples if care is not taken in randomization. Often, one must compromise between theory and practice. In a multicenter clinical trial where there is a choice of large numbers of patients from perhaps thousands of clinical sites, the sample chosen may be based more on convenience than randomness. If an effort is made, however, to choose sites based on relevant factors such as geographic location, for example, one might practically consider this a random sample. Certainly, choosing tablets for inspection cannot be done both conveniently and strictly at random. How would one identify the $1{,}565{,}387^{th}$ tablet if it was to be included in the sample? An alternative is to select samples stratified over various time periods during the production run. Ingenuity can often be used to devise a “pseudorandom” sampling procedure that can be satisfactory in difficult situations.

*Stratified sampling* is a procedure in which the population is divided into subsets or strata, and random samples are selected from each strata. Stratified sampling is a recommended way of sampling when the strata are very different from each other, but objects within each stratum are alike. Stratified sampling will be particularly advantageous when this within-object variability is small compared to the variability between objects in different strata. In quality control procedures, items are frequently selected for inspection at random within specified time intervals (strata) rather than in a completely random fashion (simple random sampling). Thus we might sample 10 tablets during each hour of a tablet run. *Systematic sampling* is often used as an improvement over random sampling. Every nth sample is chosen for inspection, testing, or analysis, which ensures a regular sampling throughout a process such as the manufacture of tablets. If the process is cyclic of periodic in nature, and, if by chance the Sampling corresponds to the period, this method of sampling can lead to erroneous conclusions.

*Cluster sampling* involves the division of population into groups or clusters each of which contain ‘ subunits’. For example, suppose that one wishes to inspect tablets visually, packaged in the final labeled container. The batch consists of 10,000 bottles of 100 tablets each. Cluster sampling, in this

example, might consist of randomly selecting a sample of 100 bottles (primary units), and then inspecting a random sample of 10 tablets (subsample units) from each of these bottles. Non-probability samples often have a hidden bias, and it is not possible to apply typical statistical methods to estimate the population parameters. Non-probability sampling methods often are convenient and simple and should not be used unless probability sampling methods are too difficult or too expensive to implement.

*Sampling for quality control* should be carefully implemented to ensure that the samples selected are representative. A representative sample can be regarded as one that is carefully chosen, and perhaps subsequently treated or modified (e.g. by compositing in the case of bulk powders) so that the sample has the same characteristics as the bulk material (powder) if it were homogeneous. For example, sampling from the top of a large container might not be representative and can result in samples that are different from those that might have been obtained from the main bulk of material. It is not easy to define exactly (statistically) a representative sampling scheme for testing bulk materials. The amount of material to be inspected, as well as the number and kind of samples to be chosen, is often based on experience and empirical rules. One method of sampling bulk raw materials from drums, for example, is to sample $\sqrt{N}$ containers (add one container if there is a remainder) where $N$ is the number of containers. The material may then be taken from different parts of the container using a *thief*, a sampling device that is basically a long hollow tube inserted into the powdered material. After collection, the various samples are often thoroughly mixed, and samples for assay are taken from this homogeneous mixture. If carefully implemented, such procedures for finished dosage forms or powdered raw materials and intermediates should result in representative samples for analysis.

## *Independence and Bias*

Many analyses and interpretation of data assume that the sample data points are independent of one another, i.e. that the result of one observation does not influence the result of another concurrent, future, or past observation. Examples of dependent or correlated observations are (1) blood pressure of the same individual taken on two or more occasions, (2) assays from the same small portion of material repeated over time and (3) responses to a flavor preference from the same individual on multiple occasions. In these cases, the same experimental Unit being used for more than one observation

is responsible for the dependence, i.e. each observation has a common component as part of its variability. Note that in (2) the nature of the data is different if the assay is performed on different portions of material rather than on a single homogeneous portion.

The same data may be treated as independent in one situation and correlated in another. For example, to assess the stability of a product, a single bottle from a batch is assayed over time with duplicate assays at each time period. Are the duplicates assayed at a given time period independent? If they are analyzed concurrently, the duplicates are probably not independent because of the common conditions existing at the time of the assay. These common conditions could include, for example, the same analyst, the same reagents, and the same instrument standardized with the same standard material. Duplicates performed at two different laboratories by two different analysts or assays done at different points in time are probably independent if proper care is taken. Since the tablets come from a single bottle in which the material is more homogeneous than the batch as a whole, can any of the assays be considered independent? These are rhetorical questions and serve to illustrate the complexity of a seemingly simple concept. If observations are correlated, special analyses may be necessary to account for the correlation.

*Bias* is another term common to statistics. If samples are not carefully chosen, bias, which may not be obvious, can easily be introduced. Human beings cannot always be objective, and if some controls are not imposed, what is observed is often what one wishes to observe. Try to make up a series of “random” numbers from 0 to 9 or letters of the alphabet quickly, without too much linking, and you will observe your own bias. Randomization and blinding are ways to overcome bias. In blinded experiments, the person observing and recording experimental results is not aware of the source of the data. This procedure is especially important in experiments with subjective evaluations. Some examples of how bias may enter an experiment may seem obvious to an objective outsider, but may often not be obvious to the scientist closely involved with his experiment. In an open (not blinded) clinical trial, the more severely affected patients may be selected for the treatment that the investigator feels (perhaps erroneously) is the better one. During duplicate assays, a knowledge of the results of the first assay may serve as a criterion for accepting or rejecting the second assay. Data used in curve fitting, e.g. as a function of time, may be rejected because the fit is not as good as expected.

## *Sampling for Attributes and Operating Characteristic Curves*

The binomial distribution can be used to construct acceptance/rejection sampling plans in quality control. MIL STD 105D is an excellent document describing sampling plans for attributes. These plans recommend the number of items to be inspected, and the number of rejects that are observed determine whether or not the lot will be accepted. The plans are more or less stringent depending on the seriousness of a defect and the risk of making a wrong decision. A wrong decision is either (1) to pass a poor lot or (2) reject a good lot.

As an example, Plan *N* from MIL-STD-105D is devised such that if there are 0.25% defects in a lot, there is a good chance of passing the lot. This corresponds to an acceptable quality limit (AQL) of 0.25. In this plan, 500 samples are taken at random, and if 3 or fewer defects are found, the lot passes; otherwise, the lot is rejected. If 0.25% of the lot is defective, the probability of the lot passing this inspection is equal to the probability of finding either 0, 1, 2, or 3 defects in the 500 samples inspected, since any one of these observations will result in a decision to pass the lot. This probability can be calculated using the binomial formula given in equation (1):

$$\sum_{X=0}^{X=3}\binom{N}{X}p^x q^{N-X} = P(0)+P(1)+P(2)+P(3)$$

where, $P(0)$, $P(l)$, $P(2)$, and $P(3)$ are the probabilities of finding 0, 1, 2, and 3 defects, respectively. The sum of these four probabilities is

$$\binom{500}{0}(0.0025)^{\circ}(0.9975)^{500} + \mathrm{L}\binom{500}{3}(0.0025)^{3}(0.9975)^{497} = 0.29+0.36+0.22+0.09+0.96$$

Therefore, the probability of passing the lot is at least 0.96 if there are 0.25% or less rejects in the batch. Conversely, the probability of rejecting such a lot is 0.04 (1–0.96). Four percent (0.04) is equivalent to the a error in hypothesis testing, if 0.25% or less defects characterize an acceptable lot.

What about the probability of accepting lots of bad quality? This is an important aspect of a sampling plan and is described by an operating characteristic (OC) curve as shown in Fig. 11.9. The OC curve shows the probability of accepting a lot for a specified plan, given the true percentage of

defects in the lot. Note that the probability of rejecting a lot of specified quality is simply (1-probability of acceptance of that lot). To construct the OC curve, it is sufficient to calculate the probabilities of acceptance at various "percent defective" values (as was done for 0.25%) and draw a smooth curve through these points. For example, if there are 1% rejects in the lot, the probability of acceptance can be calculated from equation (1).

$$\sum_{X=0}^{X=3} \binom{500}{3} (0.01)^X (0.99)^{N-X} = 0.26$$

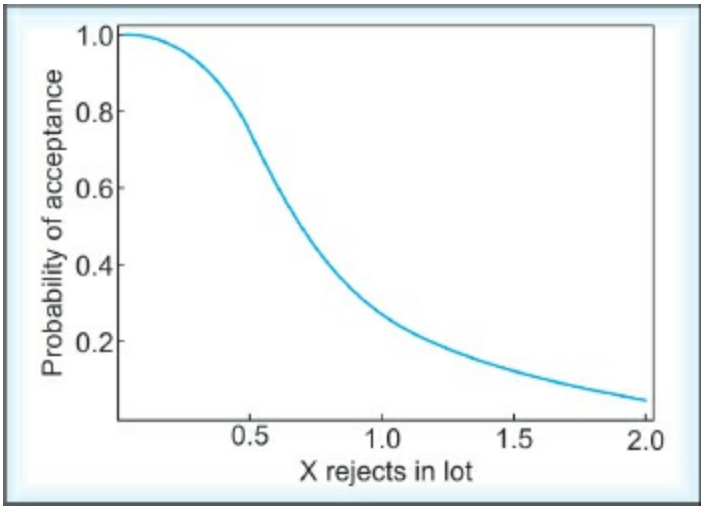



**Fig. 11.9:** Operating characteristic curve (OC) for plan N, AQL = 0.25. (MIL-STD-105D)

Note that $p$ = 0.01 (probability of observing a reject), and that $q$ = 0.99 (1 – $p$). Since 0.26 is the probability of accepting such a lot using Plan $N$, the probability of rejecting the lot is (1–0.26) = 0.74. The interpretation is that a lot with 1% chipped tablets, for example, will be rejected about 3/4 of the time and will pass the test 1/4 of the time using this plan.

## *Sample Size*

"What size sample do I need?" and "How many patients should I recruit?" are common questions that arise during planning of experiments. Estimation of the sample size needed to show a statistically significant difference (if at least some predetermined true difference exists) is an important problem in pharmaceutical and clinical studies. When testing means, the difference to be detected (d) under $H_A$, the $\alpha$ and $\beta$ errors, and the sample size, N, are closely

related. Given three of these values, the fourth is fixed. To calculate the sample size, one must specify (1) the variance, (2) $\alpha$, the risk of erroneously declaring significance and rejection of the null hypothesis, (3) $\beta$, the risk of erroneously accepting the null hypothesis given an alternative and (4) the "difference to be detected." As the experienced researcher knows, these risks and the meaningful differences are not easy to assess and are often a matter of good judgment.

If the sample size is large or if $\sigma^2$ is known, the computation of a sample size, $N$, for a single (or paired) sample experiment is:

$$N = \sigma^2[(Z_\alpha + Z_\beta)/d]^2 \quad \ldots (2)$$

where, $d$ is the difference to be detected, and $Z_\alpha$ and $Z_\beta$ are the appropriate normal deviates for the α level and $\beta$ level, respectively. For a two-sided test, $Z_\alpha$ is the value of $Z$ above which α/2% of the area is found in the normal curve (*see* Table 11.2). These are the same values used for hypothesis testing, e.g. 1.96 at the 5% level of significance, and 2.58 at the 1% level. $Z_\beta$ is the value of $Z$ above which $\beta$ of the area is found in the upper tail. For example, for $\beta = 0.2$, $Z = 0.842$; for $\beta = 0.1$, $Z = 1.28$; and for $\beta = 0.05$, $Z = 1.645$. Although this may appear complicated, the examples that follow should clarify the use of this equation.

For example, a question regarding sample size may be posed as follows. It is important to detect a mean tablet weight that is 3 mg or more different from the target weight. If such a difference exists, there should be a 90% chance that a statistical test will show significance ($\beta = 0.10$; power = $1\text{-}\beta = 0.90$). The chance of concluding that a difference of 3 mg or more exists when in fact the batch is on target should be small, e.g. 5% (α = 5%). The difference, 3 mg, and the $\alpha$ and $\beta$ risks described previously are not preordained. These values are a result of careful thought and experience, and are usually "ball-park" figures, unless legal or official requirements dictate exact limits and risks. If the standard deviation is 10, $N$ is calculated from Eq. (2):

$$N = 10^2[(1.96 + 1.28)/3]^2 = 117$$

This means that 117 tablets should be sampled to determine the mean weight. A statistical test comparing the observed mean weight versus the target weight would then have $\alpha$ and $\beta$ risks of 0.05 and 0.10, respectively. 

The values of 1.96 and 1.28 in the calculation of $N$ refer to the standard normal deviate for $\alpha = 0.05$ and $\beta = 0.10$.

A variation of Eq. (2) gives the approximate sample size for binomial data, replacing $a^2$ appropriately with $pq$.

For *small sample sizes* and *unknown a*, the problem is more difficult because the value of t changes considerably with changes in sample size (appropriate $t$ values must be used in the above formula, replacing $Z$). If an arbitrary value of $t$ is chosen based on a preliminary guess of df, and the calculated $N$ is substantially different from the preliminary guess, the estimate of the sample size will be incorrect. Guenther has shown that increasing the sample size as calculated from Eq. (2) by $0.5Z_\alpha^2$ gives a nearly correct value for the sample size. Davies provides tables for the sample size needed for $t$ tests given $\alpha$, $\beta$, and $d/\sigma$.

As an example of this calculation, consider a bioequivalency study in which the areas under the blood level versus time curves (*AUC*) for two formulations are to be compared. What sample size would be needed to detect a difference in the AUCs of ± 20 h mcg/ml with a power of 90% at the 5% level in a situation in which the average area is expected to be about 100 and the standard deviation is estimated to be 25? Bioequivalency tests are usually designed so that each subject receives each formulation on separate occasions (a paired design). According to Eq. (2), the sample size needed is:

$$N = (25)^2[(1.96 + 1.28)/20]^2 = 16.4$$

Since $\sigma^2$ is unknown, add $0.5Z_\alpha^2$ to 16.4:

$$(0.5) \cdot (1.96)^2 + 16.4 = 18.3$$

Nineteen subjects will satisfy the requirements for this study.

The above calculations are for a singlesample or paired-sample test. The calculations for a two-sample test are slightly different. The sample size for the two-sample case, where, $N_1 = N_2 = N$, ($N_1$ and $N_2$) are the sample sizes for the groups to be compared), is calculated as follows:

$$N = 2\sigma^{22}[(Z_\alpha + Z_\beta)/d]^2 \quad \text{... (3)}$$

If $\sigma^2$ is unknown, add $0.25Z_\alpha^2$ to N.

For example, time to dissolution (50%) is to be compared for two formulations of the same drug. How many tablets of each formulation should be used if a true difference of 15 min or more is to be detected with a power

of 80% at the 10% level of significance (two-sided test)? The standard deviation is approximately 10. In comparison with the previous example, the experimenter is willing to take greater risks of making errors of the first and second kinds ($\alpha$ and $\beta$ are 0.1 and 0.2, respectively). When the risks of making these errors are larger, the sample size needed to meet these criteria is smaller. Using equation (3), the sample size is:

$$N = 2(10)^2[(1.645 + 0.842)/15]^2 = 5.5$$

Adding $0.25Z_\alpha^2\,(0.25 \times 1.65^2 = 0.7)$ to $N$ results in 6.2. Seven tablets from each formulation should be sufficient to satisfy the above conditions.

For one-sided tests, the same formulas are used, but $Z_\alpha$ has $\alpha$ percent (%) of the area in the upper tail (e.g. at the 5% level, $Z_\alpha$ = 1.645, and at the 10% level, $Z_\alpha$ = 1.28). In the previous example, if a one-sided test were appropriate, using Eq. (3) results in the following calculation of $N$:

$$N = 2(10)^2[(1.282+0.842)/15]^2+0.25\,(1.282)^2 = 4.4$$

Five tablets of each formulation would be adequate.

## Power

The power of a test is its ability to detect a difference if such a difference truly exists. The power can be calculated solving for $Z_\beta$ from Eqs (2) or (3), specifying values for $N$, $\alpha$ and "d". In the one-sample (or paired-sample) test, for example, from Eq. (2):

$$Z_\beta = d\sqrt{N/\sigma^2} - Z_\alpha \qquad \ldots (4)$$

If $N$ is small, substitute appropriate values of $t$ for $Z$, or make the inverse adjustment for $N$ as discussed previously, i.e. make the computations using $Z$, but subtract $0.5Z_\alpha^2$ from $N$.

Consider a bioavailability study in which the average of the ratios of *AUC* of a tablet formulation and solution of the same drug is to be assessed. If the bioavailabilities of the tablet and solution are the same, the ratios of the areas should be equal to 1, on the average. Therefore, the null hypothesis is $H_0$: Ratio = 1. The two-sided alternative is $H_A$: Ratio ≠ 1. The FDA is interested in knowing the power of such tests because small sample sizes that result in nonsignificant differences may have little power. What is the power of the test with a sample size of 12 in which protection is to be provided against erroneous acceptance of alternatives for which the true ratio differs from 1 by 0.2 (20%) or more? The test is performed at the 5% level, and the s.d. is approximately 0.25. Since $N$ is small, subtract $0.5Z_\alpha^2$ from $N$ ($Z_\alpha$ = 1.96; 12-$0.5Z_\alpha^2$ ~ 10), and use this value in Eq. (4). (This is the inverse of the procedure used previously in calculating the sample size.) The calculation using Eq. (4) follows:

$$Z_\beta = 0.2\sqrt{10/0.25^2} - 1.96 = 0.57$$

$\beta$ = 0.285 and the power is 71.5% as determined from Table 11.2.

In a two-sample test, the following is used:

$$Z_\beta = d\sqrt{N/2\sigma^2} - Z_\alpha \qquad \ldots (5)$$

Suppose, in a two-sample, two-sided test, $N = N_2 = 10$, $\alpha = 0.05$, $d = 2$, and $\sigma^2 = 3$. Again, subtracting 0.25 $Z_\alpha^2$ from $N$ as before (10–1 = 9), and using Eq. (5):

$$Z_\beta = 2\sqrt{9/(2)(3)} - 1.96 = 0.49$$

$Z_\beta$ is 0.49 and the power is approximately 69% (Table 11.2). This means that if a statistical test (*t* test) is performed comparing the means of the two groups at the 5% level with 10 experimental units in each group, if the true difference (d) between the group means is 2 or more, the statistical test will show a significant result with a probability of at least 69%. The calculation of power is discussed more fully in Pharmaceutical Statistics by Bolton (*see* General References).

## Consumer Acceptance Testing

Before a product is introduced to the market, it may be advisable to assess patient or consumer acceptability of one or more formulation attributes, such as taste, color, packaging, and physical characteristics (e.g. viscosity of a liquid suspension or thickness of an ointment). The formulation section is often involved in implementing and evaluating these tests, and a familiarity with common designs such as monadic tests, paired comparisons, and triangle tests is important.

### *Monadic or Single-product Test*

In this test, the attributes of two or more products are compared, and each individual evaluates only a single product. To analyze the resulting data, some quantitative measurement must be associated with the test since such tests are often qualitative in nature. A number or score may be assigned to a descriptive term related to the attribute being assessed. For example, “excellent” = 1; “good” = 2; “fair” = 3; and “poor” = 4; or “I would buy this product” = 3, “I might buy this product” = 2; and “I would not buy this product” = 1. The scoring systems are often arbitrary, but research on the development of such scoring schemes considers the following: (1) How many choices should be given to the test panelist? (2) How should the evaluation statements be expressed? (3) Are the intervals between adjacent statements equal? For example, is the difference between “poor” and “fair” the same as the difference between “fair” and “good”? Usually, an arbitrary equiinterval linear scale is used despite its theoretical shortcomings.

Use of six or seven reasonably spaced evaluation statements with a sufficiently large panel (30 or more) results in data that can be reliably analyzed. Snedecor and Cochran discuss such scaled data and conclude that the ordinary *t* test is applicable (with a small continuity correction) provided a reasonable sample size is used. The test for two products can be analyzed using a two-sample *t* test. For more than two products, analysis of variance is used. The design, analysis, and interpretation of such experiments are the same as for other similar tests described in this chapter, e.g. *comparison of two drugs in independent groups*. Consider the following example.

An antacid product is reformulated with a new less expensive flavoring agent. Fifty subjects, users of this product, were randomly divided into two

groups of 25 each, with each subject evaluating either the new or old formula. A "thermometer" scale (Fig. 11.10) was used, and the results are shown in Table 11.10. Testing the hypothesis of equality of means, $H_0$: $\mu_1 = \mu_2$, use the t test for two independent groups at the 5% level. The procedure and calculations have been described in the section "Hypothesis Testing for Statistical Significance" for the case of the independent two-sample *t* test.

$$t_{47} = |\bar{X}_1 - \bar{X}_2 - 0| / S_p \sqrt{1/N_1 + 1/N_2}$$

$$S_p = \sqrt{[24(1.94)^2 + 23(1.90)^2]/47} = 1.917$$

$$t_{47} = |7.44 - 6.88| \left[1.917\sqrt{1/25 + 1/24}\right] = 1.02$$

**Table 11.10:** Comparison of flavor using thermometer scale

| | Old product | New product |
|---|---|---|
| Number of subjects | 25 | 24* |
| Average score | 7.44 | 6.88 |
| Standard deviation | 1.94 | 1.90 |

* One subject was ill andcould not evaluate the product


**Fig. 11.10:** Typical thermometer scale for evaluating consumer products

Although the old product received a higher rating, the two formulations are not significantly different. The *t* value needed for significance is 2.01 (*see* Table 11.5).

The action based on these results depends on many factors, including common sense. In this situation, the new product might be marketed because of the "nonsignificant"-difference and the decreased production costs. Alternatively, the marketing group might feel that the existing franchise is so good that any product that is conceivably not as good (as in the case here), or different in any way for that matter, would not be a viable substitute. In the latter case, an experiment designed to test if products are distinguishable might be preferred.

The independent two-sample *t* test described in this section, sometimes known as a "monadic" test, may lack sensitivity because of large intersubject variability. Nevertheless, this test is often preferred, depending on product type as well as on cost and time restraints. Such tests can be completed more quickly than paired tests in which each subject evaluates two products. Also, the procedure used in this test may be more realistic when related to how products are actually used in the marketplace.

## *Paired Tests*

A paired test in which each subject compares two or more products is often desirable because of convenience and the improved sensitivity resulting from the intrasubject comparison. Products may be evaluated by (1) preference whereby the subject notes which product is preferred, (2) a ranking procedure or (3) a scoring or rating system as discussed under "monadic" tests. If more than one product is to be compared, "round robin" tests can be used, whereby each subject tests each and every possible pair of products. A variation is the incomplete block design, in which a balanced subset of the products are tested by each individual. In repeat tests, products are evaluated sequentially, and sufficient time should be allowed between assessments for the subject to return to his normal state. For example, in taste or smell tests, the sensory organs can be over stimulated and a recovery period should be part of the experimental design. Order effects occur often in such tests because of unconscious bias as well as sensory fatigue, and principles of good design including blinding and randomization are particularly important. Testing using a response scale (e.g., 1 to 10) can be analyzed by paired *t* test

procedures or by a two-way analysis of variance.

If the number of subjects is small and the experimenter feels that assumptions such as normality that underly the statistical analysis are not valid, suitable nonparametric techniques (such as ranking tests) can be used (these tests are described at the end of this chapter). As an example consider a test in which 60 subjects were asked to use two variations of a formula, A and B. Preference was indicated based on the formula's "feel" in the mouth, and the results were the following:

| Prefer formula A | Prefer formula B | No preference |
|---|---|---|
| 32 | 16 | 12 |

The question of what to do with "no preference" decisions is controversial. Should subjects be forced to state a preference? Although there is no definitive answer, it appears best to allow subjects to give a "no preference" decision, but not to include this data in the statistical analysis. Certainly, the "no preference" data should not be ignored when the recommended action is finally made. It would certainly make a difference, for example, whether either 90% or 10% of the subjects made a "no preference" decision, regardless of statistical significance.

These data can be analyzed by *chi-square or equivalent binomial techniques*. The hypothesis to be tested is that among the subjects who express a preference, there is equal preference for the two formulas.

$$H_0: p = 0.5 \qquad H_A: p \neq 0.5$$

$$\text{Percentage preferring A} = 32/48 = 66.7\%$$

$$Z = [\,|p_A - 0.5| - 1/(2N)]/\sqrt{pq/N}$$

$$= (0.667 - 0.5 - 1/96/\sqrt{(0.5)(0.5)/48} = 2.17$$

There is a significant preference ($P < 0.05$) for Formula A among those who express a preference. As just explained, when the results of such an experiment are reported, the number of consumers who had no preference should be acknowledged. In this example, the conclusion is that of 48 subjects who expressed a preference, 2/3 preferred Formula A ($P < 0.05$), and there were 12 subjects who had no preference.

## *Triangle Tests*

Preference tests are primarily designed to predict the proportion of the consumer population that will prefer one preparation to another. However, consumers may perceive two formulas as being very different, but they could be segmented into two equal groups, each of which prefers the alternate product. Thus, equality of preferences does not distinguish between two possible situations (1) the existence of two distinct but equal groups, each of which prefers one or the other product and (2) one homogeneous group that simply cannot differentiate the products.

The triangle test is designed to assess if products are in fact distinguishable. Three samples, similar in appearance, are submitted for testing, two of one product and one of another. The consumer is asked to choose the product that is different from the other two (optionally, a preference can also be requested). If there is a "significant" number of correct guesses, the products are considered distinguishable, and preference data, if requested, may be analyzed to determine the preferred product.

Order effects are important in such a test, and the order of presentation should be randomized or balanced. There are six possible ways of presenting three products of which two are identical. If the products are called A and B, they can be presented in the following orders. (The products should be labeled in some random manner, and the six orders should be randomized.)

| Subject | 1st | 2nd | 3rd |
|---|---|---|---|
| 1 | *A* | B | B |
| 2 | A | A | B |
| 3 | B | B | A |
| 4 | B | A | A |
| 5 | B | A | B |
| 6 | A | B | A |

If order effects are present (e.g. the choice by the panelist of the odd, or different, product tends to be the second product tested), the balanced design protects against bias.

Carry over effects cause a problem. For example, in a taste test, if one or more products has so strong a taste as to numb the taste buds, this would

influence the evaluation of the subsequent products. Sufficient recovery time in between assessments would be an important part of the design. On the other hand, too long a time interval between tastes could involve a "memory" factor, thus introducing excess variation. One could, however, reason that this situation is more realistic in that consumers do not normally compare products side by side. These are all valid arguments to be individually assessed for each product and marketing situation.

Suppose that a variation of the base of a marketed ointment is made to improve its kinesthetic properties and that 30 panelists are to compare the "feel" of the old and new products using a triangle test design. After proper randomization of the products, 15 of the 30 panelists make the correct choice, i.e. choose the correct "different" product.

The statistical analysis is based on a test of the following hypothesis:

$$H_0: p = 1/3$$
$$H_A: p > 1/3$$

This means that if there is no difference, the correct product will be chosen one-third of the time by chance. This is a good example of a one-sided test, i.e. a difference can be tested for significance only if correct choices are made by more than one third of the panelists. If less than one third choose the odd product, this result can be due only to chance. If the experiment results in significantly less than one-third correct choices, the experimental procedure should be questioned. The test of hypothesis uses the binomial distribution with $p = 1/3$ and N = 30. The test statistic is:

$$Z = [\,|\,p_{obs} - 1/3\,| - 1/(2N)]/\sqrt{(1/3)(2/3)/N}$$

where, $p_{obs}$ is the observed proportion of correct choices, and $N$ is the number of responses. For this example:

$$p_{obs} = 15/30 = 0.50$$

and

$$Z = (|1/2 - 1/3| - 1/60)/\sqrt{(2/9)/30} = 1.74$$

At the 5% level (one-sided test), a value of $Z$ equal to or greater than 1.65 is significant (*see* Table 11.2). In this example, the conclusion is that the products are distinguishable at the 5% level.

## *Other Tests*

Other designs used in consumer testing include *round-robin*, *sequential*, and *repeat paired preference tests*. The ***round-robin*** used for more than two products compares all preparations in all possible pairs. The ***sequential design*** is based on the idea that results can be analyzed sequentially after each preference, so that if large enough differences exist, a smaller sized panel can be used to come to a decision more quickly than if a fixed sample size is chosen in advance. The advantage of this design depends on many factors, including anticipated differences, the nature of the products, and the setting of the test. ***Repeat paired comparison tests*** consist of a paired preference test repeated on a second occasion. The analysis can result in a segmenting of the population into two groups preferring each product as well as a group who cannot distinguish the products. This test may be physically difficult to implement, however, and the data are sensitive to deviations from the model.

## Analysis of Variance (ANOVA) and Experimental Design

ANOVA is inextricably connected to experimental design. Experiments that are conceived to compare, estimate, and test such effects as drug treatments, formulation differences, and analytical methods can be designed to yield an optimal return for effort expended. The experimental results may then be analyzed by ANOVA techniques. A good experiment speaks for itself; the conclusions are often obvious without complicated mathematical treatment. With sophisticated calculators and computers readily available, however, most experiments, if designed properly, can be easily analyzed. In a poorly designed experiment-on the other hand, more than one factor may contribute to an experimental result with no way of untangling the effects of the factors (this is called “confounding”). Examples of obvious confounding are (1) a clinical study in which patients are allowed to take various concomitant drugs other than the test drug that affect the condition being treated and (2) a comparison of a new tablet formulation to the former formulation for dissolution, with the tablets prepared on two different tablet presses, one formulation on each press. Differences in the performance of the presses (pressure, for example) can contribute to differences in dissolution in addition to differences due to formulation changes.

Analysis of variance separates the total variation in the data into parts, each of which represents variation caused by factors imposed on the experiment. A properly designed experiment allows a clear unconfounded estimate of such variation or, at least, can identify the confounding factors, if present. Consider an experiment to assess the effects of lubricating agent and disintegrating agent on the dissolution of a tablet. The final analysis of variance would separate the effects of these factors by computing that part of the total variation attributable to the lubricating and disintegration agents isolated from that variation due to experimental error. This separation serves as a basis for testing statistical hypotheses.

### *One-way Analysis of Variance*

One-way ANOVA can be considered an extension of the independent two-sample *t* test to more than two groups. In the *t* test, the following statistic compares the difference of two sample means to the standard deviation of the difference (the denominator of the *t* statistic):

$$(\overline{X}_1 - \overline{X}_2)/\sqrt{S_p^2(1/N_1 + 1/N_2)}$$

The pooled variance, $S_p^2$, depends on the variability within each of the two groups. If the two means come from the same normal distribution, the difference between these means (suitably weighted) should also be a measure of the variability of the data. Since the variance of a mean, $S_{\bar{x}}^2$, is equal to $S^2/N$, then $S^2$ is equal to $NS_{\bar{x}}^2$. A comparison of the estimate of $S^2$ from the difference of the two (or more) means to the within group variability ($S^2$ pooled) is a measure of the difference between means. If these two estimates of variation are similar, one may conclude that the means of the groups do not differ, i.e. that the difference between the means can be accounted for by ordinary variability. If the variability due to the difference between means is "significantly" larger than that within groups, one can conclude that the means of the groups differ.

A large variability of the means is associated with large differences between the means. Since as noted previously, the variance of a mean is $\sigma^2/N$ (where $N$ is the sample size), the sample variance of the mean is weighted by $N$ to obtain an estimate of $\sigma^2$. Thus, a statistic that can be used to test treatment (group) differences is formed by the following ratio:

$$\sum \frac{N_i(\overline{X}_i - \overline{X}_0)^2/(G-1)}{S^2} = F$$

where, the numerator is the variability (variance estimate) due to the different means, and the denominator is the pooled within group variance. $\overline{X_i}$ is the mean of the *ith* group, $\overline{X_0}$ is the overall mean of all the groups, and $G$ is the number of groups.

The distribution of this $F$ ratio under the hypothesis that the group means are equal is the same $F$ probability distribution previously mentioned when testing for the equality of two variances. The $F$ distribution has $\upsilon_1$ and $\upsilon_2$ degrees of freedom, where $\upsilon_1$ is the degrees of freedom associated with the numerator ($G$–1), and $\upsilon_2$ is the *degrees of freedom* associated with the denominator equal to $\Sigma(N_i)$-$G$. For two groups, analysis of variance results in exactly the same probability level as the two-sample $t$ test. (The $F$ statistic with 1 df in the numerator, which is the case for two groups, is the square of the $t$ statistic with degrees of freedom equal to that in the denominator of the

*F* ratio.) The following example should clarify some of the foregoing concepts and computations.

In a preclinical study, animals were treated with two antihypertensive experimental drugs and a control drug, with 12 animals randomly assigned to the three groups, four per group. One animal died from a nondrug related cause and was lost to the experiment. The results (change in blood pressure from baseline) are shown in Table 11.11. Are the treatment means different, or do the observed differences merely reflect the inherent variation of the animals' response to such treatments? Under the assumptions that the *variances within each group are equal*, and that the data are *independent and normally distributed*, a test for equality of means can be performed using analysis of variance. The reader should refer to Table 11.11 as an aid in following the calculations needed for the ANOVA.

**Table 11.11:** Change of blood pressure in preclinical study comparing two drugs and control

| | Drug 1 | Drug 2 | Control |
|---|---|---|---|
| | 15 | 8 | |
| | 12 | 14 | 16 |
| | 19 | 13 | 20 |
| | 11 | 6 | 22 |
| Sum | 57 | 41 | 58 |
| Mean ($\overline{X}$) | 14.25 | 10.25 | 19.33 |
| *S* | 3.59 | 3.86 | 3.06 |

The overall mean is:

$$\overline{X_0} = 156/11 = 14.18$$

The Between Treatment Mean Square (BMS) is:

$$\begin{aligned} & SN_t(\overline{X}_t - \overline{X}_0)^2/(G-1) \\ &= [4(14.25-14.18)^2 + 4(10.25-14.18)^2]/2 \\ &= 70.74 \end{aligned}$$

The Within Treatment Mean Square (WMS) is the variance pooled from within each group.

$$\begin{aligned} & [3(S_1^2)+3(S_2^2)+2(S_3^2)]/(N_t-3) \\ &= 102.167/8 \\ &= 12.77 \end{aligned}$$

where, $N_t = \Sigma N_i = 11$.

The $F$ ratio with 2 and 8 df, a test of differences among treatment means, is BMS/WMS.

$$F_{2,8} = 70.74/12.77 = 5.54$$

If the hypothesis that all three means are the same is true, the ratio BMS/WMS shoul *d* be equal to 1, on the average. If the computed $F$ ratio is less than 1, the means are not significantly different. If the $F$ ratio is greater than 1, an $F$ table should be used to determine if the ratio is sufficiently large to declare significance. The cutoff point for significance at the 5% level for $F_{2,8}$ is 4.46 (*see* Table 11.9). Since the calculated $F$ (5.54) is larger than 4.46, the conclusion is that at least two of the means differ from each other at the 5% level.

**Analysis of Variance (ANOVA) Table**

| Source | *Df* | Sum of Squares (SS) | Mean Square (MS) |
|---|---|---|---|
| Between groups | 2(*G*–1) | 141.47 | 70.74<br>$F^{2,8} = 5.54$ |
| Within groups | 8(*N*–3) | 102.17 | 12.77 |
| Total | 10(*N*–1) | 243.64 | |

Computations for ANOVA, when done by hand, routinely use shortcut formulas, and the results are presented in an analysis of variance table. For the above example, the calculations can be simplified as shown below (*refer* to Table 11.11 to help in following the calculations).

$$\text{Total sum } (\Sigma X) = 156$$

$$(\Sigma X)^2/N_t = \text{correction term } (C.T.) = 156^2/11 = 2212.36$$

$$\Sigma X^2 = 2456$$

*Total Sum of Squares* = TSS

$$= \Sigma X^2 - \text{C.T.} = 2456 - 2212.36 = 243.64$$

*Between Treatment sum of Squares* = BSS

$$ST_i^2/N_i - \text{C.T.}$$

$$= \frac{57^2}{4} + \frac{41^2}{4} + \frac{58^2}{3} - 2212.36$$

$$= 141.47$$

where, $T_i$ is the total sum of data in *ith* treatment group.

Within Treatment Sum of Squares
= WSS = TSS BSS
= 102.17

The analysis is typically presented in an "analysis of variance (ANOVA) table."

The total variation is $S(X - \bar{X})^2$ and is separated into two parts, that due to differences of the means of the groups, and that due to the pooled variation within the groups. If the groups are not different, the variation among different groups should be no greater than that due to variation among individuals within groups, the "within" variation.

In general, there are "*G*" groups with replicate measurements in at least one of the groups. Although calculations are simplified and efficiency is usually optimal when comparing means if the number of observations in each group is equal, this is not a necessary condition for this analysis. In many experiments, especially in those involving humans, the original plan usually provides for equal numbers in each group, but life circumstances intervene, resulting in dropouts and lack of a symmetric design. The imbalance presents no problem, however, in the analysis of this one-way design.

Consider another example of this design in which an analytical method is tested by sending the same (blinded) sample to each of seven laboratories from the same company, located at different sites. Each of laboratories 1 through 6 has three analysts perform the analysis. Laboratory 7 reports only

two results because only two analysts are available. In this experiment, a problem would result if one of the two analysts in laboratory 7 does a third analysis to present three results, since then the data for that laboratory would not be independent. Independence, in this example, refers to the fact that within each laboratory, analysts perform their analyses independently. The results are shown in Table 11.12. The computation is identical to that described in the previous example.

**Table 11.12:** Assay results for seven laboratories Laboratory

| | Laboratory | | | | | | |
|---|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **6** | **7** |
| | 9 | 11 | 6 | 10 | 5 | 7 | 12 |
| | 8 | 9 | 9 | 10 | 3 | 7 | 10 |
| | 7 | 13 | 9 | 7 | 4 | 7 | – |
| Average | 8 | 11 | 7 | 9 | 4 | 7 | 11 |
| | Overall average = 8.0 | | | | | | |

Table 11.9 shows $F_{6,13}$ = 2.9 at the 5% level. Therefore, the ratio 8.5 is significant, and at least two of the laboratories are considered to be different ($P < 0.05$).

**ANOVA**

| Source | DF | SS | MS | |
|---|---|---|---|---|
| Between labs | 6 | 102 | 17 | |
| Within labs | 13 | 26 | 2 | $F_{6,13} = 8.5$ |
| Total | 19 | 128 | | |

Confidence limits on the overall average can be constructed to give a range for the true mean of the analysis: $\bar{X}_0 \pm t(S_{\bar{X}0})$. For 95% confidence limits, the value of $t$ is 2.16 and $S_x^2$ = 2/20. (From the ANOVA Table, df = 13 and $S^2$ = 2; the total number of observations is 20.) The 95% confidence limits are $8 \pm 2.16\sqrt{2/20} = 8 \pm 0.7$. Note that the variance estimate for the computation is the *within error*, or *variance*. This is the correct error term because all of the laboratories of interest were included in the experiment (*a fixed model*). If the laboratories were only a sample of many possible laboratories, then the correct error term would be the between mean square with (C–1) degrees of freedom. (This is known as a *random model*.) In 

general, the between mean square is larger and has less degrees of freedom than the within mean square, and confidence limits are wider in this case. This less precise estimate is to be expected, because in the random case, not all members of the population have been sampled. The overall mean is estimated from a small sample of the possible laboratories. In the fixed model, each of all of the possible laboratories (7) havehas been sampled with replicate determinations (analysts) obtained from each laboratory. If the number of observations in each group is not equal in the random model, construction of a confidence interval for the overall mean is difficult.

## *Multiple Comparisons*

With more than two treatments, if the *F* test is significant in the analysis of variance, one must determine which of the treatments differ. If a separate test is done comparing each pair of treatments, the chances of finding significance when the treatments are really identical are greater than that indicated by the $\alpha$ level. If the $\alpha$ level is 0.05, by definition, one time in twenty a difference will be found to be significant when the treatments are truly identical. If more than one pair of treatments are tested for significance in the same experiment, significant differences are found in more than 5% of such experiments, if treatments are truly identical. This concept may be better understood if one thinks of a large experiment in which 20 independent comparisons are to be made at the 5% level. On the average, one significant difference would be expected in each experiment of this kind, if the hypothesis of equal treatment means is true. The problem of multiple comparisons is complex, and many solutions have been proposed.

The simplest method of dealing with multiple comparisons gives more significant results than would be expected from the ANOVA $\alpha$ level. At test is constructed for differences between means using the degrees of freedom and mean square error from the ANOVA. If the sample sizes are the same in each group, a single least significant difference (*LSD*) can be constructed:

$$LSD = t_{df,\alpha}\sqrt{S^2(2/N)}$$

where, *t* is the tabled value of *t* with appropriate df at the *a* level of significance. Any difference exceeding the *LSD* can be considered to be significant. The *LSD* test should be used *only if the F test from the ANOVA is significant.*

In the above example of seven laboratories, comparisons of laboratories with three observations (Laboratories 1 through 6) would result in an *LSD* equal to $2.16\sqrt{2(2/3)} = 2.49$, where, 2.16 is the value of t with 13 df at the 0.05 level, and 2 is the within mean square. Laboratories 1 and 2 are significantly different from each other, the difference of their means (11–8 = 3) exceeding the *LSD* (2.49). Laboratories 1 and 3, for example, are not significantly different. In comparing any of laboratories 1 through 6 to laboratory 7, an ordinary two-sample *t* test is used. For example, the comparison of laboratories 1 and 7 is performed as follows:

$$t = (11-8)/\sqrt{2(1/3+12)} = 2.32$$

Since the calculated *t* (2.32) is greater than the tabulated *t* at the 5% level with 13 df (2.16), the difference is significant.

Other tests take into account the multiplicity of comparisons and impose a penalty so that differences greater than that calculated for the *t* test are required for significance. One commonly used method is ***Tukey's multiple range method:***

$$\text{Compute } |\text{Difference}| / \sqrt{S^2/N} = q \quad \ldots (6)$$

Text by Snedecor and Cochran shows significant values of *q* depending on the number of treatments, df and $\alpha$ level.

From Eq. (6) the minimum difference for significance for any pair of treatments is is $q\sqrt{S^2/N}$. Strictly speaking, when using this formula, *N* should be the same for each treatment. In the laboratory example, laboratory 7 has two observations, compared to three observations in the other 6 laboratories. A slight adjustment for *N* would be necessary in this example, but for purposes of illustration, assume that there are three observations for all laboratories ($N = 3$).

For this example at the 5% level, *q* equals 4.88, and the minimum difference needed for significance is:

$$4.88\sqrt{2/3} = 3.98$$

This difference is larger than that computed by the LSD procedure. Laboratories 1 and 2 are not significantly different using this method. Laboratory 5 is significantly different from all the other laboratories except for laboratories 3 and 6. Significant difference in laboratory results may require an action to determine and correct the cause; or, perhaps, it might be

important just to know that differences exist.

Another commonly used method is Duncan's multiple range test, which is considered to have excellent properties.

## *Two-way Analysis of Variance (Randomized Blocks)*

The two-way model is an extension of the paired t test in which more than two groups or treatments are compared. As in the paired *t* test, each individual (often referred to as a "block") is subjected to every treatment. Sometimes this model is described as a design in which each individual acts as his own "control."

In general, the order in which treatments are assigned to individuals is randomized unless a special design such as a crossover is used, which is discussed in a subsequent section of this chapter. A table of random numbers can be used to randomize the order of treatments to be tested on each individual or experimental unit.

The analysis of the two-way design is similar to the one-way ANOVA except that an additional source of variation is present, that due to the differences among the blocks. For example, if three assay methods are to be compared on six batches of granular material, the variation due to blocks (the batches) is associated with the differences in concentration of active material in the six batches. The following example should clarify the design and data analysis. The time to 10% decomposition at accelerated conditions was compared for five batches of tablets using three different kinds of packaging material, with the results shown in Table 11.13. The computations for the ANOVA are shown below. Note that the computation of the between batch and between packages sum of squares is performed in the same way as the one-way analysis of variance. The difference between the *total sum of squares* and the *sum of the between packages and between batches sum of squares* is the *error sum of squares*.

**Table 11.13:** Stability of five batches of tablets using three kinds of packaging material

| Batches | Packaging Material A | B | C | Mean |
|---|---|---|---|---|
| 1 | 96 | 101 | 89 | 95.33 |
| 2 | 89 | 99 | 80 | 89.33 |
| 3 | 82 | 88 | 83 | 84.33 |
| 4 | 94 | 94 | 90 | 92.67 |
| 5 | 93 | 90 | 89 | 90.67 |
| Mean | 90.8 | 94.4 | 86.2 | |

In this design, the error sum of squares is less than that which would be computed from the same data without the "blocking" factor of batches. If the batches sum of squares were not included in the analysis, the error sum of squares would be increased by that amount. The inclusion of the batch, or blocking factor, usually results in a smaller mean square for error, and thus a more precise experiment.

Total SS

$$= \Sigma X^2 - C.T.$$

$$= 96^2 + 101^2 + \text{L} + 89^2 - \frac{(96 + 101 + \text{K} + 89)^2}{15}$$

$$= 123{,}259 - 122{,}763.27$$

$$= 495.73$$

Between package $SS = R\Sigma(\bar{C}_i)^2 - C.T.$
$= 5(90.8^2 + 94.4^2 + 86.2^2) - 122{,}76.27 = 168.93$
where, $R$ is the number of rows (batches), and $\bar{C}_i$ is the column (packages) means.

Between Batch $SS = C\,\Sigma(\bar{R}_i)^2 - C.T. = 3(95.33)^2 + 89.33^2 + \ldots + 90.67^2) - 122{,}763.27 = 202.40$
where, $C$ is the number of columns (packages), and $\bar{R}_i$ is the row (batch) means.

Error $SS$ = Total $SS$ - Package $SS$ - Batch $SS$
$= 495.73 - 168.93 - 202.40$
$= 124.40$

As in the case of the one-way ANOVA, *F* ratios are referred to in *F* Table (*see* Table 11.9) with appropriate degrees of freedom for tests of significance. The *F* ratio for packages, with 2 and 8 df, is 5.43 (between packages MS/error MS), which is significant at the 5% level. Therefore, at least two of the packages differ. The least significant difference (*LSD*) is $t\sqrt{S^2(2/N)} = 2.31\sqrt{15.55\,(2/5)} = 5.76$. In this case, packages B and C are

significantly different, with package B resulting in the least degradation.

## *The Question of a Fixed or Mixed Model*

The F ratio for batches, with 4 and 8 df, is 3.25, which is significant at the 10% level but not at the 5% level (*see* Table 11.6). A relevant question is, "Is the fact that these batches may differ from each other with regard to their stability an important consideration, or are the five batches being used merely as a means of obtaining replication for the assessment of the three packages?" If the batches were somehow special, perhaps prepared so that they differed in some known way, it would be of interest to know if the stability of these batches were different. In these cases, "batches" is considered a *fixed* effect, i.e. there is a concern in the results of *only* these five special batches (or the method by which each was prepared), and all the batches of interest have been tested. Inferences about future batches made under unknown conditions are not of current interest.

If the batches are chosen randomly merely to provide replication, however, then the batch differences *per se* are not of primary concern. In this case, the ANOVA is referred to as a *mixed model:* the *batches are random* and the *packages are fixed*, i.e. in this experiment, possible differences among the three packages are of interest, and inferences about other yet untested packages are of no concern. If batches are random, the *F* test for batches described previously may be incorrect in the presence of batch *X* package interaction. Interaction, in this example, would be evident if package differences depended on which batch was being tested. A more correct error term (the denominator of the *F* ratio) for batch differences would come from replicate determinations within each batch, e.g. a repeat determination obtained by assaying a duplicate separate package. In any event, if batches are *random*, serving only as replicates for assessing package differences, there is usually little interest in whether or not batches differ. In fact, batch differences are known to exist, and that is why each package is tested on the different batches.

The two-way design is used in preclinical and clinical studies in which two or more drugs and/or placebos are to be compared. In these cases, animals or humans represent the rows or blocks that are considered to be random, i.e. the subjects are chosen as a means of replication to estimate the error in the experiment. The question of mixed and fixed effects models is an

important consideration in the analysis of multicenter clinical trials. Suppose a clinical study comparing the effect of an antihypertensive drug to placebo included 20 patients (10 on drug and 10 on placebo) at each of eight clinical sites, with results shown in Table 11.14. The difference between treatments is not quite significant at the 5% level as indicated by the (Treatment)/(Sites × Treatment) ratio of 4.68. An *F* ratio of 5.59 with 1, 7 df is needed for significance.

In this example, there is replication in each cell of the two-way design. (A cell is defined as the intersection of a row and column, e.g. site 3 and placebo in the 8 × 2 table, Table 11.14.) Each cell consists of ten patients. There are now two error terms in the analysis of variance, the interaction term and the within error term, the variability estimate from replicate determinations within each cell. The variation due to replicates (within mean square) can be estimated by pooling the variance between patients within each treatment group at each center ($N$ = 10). This estimate has 144 df, 9 df within each treatment and 18 df from each of the eight centers.

| *ANOVA* | | | | |
|---|---|---|---|---|
| **Source** | ***DF*** | ***SS*** | ***MS*** | |
| Between batches (rows) | $R-1=4$ | 202.40 | 50.6 | $F_{4,8}=3.25$ |
| Between packages (columns) | $C-1=2$ | 168.93 | 84.5 | $F_{2,8}=5.43$ |
| Error (row × column) | $(R-1)(C-1)=8$ | 124.40 | 15.55 | |
| **Total** | **14** | **495.73** | | |
| Sites | 7 | 137.05 | 19.58 | |
| Treatments | 1 | 37.82 | 37.82 | $F_{1,7}=4.68$ |
| Sites × Treatments (interaction) | 7 | 56.52 | 8.07 | $F_{7,144}=1.47$ |
| Within sites* | 144 | 780.56 | 5.49* | |

* Between subject variance (*see* text), within sites and treatments, is adjusted to be comparable to other terms in the ANOVA

**Table 11.14:** Multicenter trial of an antihypertensive drug and analysis of variance

| **Average blood pressure change (mm Hg)** | | |
|---|---|---|
| **Site** | **Placebo** | **Active** |
| 1 | -2.6 | -9.3 |
| 2 | 0.8 | -6.4 |
| 3 | -5.9 | -5.4 |

| | | |
|---|---|---|
| 4 | -3.2 | -8.4 |
| 5 | 2.5 | -1.2 |
| 6 | -8.2 | -4.2 |
| 7 | -0.9 | -6.8 |
| 8 | -10.0 | -10.4 |
| | -3.44 | -6.51 |

If interaction is present, the “interaction” variance is a composite of the “within” variance and the interaction of sites and drug treatments, and thus is greater than the “within” error. The interaction represents the degree to which the sites differ in their ability to differentiate the drugs. Interaction is great if some sites favor drug and others favor placebo. *If a sites × treatment interaction exists*, “within” error is the correct error term for treatments if sites are considered to be *fixed;* that is, the correct *F* ratio for treatments in the fixed model is 37.82/5.49 = 6.89. This *F* has 1 and 144 df and is significant at the 1% level. Which, then, is the correct *F* test? Is the drug significantly better than the placebo or not? With the assumption that the interaction variance is not zero, (that is, interaction is present), the answer depends on whether clinical sites are *fixed* or *random*. If fixed, the result is highly significant; if random, the result misses significance at the 5% level, an important significance level for governmental acceptance of data for a drug submission.

Practically speaking, the sites are not randomly selected; on the other hand, they may not be considered fixed. A “*fixed*” effect means that all members of the population have been sampled. In the present case, is not the purpose of the experiment to make inferences to other clinical sites? Although rhetorical, this dilemma is real, and two points should be carefully considered and understood (1) An assumption of a *fixed model* makes it *easier to obtain a significant difference between drug and placebo* and (2) if the *mixed model* shows no significance and the *fixed model* shows *significance*, the possibility exists that a significant interaction is present. A significant interaction means that clinical sites do not differentiate treatments equally, an important consideration if some clinical sites favor one treatment, and others favor the other treatment (as opposed to interaction in which one treatment is consistently favored, and only the *degree of difference is*

*different*). The implication of interaction in this context is that the effect of the treatment is not consistent but depends on the clinic, patient population, and so forth.

## *Missing Values*

Although data may be inadvertently lost from any experiment, human clinical trials are particularly susceptible to data loss. Even the most well-designed, well-intentioned study cannot enforce exact adherence to the plan, owing to the usual vagaries of life. Patient dropouts due to noncompliance, adverse effects, and missed visits due to illness or forgetfulness are part of almost every clinical study. In cross-classified designs, e.g. twoway, factorial—the usual analysis is not valid when pieces of data are missing. Various options for the analysis exist depending on the design, the quantity of missing data, and the nature of the data. Alternatives include estimation of the missing data, curve fitting, and the use of special complex computer programs, which make adjustments for the unbalanced design. The analyses are usually not simple and often are difficult to interpret. Herein lies much of the art of statistics, and an experienced statistician should be consulted in such matters.

## *Bioavailability and Crossover Designs*

Evaluation and analysis of bioavailability and pharmacokinetic parameters are important considerations for the pharmaceutical scientist. Usually, a finished dosage form is compared to some preliminary formulation or marketed formulation regarding relative absorption. The usual method of comparison consists of a clinical study in normal volunteers in which single doses of the experimental and comparative formulations are taken in a crossover design. In this design, half of the subjects are randomly chosen to take either one or the other of two formulations on the first experimental occasion (also known as period, leg, or visit), and the remaining formulation on the second occasion. A sufficient period of time should intervene between the two periods so that “all” of the drug is eliminated before the second dose is administered. It is important that power considerations be taken into account when determining sample size; the FDA recommends that the experiment be of sufficient sensitivity to have 80% power to detect a difference of 20% or more between formulations at the 5% level. Usually, 12 to 24 subjects satisfy these conditions using a crossover design.

Various parameters can be considered in the comparison of formulas for bioavailability, including (1) total amount of drug absorbed (from area under the blood level versus time curve, for example), (2) peak blood level and (3) time to peak. Analysis of the results at each plasma (or urine) sampling time should be discouraged because the multiplicity of tests and correlation of data at proximal time points lead to confusion regarding the true significance level of such tests. Also, interpretation is difficult. Other analyses for bioavailability data have been proposed that may be more relevant depending on the drug under consideration. Although these alternative analyses have much merit, bioavailability data are usually reduced at present to the comparison of the three parameters just mentioned; this procedure is acceptable to the FDA in most cases.

The simplest analysis of bioavailability comparisons uses an analysis of variance for a crossover design. Statistical tests of hypotheses may be done separately for each parameter, e.g. relative absorption, time to peak, and peak plasma level. This method involves the easiest computation and interpretation but does not take into account multiple statistical tests and correlation of the parameters. Thus, although a multivariate test may be more appropriate, it may often be difficult to interpret. Some statisticians have suggested that hypothesis testing is not appropriate in, bioavailability tests, because acceptance of the null hypothesis cannot logically lead to the conclusion that two formulations are exactly the same, since two different formulations cannot be *identically* equivalent. Rather, confidence intervals provide more useful information. For example, if two formulas can be said to be within 25% of each other with a high degree of confidence, the clinician can then decide whether the two formulations are similar from a practical point of view.

*Example:* The following example concerns the analysis of only a single parameter for purposes of illustration. The bioavailability of a tablet formulation of a new drug is to be compared to an equivalent amount of drug given in solution. The values in Table 11.15 for area under the blood level versus time curve were obtained from 12 subjects in a crossover design. The analysis of variance accounts for *order of administration* with 1 df as well as “subject” and ‘formulations” variance. The error sum of squares is the total sum of squares minus the sum of squares due to order, subjects, and formulations. Otherwise, the computations are similar to those previously described:

**Table 11.15:** Results of bioavailability study—area under blood level vs. time curve

| Subject | Tablet | Solution | Order of administration |
|---|---|---|---|
| 1 | 76 | 83 | T-S* |
| 2 | 59 | 54 | T-S |
| 3 | 84 | 95 | S-T |
| 4 | 96 | 81 | T-S |
| 5 | 50 | 64 | S-T |
| 6 | 61 | 66 | S-T |
| 7 | 48 | 57 | S-T |
| 8 | 68 | 61 | T-S |
| 9 | 74 | 70 | S-T |
| 10 | 86 | 79 | T-S |
| 11 | 91 | 88 | S-T |
| 12 | 57 | 68 | T-S |
| Average | 70.83 | 72.17 | |

*T-S = Tablet first, solution second

Between subjects

$$SS = \Sigma R_i^2/2 - C.T.$$

$$= \frac{(159^2 + 113^2 + \text{K } 125^2)}{2} - \frac{1716^2}{24}$$

$$= 4225$$

Between formulation

$$SS = \frac{\Sigma C_i^2}{12} - C.T.$$

$$= \frac{(850^2 + 866^2)}{12} - 122{,}694$$

$$= 10.67$$

Order $SS = \frac{\Sigma O_i^2}{12} - C.T.$

$$= \frac{882^2 + 834^2}{12} - 112{,}694$$

$$= 96$$

where, $O_1$ is the sum of results of first visit and $O_2$ is the sum of results for second visit.

$$\text{Total } SS = \Sigma X_i^2 - C.T.$$
$$= 127,402 - 122,694$$
$$= 4708$$

Error $SS$ = Total $SS$ – Subject $SS$ – Formulation $SS$ – Order $SS$

$= 4708 - 4225 - 10.67 - 96$

$= 376.33$

**ANOVA**

| Source | *DF* | *SS* | *MS* | |
|---|---|---|---|---|
| Subjects | 11 | 4225 | 384.1 | |
| Formulations | 1 | 10.67 | 10.67 | $F_{1,10} = 0.3$ |
| Order | 1 | 96 | 96 | $F_{1,10} = 2.6$ |
| Error | 10 | 376.33 | 37.63 | |
| Total | 23 | 4708 | | |

The *F* tests for formulations and order of administration are not significant ($F_{110}$ is 4.96 at the 5% level), suggesting that the values of total absorption from both formulations are similar. A significant order effect would occur if the average results of the first visit differed from the average results of the second visit. Such a difference could be caused by systematic differences in the experimental procedure, the assay procedure, or the state of the subjects. A significant order effect does not invalidate the experiment. The design separates the variability due to order of administration from the residual experimental error.

The *crossover design* is a special case of the *Latin square design*, and the same deficiencies (and advantages) are present in both. A discussion of Latin square designs can be found in standard texts on experimental design (*see* "General References"). If differential carryover effects exist (i.e. the effect of one drug preceding the other has a carryover effect different from that which occurs if the order of administration is reversed), the treatment effects are confounded and inferences can be misleading. In this case, confounding is interpreted as a confusion of treatment differences with differences due to the different carryover effects. Also, if interactions exist, the error is inflated, and interpretation of the results is unclear. Therefore, if either carryover effects or interactions are suspected, crossover designs should not be used. Special Latin square designs can be used to estimate and account for carryover

effects, or a simple *parallel group (one-way) design* can be used.

The advantage of the crossover design as compared with a parallel group design is that the sensitivity of the experiment is increased, owing to the use of within subject variance as the error in the crossover design, which is smaller than the between patient variance measured in the parallel groups design. Grizzle has proposed an analysis that detects carryover effects. If such effects are present, the data from only the first visit are used and then analyzed as a one-way ANOVA, disregarding the data obtained from the second visit. (In the case of two treatments, this analysis is the same as an independent two-sample *t* test.)

*Incomplete block designs* may be used for bioavailability studies when more than two formulations are to be compared. For example, if three formulations are included in a study, a full crossover would require three visits. An incomplete block design would have each subject take only two of the possible three formulations in a symmetric pattern. For formulations A, B, and C, six subjects taking A-B, A-C, B-C, B-A, C-A and C-B, in the order specified, results in an incomplete block design, balanced for order. Elementary discussions of incomplete block designs can be found in the works of Davies and Cox.

Bioavailability data are often analyzed by using a log transformation or by computing a ratio of parameter estimates (e.g. $AUC_{\text{tablet}}/AUC_{\text{solution}}$) as the test statistic. These techniques usually result in similar conclusions although the use of the log transformation results in an asymmetric confidence interval when the antilog is calculated to back transform the data. An advantage of using ratios is that the statistic may be more easily interpreted by the clinical scientist. As previously noted, use of confidence intervals may be a more appropriate way of expressing the difference between two formulations with regard to a given parameter. For the example shown in Table 11.15, 95% confidence limits on the difference between the formulations for “area under the curve” can be constructed as follows (*t* with 10 df = 2.23):

$$(72.17 - 70.83) \pm 2.23\sqrt{37.63(1/12 + 1/12)}$$
$$= 1.34 \pm 5.58 = (-4.24 \text{ to } 6.92)$$

The true difference between the *AUC*s for the two formulations lies between −4.24 and 6.92 with a probability of 95%.

## Transformations

Data analysis can often be improved by means of transformations, which result in a better fit of the data to the statistical model, or an improvement in properties that satisfy the test assumptions, such as variance homogeneity.

Probably the most frequently used transformation is the logarithmic transformation, obviously restricted to data with positive values. Examples in which the log transformation is recommended include the following situations. (1) The data have a lognormal distribution. (The transformation results in normally distributed data.) (2) The coefficient of variation ($S/\bar{X}$) of the observations, $X$, is constant. (Log $X$ has approximately constant variance.) (3) If the data consist of a simple exponential function of the independent variable $X$ (that is, $Y = Ae^{BX}$), then log $Y$ is linear with respect to $X$. One should consider the effects of the log transformation on the variance as well as on the distribution of the data. The transformation may make skewed data appear more normal, but at the same time, it may result in nonhomogeneity of variance, heteroscedasticity. Ideally, the transformation results in satisfying both the normality and variance homogeneity assumptions implicit in the usual tests. Fortunately, this transformation often results in data that approximately satisfy both of these assumptions.

Other useful transformations are the arcsin transformation, which is used for proportions, and the square root transformation, which stabilizes the variance if the variance is proportional to the mean.

## *Detection of Outliers*

Although many rules have been proposed to detect aberrant values, or outliers, considerable judgment should be exercised before discarding suspect data. For example, if replicate values from an analysis were 5.10, 5.11, 5.09, and 5.30, the last value would immediately be suspected with or without a statistical test. When doubt arises, Table 11.16, “Dixon criteria for Testing an Extreme Mean,” can be used to decide if an extreme value belongs with the rest of the data. The data are first ordered numerically from low to high, $X_1$, $X_2$,... $X_k$, where there are $k$ values. For example, if $k$ is between 3 and 7, and $X_k$ is the extreme value, form the ratio $(X_k - X_{k-1})/(X_k - X_1)$. If this ratio is greater than the tabulated value, the extreme value is considered aberrant

(Table 11.16).

**Table 11.16:** Dixon criteria for testing an extreme mean

| ***k*** | | **Significance level 5 percent** |
|---|---|---|
| 3 | $r_{10} = (X_2 - X_1)/(X_k - X_1)$ if smallest value | 0.941 |
| 4 | is suspected; | 0.765 |
| 5 | $= (X_k - X_{k-1})(X_k - X_1)$ if largest value | 0.642 |
| 6 | is suspected. | 0.560 |
| 7 | | 0.507 |
| 8 | $r_{11} = (X_2 - X_1)/(X_{k-1})$ if smallest value | 0.554 |
| 9 | is suspected; | 0.512 |
| 10 | $= (X_k - X_{k-1})/(X_k - X_2)$ if largest valueis suspected. | 0.477 |
| 11 | $r_{21} = (X_3 - X_1)/(X_{k-1} - X_1)$ if smallest value | 0.576 |
| 12 | is suspected; | 0.546 |
| 13 | $= (X_k - X_{k-2})/(X_k - X_2)$ if largest value is suspected. | 0.521 |
| 14 | $r_{22} = (X_3 - X_1)/(X_k\text{-}_2 - X_1)$ if smallest value | 0.546 |
| 15 | is suspected; | 0.525 |
| 16 | $= (X_k - X_{k-2})/(X_k - X_3)$ if largest value | 0.507 |
| 17 | is suspected. | 0.490 |
| 18 | | 0.475 |
| 19 | | 0.462 |
| 20 | | 0.450 |
| 21 | | 0.440 |
| 22 | | 0.430 |
| 23 | | 0.421 |
| 24 | | 0.413 |

| | |
|---|---|
| 25 | 0.406 |

In the foregoing example, 5.30 appears to be an outlier. Since $k = 4$, form the ratio (5.30 – 5.11)/(5.30–5.09) = 0.9. Table 11.16 shows that at the 5% level with $k = 4$, a value of 0.765 or greater is significant. The value 5.30 is deemed to be an outlier and is rejected. An intelligent discussion dealing with outliers and various recommended procedures is presented in ASTM, E178–75.

## Regression

Regression is a form of curve fitting that can be used for descriptive or predictive purposes. The theory of regression analysis allows equations relating variables to be established; it also aids in understanding the behavior (reliability) of the estimated equation parameters.

## *Simple Linear Regression*

Simple linear regression is concerned with the fitting of straight lines, $Y = A + BX$, where A and B are the parameters of the line, the intercept, and slope, respectively. The dependent variable $Y$ is a response that is the outcome of an experiment and is variable, i.e. its outcome cannot be exactly predicted. The independent variable is considered to be known precisely. For example, if blood pressure is the dependent variable, any or all of the following may be independent variables, depending on the objective of the experiment: dose of drug, length of treatment, and weight of patient. Some examples in which regression analysis would be appropriate are (1) cholesterol lowering as a function of dose, (2) log plasma concentration as a function of time, (3) optical density as a function of concentration, (4) dissolution as a function of stearate concentration and (5) tablet assay as a function of tablet weight.

The objective of linear regression analysis is to fit the best straight line, $Y = a + bX$, from the experimental data, using least squares. The least squares line is defined as the line that makes [Σ(observed $Y$-calculated $Y$)$^2$] a minimum, where the "calculated $Y$" is obtained from the least squares line. From the methods of calculus, it can be shown that:

The slope, $b = [S(X - \bar{X})(Y - \bar{Y})]/[S(X - \bar{X})^2]$
The intercept, $a = \bar{Y} - b\bar{X}$

Various significance tests may be performed on the estimates of the parameters a and $b$ if the following assumptions are met:

1. $X$ is measured without error. (If both $X$ and $Y$ are subject to error, line fitting may be accomplished by other minimizing techniques.)
2. $Y$ is a variable with a normal distribution at each X, and the observed $Y$s are statistically independent.
3. The variance of $Y$ is the same at each $X$.

4. The *true* relationship between $X$ and $Y$ is a straight line.

With these assumptions, the least squares fit can be used to estimate the variance; the variance of the slope; the variance of the intercept; a predicted value of $Y$; and confidence limits at a new value of $X$. It can also be used to determine if the relationship is a straight line. (Multiple observations of $Y$ at least one value of $X$ are needed for the latter test.) All of the statistical tests and confidence intervals are based on normal curve theory. If the data are normal, it can be shown that estimates of the parameters $a$ and $b$ have normal distributions. The procedure for fitting straight lines and some relevant statistical tests are presented, using data derived from stability studies for illustrative purposes.

## *Fitting Lines to Stability Data*

Because federal regulations now require expiration dating of pharmaceutical products, statistical procedures play an important role in the analysis of stability data. The fitting of data to straight line models and determination of variability allow inferences and predictions of future potency, as well as of the time to a given level of degradation, to be made with probability qualifications. Proper design of such studies is extremely important. Design considerations include the number of points in time at which the drug will be analyzed, various storage conditions, the number of samples to be analyzed at each point in time, and the source of these replicates. FDA statisticians recommend that three batches at ambient conditions be used to estimate stability characteristics. Tablets should be taken from more than one bottle at each time period, especially during early development or marketing stages. Recommendations regarding an optimal choice of time periods for assay to reduce the variability of the regression line have appeared in the literature; however, in practice, the choice of time points seems to involve more than just this kind of optimality. Designs should include appropriate observation intervals to reveal possible lack of linearity in the stability plots.

Table 11.17 shows the concentration of intact drug in tablets as a function of time at 25°C. Three “randomly” selected tablets from a single batch were assayed at each sample time. (In the case of tablets taken from each of three batches, the statistical analysis would take the different batches into account, and would be more complex than what is presented here). The results are plotted in Fig. 11.11.

| **Table 11.17:** Stability data with triplicate assays from a single batch | |
|---|---|
| **Time** | **Concentration (% of label)** |
| 0 weeks | 102,102,104 |
| 8 weeks | 100,99,101 |
| 16 weeks | 98,99,98 |
| 24 weeks | 94,97,96 |
| 32 weeks | 97,95,93* |
| $\Sigma X = 240$* | $\Sigma Y = 1473$ |
| $\Sigma(X - \bar{X})^2 = 1920$ | $\Sigma(Y - \bar{Y})^2 = 137.33$ |

* Note that the sum of "time" values ($X$) equals 3×80 = 240 because each "time" appears 3 times, one for each value of the concentration ($Y$)

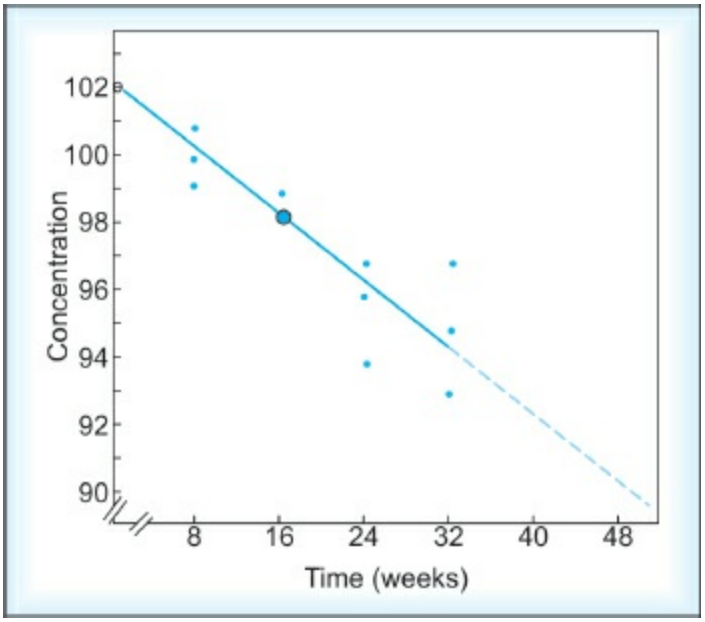



**Fig. 11.11:** Stability plot showing loss of drug with time in a tablet formulation

Some pertinent questions are (1) Do these data represent a straight line? (Does the decomposition follow zero-order kinetics?) (2) If so, at what time will the product be 10% decomposed? (3) How much intact drug will be present in one year? To answer these questions, estimates of the slope (b) and intercept (a) that define the straight line, and the variance estimate are

obtained using the least squares procedure as follows (*see* Table 11.17).

The sample estimate of the *slope*, *b*, is:

$$b = \Sigma(X - \bar{X})(Y - \bar{Y})/\Sigma(X - \bar{X})^2$$

A shortcut computing formula for *b* is:

$$b = (N\Sigma XY - \Sigma X\Sigma Y)/[N\Sigma X^2 - (\Sigma X)^2]$$

$$= \frac{15.[0(102) + 0(102) + \text{K } 32(93)] - (240 \times 1473)}{[(15)(0^2 + 0^2 + \text{K } 32^2) - (240)^2]}$$

$$= (346,920 - 354,000)/(86,400 - 57,600) = 0.246$$

The sample estimate of the *intercept*, *a*, is:

$$a = \bar{Y} - b\bar{X} = 98.33 - (0.246)(16) = 102.27$$

Therefore, the *equation of the fitted line* is:

$$C_T = 102.27 - 0.246T \quad \ldots (7)$$

where, $C_T$ is the concentration of drug at time T. The variance estimate, $S^2$, is:

$$S^2 = [\Sigma(Y - (a + bX))^2]/(N - 2)$$

A shortcut formula is:

$$S^2 = [\Sigma(Y - \bar{Y})^2 - b^2\Sigma(X - \bar{X})^2]/(N - 2)$$
$$= [137.33 - (-0.246)^2(1920)]/13$$
$$= 1.638$$

The variance, as computed here, is an estimate of the error in the line fitting, which includes tabletting variation (weight variation, mixing heterogeneity, and other random errors) and assay error, as well as variation due to the fact that a straight line might not be an accurate representation of the data. In the present example, an independent estimate of the tablet variation is available from the replicates at each observation point. A oneway ANOVA of these data is the first step in separating the sum of squares into its various components. There are five observation times with three replicates at each time.

The between times sum of squares can be further subdivided into two parts.

| ANOVA | | | |
|---|---|---|---|
| **Source** | ***DF*** | ***SS*** | ***MS*** |
| Between times | 4 | 119.33 | 29.83 |
| Within times | 10 | 18.00 | 1.80 |
| Total | 14 | 137.33 | |
| **Source** | ***DF*** | ***SS*** | ***MS*** |
| Regression | 1 | 116.03 | 116.03 |
| Deviations from regression | 3 | 3.30 | 1.10 |

The *regression* sum of squares is calculated as:

$$b^2\Sigma(X - \bar{X})^2 = (-0.246)^2 \times (1920)$$

and is the sum of squares due to the slope of the line. A slope of zero would result in a zero regression sum of squares; a large slope results in a large sum of squares. The deviations sum of squares is equal to between times sum of squares minus regression sum of squares, (119.33–116.03), and represents the variance due to the deviations of the average results at each time period from the fitted line.

**Test for linearity:** The *F* test for linearity has 3 and 10 degrees of freedom ($F_{3,10}$) and is equal to deviations MS/within times MS, or 1.10/1.80 = 0.61. This is not significant ($F_{3,10}$ = 3.71 at the 5% level. *See* Table 10–11-6). There is no evidence for lack of linearity. The nonsignificant test for linearity in this context suggests that the use of a straight line as a model for these data is reasonable. The F ratio compares the deviations of the means of the observations from the fitted line at each time period to the error determined from replicates within each time period. If this ratio is small, there is no reason to believe that the relationship is not linear. A lack of linearity would result in an inflated deviations mean square, because the data representing a nonlinear function would be far removed from the least squares line.

**Test of slope:** It is of interest to test if the slope differs from zero, i.e. the drug is indeed degrading. A zero slope means that no degradation is occurring. The test compares the sample estimate of the slope to its variance using a t test. The variance estimate of a slope is $S^2/\Sigma(X - \bar{X})^2$.

$$H_0{:}B{=}0 \qquad H_0{:}B \neq 0$$

$$t_{13} = |b - 0| / \sqrt{S^2 / \Sigma(X - \overline{X})^2}$$

$$= 0.246 / \sqrt{1.638^* / 1920} = 8.416$$

* 1.638 is the variance estimated from the original least squares fit. A safer estimate of error is the within mean square with 10 df equal to 1.80.

(Equivalently, $F_{1,13}$ = Regression *MS*/Error *MS* = 70.84 = $t^2$). The slope is significant (*see* Table 11.5); the drug is degrading.

**Test of intercept:** Although it may not be of special interest in this example, a test for the significance of the intercept may also be performed. In this example, one might wish to test if the intercept is different from 100. A *t* test is performed comparing the difference between the intercept estimate, a and 100 to its standard deviation.

$$H_0{:}A = 0 \qquad H_A{:}A \neq 100$$

$$t_{13} = |a - 100| / \sqrt{S^2[1/N + \overline{X}^2 / \Sigma(X - \overline{X})^2]}$$

$$= (2.27) / \sqrt{1.638\,(1/15 + 256/1920)} = 3.97$$

Thus, the intercept, 102.27, (potency at time 0) is significantly greater than 100, $P < 0.05$ (*see* Table 11.5).

**Prediction:** Based on the data, the best prediction for the time at which 10% of the drug is decomposed (or when 90% of the intact drug is present) is determined by rearranging Eq. (7).

$$T_{90} = (90 - 102.27)/(-0.246) = 49.9 \text{ weeks}$$

where, $C_T$ is 90 and $T_{90}$ is the time when 10% of the drug is degraded.

The amount of active drug predicted to be present after one year is calculated from Eq. (7):

$$C_{52} = 102.27 - 0.246\,(52) = 89.5\%$$

The estimates of both $T_{90}$ and $C_{52}$ just described are variable and have error associated with them. The error of a predicted value of *Y* (concentration, $C_{52}$, for example), where a *predicted value is an actual observation at time X*, depends on the magnitude of the variance, and how far away the new time is from the mean of the time values ($\overline{X}$) used to compute the least squares equation. The further the new value of *X* is from, ($\overline{X}$) the more variable is the

estimate of the predicted value. The variance of a predicted value is:

$$S_p^2 = S^2(\textit{predicted value})$$
$$= S^2[1+1/N+(X_T-\overline{X})^2/\Sigma(X-\overline{X})^2]$$

where, $X_T$ is the time of prediction. The predicted value, an assay actually determined at the time of prediction, has a variance that consists of the error due to the estimation involved in the line fitting *plus* the error associated with the new assay at the predicted time. A 95% confidence interval for a single assay performed at 52 weeks can be computed once the variance estimate of the predicted value is calculated. A 95% confidence interval equals $C_{52} \pm t_{df,0.05}(S_p)$. In the example, the confidence interval is:

$$89.5 \pm t_{13,0.05}\sqrt{1.638\,[1+1/15+(52-16)^2/1920]}$$
$$= 89.5 \pm 2.61\sqrt{2.85} = 89.5 \pm 3.65$$

A confidence interval for the "true" potency at 52 weeks (a sample would not actually be assayed at this point) is:

$$C_{52} \pm t_{13,0.05}\sqrt{S^2[1/N+(X_T-\overline{X})^2/\Sigma(X-\overline{X})^2}$$
$$= 89.5 \pm 2.38$$

The time at which 90% remains, for example, is known as inverse prediction, or a linear calibration problem, in which it is of interest to estimate the time, $T = (C_T - C_0)/b$, with an associated confidence interval. It can be shown that the confidence interval for $T$ is:

$$\left[(T-g\overline{T}) \pm (tS/b) \times \sqrt{[N+1)/N](1-g)+(T-\overline{T})^2/\Sigma(T-\overline{T})^2}\right] \div [(1-g)] \quad \ldots (8)$$

where, $g = (t^2)(S^2)/[b^2\ \Sigma(T - \overline{T})^2$ and this the appropriate tabled "$t$" value for the confidence interval. In the present example, the time for 10% decomposition, $T_{90}$, is 49.9 weeks. For a 95% confidence interval:

$$g = (2.16)^2\ (1.638)/(0.246)^2\ (1920) = 0.0658$$
($t_{13}$ for $\alpha = 0.05$ is 2.16)

(Note that if the slope $b$ is not significantly different from zero, $g$ will be greater than 1 and confidence limits cannot be obtained). A 95% confidence interval for $T_{90}$, the time for 10% decomposition, using equation (8) is:

$$[48.85 \pm (2.16 \times 1.28/0.246)$$
$$\times \sqrt{(16/15(0.9342 + (49.9 - 16)^2/1920}]$$
$$\div [0.9342] = 37.1 \text{ to } 67.5$$

Thus, the time for 10% decomposition probably occurs between 37.1 and 67.5 weeks. A conservative expiration date based on the time for 10% decomposition is 37.1 weeks, according to the lower limit of the confidence interval.

**Allocation of *X*:** The slope is estimated better if the $X$ values are maximally spread apart. This can be seen from inspection of the variance of the slope, which is equal to $\sigma^2$ divided by $\Sigma(X - \bar{X})^2$ This quantity is maximized if the assay points are equally divided between points at zero time and the last assay time. This is not usually done, however, because considerations are important in addition to this "optimal" allocation. For example, in stability testing, data are usually obtained at intermediate points to observe the functional relationship between concentration and time.

**Weighting in regression:** If the reaction is first order, a least squares procedure is followed using log $C$ for concentration. This transforms the exponential equation, $C = C_0 e^{-kt}$, to a linear function, $\log C = \log C_0 - kt$ 2.303. If the data are log-normal, i.e. if log $C$ is normal, then the variance homogeneity assumption is satisfied as a result of the log transformation if the coefficient of variation, $S/\bar{X}$ of the original untransformed data is approximately constant. If a log transformation is inappropriate (as in zero-order reactions) and the variance is not constant but depends on the magnitude of the value (e.g. the coefficient of variation is constant), a weighting procedure should be used when fitting the least squares line, with weights equal to the inverse of the variance.

The weighting procedure is used for any least squares fit in which the variance is not constant, but the procedure is more complex than the examples considered here. When fitting lines to the Arrhenius relationship, $\log k = -H_A/T + K$, a weighted least squares procedure should be used, because in general, the log $k$'s do not have equal variance. A nonlinear approach to the statistical analysis of stability data combining the first order and Arrhenius equations has recently been described.

## *Correlation*

Figure 11.12 is a scattergram showing the relationship of change in blood pressure after treatment and the pretreatment blood pressure measurement. If both variables are subject to error and are distributed bivariately normal, a correlation coefficient r can be computed and tested for significance:

$$r = \Sigma(X - \bar{X})(Y - \bar{Y}) / \sqrt{\Sigma(X - \bar{X})^2 \Sigma(Y - \bar{Y})^2}$$

where, the *X*'s and *Y*'s are paired values, e.g. change of blood pressure and baseline blood pressure values. The value of *r* lies between +1 and 1 and measures the linear relationship between two variables, *X* and *Y*. A correlation coefficient of +1 and 1 indicates that all points fall exactly on a straight line of positive slope or negative slope, respectively. A correlation of 0 indicates independence of the two variables, a zero slope (if they are bivariately normal). In practice, an exact fit (r = ± 1) or a zero correlation rarely occur and one must decide if the observed *r* is large enough to be taken seriously. The test of the correlation coefficient ($H_0$: $\rho = 0$) is a *t* test: $t = r / \sqrt{(1 - r^2)/(N - 2)}$ with *N*–2 df. For data in (Fig. 11.12), where *N* = 19 and *r* = 0.478, the test of significance (17 df), *t* = (0.478 – 0)/ $\sqrt{(1 - 0.228)/(17)} = 2.24$, shows a significant correlation at the 5% level (*see* Table 11.5).

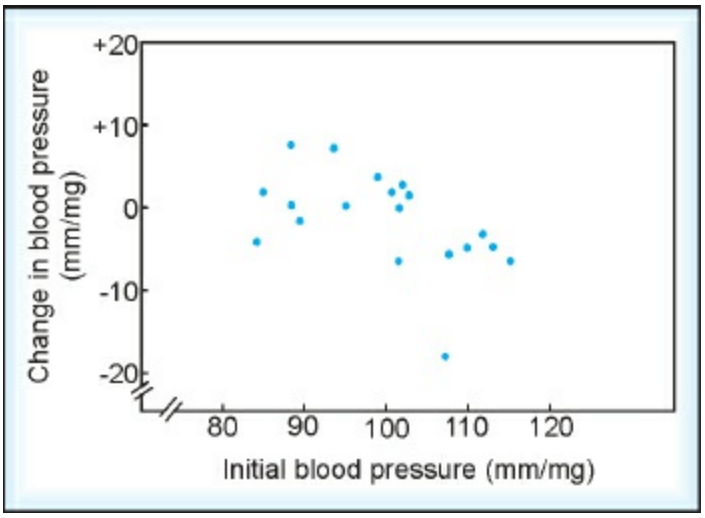



**Fig. 11.12:** Scattergram showing the relationship of posttreatment change in blood pressure to pretreatment blood pressure

Correlations should be carefully considered since a significant correlation does not necessarily indicate cause and effect. For example, a strong

correlation between plasma uric acid increase and potassium decrease in patients on diuretics does not mean that one causes the other. Also, it should be appreciated that a small, perhaps meaningless, correlation coefficient may be statistically significant when dealing with large sample sizes. In general, one should interpret correlations with caution and use such results as clues for further experiments.

## Nonparametric Statistics

Nonparametric statistics, also known as distribution-free statistics, may be applicable when the nature of the distributions are unknown and we are not willing to accept the assumptions necessary for the application of the usual statistical procedures. Most of the statistical tests described in this chapter are based on an assumption that the underlying distribution of the data is known (binomial or normal, for example). This assumption, although never exactly realized and occasions arise in which data are clearly too far from normal to accept the assumption of normality. The data may deviate so much from that expected for a normal distribution that to assume normality, even when dealing with means, would be incorrect.

Many nonparametric tests assume only that the underlying distribution is continuous. In addition, analysis of data using nonparametric methods has the advantage of using simple calculations that are often based on ordering or ranking procedures. Thus, these methods can be used to obtain a quick look before a full-fledged analysis is undertaken. Various nonparametric tests and their applications are shown in Table 11.18.

A disadvantage is that these tests lack power compared with corresponding parameteric tests; nevertheless, some nonparametric tests are surprisingly powerful. (As noted previously, power is the ability of a test to find significance, should a true difference exist.) Also, in more complex designs, nonparametric tests may not give the variety of analyses and interpretations given by parametric analyses. Several popular methods of analysis are presented in this discussion; more detail is given by Wilcoxon and Wilcox.

**Table 11.18:** Nonparametric tests and their applications

| Sr. No. | Test name | Application |
|---|---|---|
| 1. | The sign test | Used to compare two treatments in a paired design, useful for a quick assessment of the results of an experiment. |
| 2. | Wilcoxon signed rank test | Used to compare two treatments in a paired design, more sensitive than |

| | | |
|---|---|---|
| | | sign test. |
| 3. | Wilcoxon rank sum test (also known as the Mann-Whitney U-test) | Used to compare the averages of two treatments, test for differences between two independent groups. |
| 4. | Kruskal-Wallis test (one-way anova) | Used to compare the averages of more than two treatments. |
| 5. | Friedman test (two-way analysis of variance) | Used to compare the averages of more than two treatments, the treatments are ranked *within each block* (e.g., animal or person), disregarding differences between blocks. |

## *The Sign Test*

The sign test is a popular nonparametric test used to assess the significance of differences of paired data. The underlying distribution that represents the data need not be precisely defined, as opposed, for example, to the assumption of normality necessary for the *t* test. The sign test, however, has less power to differentiate the two treatments than a test in which the actual distribution of data is taken into account. This means that for a given set of data, the sign test may not result in a significant difference, whereas an appropriate parametric test (such as the *t* test) might show significance. The sign test has the advantage of using simple binomial calculations and it is useful for a quick assessment of the results of an experiment. The procedure for performing the sign test follows:

1. The differences of each of the paired samples are tabulated, indicating only whether one treatment or factor has a higher or lower value than the other. For example, a positive difference means that the second treatment gives higher results, and a negative difference means that the first treatment gives higher results. In case of a tie, the difference is ignored. With continuous data, there should be no ties, but ties do occur because of limitations of measuring techniques and/or because the data are not really continuous.

2. After tabulation, the proportion of "wins" (positive differences, for example) is calculated. The observed proportion of "wins" is compared to that expected under the null hypothesis of equality of treatments ($H_0$: $p_0$ = 0.5), using a one-sample binomial test.

In the following example (Table 11.19), tablets were taken from two different punches of a tablet press at various times during a run because a difference in weight had been suspected. In 18 of 24 cases, the tablet from the left side had a higher weight than that on the right side (a positive difference for the left side minus right side). There is one tie (7:00), which is disregarded for the purposes of the statistical test. The observed proportion p is 18/24 = 0.75 and the hypothetical proportion $P_0$ is 0.5.

$$Z = (|0.75 - 0.5| - 1/48)/\sqrt{(0.5)(0.5)24} = 2.25$$

**Table 11.19:** Weight differences of tablets taken from right and left sides of tablet press

| Time | Right | Left | Δ | Time | Right | Left | Δ |
|---|---|---|---|---|---|---|---|
| 1:00 PM | 220 | 221 | + 1 | 4:15 | 218 | 219 | + 1 |
| 1:15 | 221 | 220 | – 1 | 4:30 | 222 | 223 | + 1 |
| 1:30 | 219 | 223 | + 4 | 4:45 | 226 | 228 | + 2 |
| 1:45 | 218 | 221 | + 3 | 5:00 | 217 | 227 | + 10 |
| 2:00 | 223 | 218 | – 5 | 5:15 | 219 | 220 | + 1 |
| 2:15 | 217 | 213 | – 4 | 5:30 | 215 | 218 | + 3 |
| 2:30 | 221 | 225 | + 4 | 5:45 | 220 | 224 | + 4 |
| 2:45 | 218 | 220 | + 2 | 6:00 | 219 | 220 | + 1 |
| 3:00 | 226 | 224 | – 2 | 6:15 | 223 | 221 | – 2 |
| 3:15 | 220 | 223 | + 3 | 6:30 | 216 | 220 | + 4 |
| 3:30 | 217 | 219 | + 2 | 6:45 | 222 | 226 | + 4 |
| 3:45 | 224 | 223 | – 1 | 7:00 | 221 | 221 | 0 |
| 4:00 | 222 | 225 | + 3 | | | | |

Since $Z$ is greater than 1.96, the difference is significant at the 5% level. The tablets from the left side tend to have higher weights than those from the right side, suggesting that an adjustment on the tablet press should be made.

## *Wilcoxon Rank Sum Test*

This test is used to compare the averages of two treatments and has excellent power compared with the more powerful t test if the data are normally distributed. Consider the example in Table 11.20 showing changes in weight

of control animals compared with animals given an anorexic drug. Values for drug and control groups are ranked in order of magnitude. A rank is assigned to each value; *average ranks* are assigned in case of *ties*. The ranks are then summed within each treatment group. In most cases, a significance test may be used based on an approximation to the normal distribution:

$$Z = [T - E(T)] / \sqrt{S_T^2}$$

has a distribution that is approximately normal with variance equal to 1.

$$E(T) = N_1 (N_1 + N_2 + 1)/2 \text{ and } S_T^2 = N_1 N_2 (N_1 + N_2 + 1)/12$$

$N_1$ = Smaller sample size

$N_2$ = Larger sample size

$T$ = Sum of ranks for smaller sized sample

In this example:

$$E(T) = (8)(8 + 9 + 1)/2 = 72$$
$$S_T^2 = (8)(9)(8 + 9 + 1)/12 = 108$$

For a two sided test:

$$Z = [|94 - 72|] / \sqrt{108} = 2.12$$

**Table 11.20:** Weight change in drug and control groups and ranks

| Control | | Drug | |
|---|---|---|---|
| Weight Change | Rank | Weight Change | Rank |
| 0 | 14 | −2 | 10.5 |
| −3 | 9 | −8 | 3 |
| +9 | 17 | +1 | 15 |
| −1 | 12.5 | −19 | 1 |
| −4 | 7.5 | −4 | 7.5 |
| +3 | 16 | −2 | 10.5 |
| −1 | 12.5 | −11 | 2 |
| −5 | 5.5 | −5 | 5.5 |
| | | −7 | 4 |
| Sum of Ranks = | 94 | Sum of Ranks = | 59 |

According to Table 11.3, a value greater than 1.96 is needed for significance at the 5% level. Therefore, the drug is significantly different

from the control ($P < 0.05$); the drug group showed a greater weight loss. Wilcoxon and Wilcox provide a table that can be used to assess significance, given the sample sizes of the two groups and the rank sum.

For paired data, the non-zero differences (zero differences are not included in the analysis) of the pairs are first ranked in order, *disregarding sign*. Then, the sum of the ranks of the positive differences and the sum of the ranks of the negative differences are computed and tested for significance. The data in Table 11.21 were obtained from a bioavailability study using 15 subjects. The peak serum concentration is compared for the two products. The sum of the positive ranks is 68.5 and the sum of the negative ranks is 36.5. An approximate "normal" test is:

$$Z = [T - (N)(N+1)/4]\sqrt{N(N+1)(2N+1)/24}$$

where, $N$ is the number of pairs, and $T$ is the sum of the ranks (either positive or negative).

$$Z = [68.5 - (14)(15)/4]\sqrt{(14)(15)(29)/24} = 1.00$$

**Table 11.21:** Results of bioavailability comparison peak height of serum concentration

| Subject | Test Product | Control Product | Difference | Rank |
|---|---|---|---|---|
| 1 | 9.6 | 9.4 | +0.2 | 1.5 |
| 2 | 3.3 | 3.3 | 0 | Omit |
| 3 | 2.8 | 2.4 | +0.4 | 5.5 |
| 4 | 5.0 | 4.1 | +0.9 | 10 |
| 5 | 6.4 | 4.7 | +1.7 | 14 |
| 6 | 2.8 | 3.5 | −0.7 | 8 |
| 7 | 4.0 | 3.7 | +0.3 | 3.5 |
| 8 | 3.2 | 3.0 | +0.2 | 1.5 |
| 9 | 4.3 | 3.3 | +1.0 | 11 |
| 10 | 4.7 | 6.2 | − 1.5 | 13 |
| 11 | 3.3 | 2.9 | +0.4 | 5.5 |
| 12 | 4.6 | 3.8 | +0.8 | 9 |
| 13 | 4.0 | 5.1 | − 1.1 | 12 |
| 14 | 5.9 | 5.3 | +0.6 | 7 |
| 15 | 3.6 | 3.9 | 0.3 | 3.5 |

Since $Z$ is less than 1.96, the test product is not significantly different from the control at the 5% level. Wilcoxon and Wilcox have a table for significance testing and for small sample sizes, this table rather than the foregoing approximate formula should be used.

## *One-way and Two-way Designs*

Nonparametric tests are available for one and two-way ANOVA type designs. In the case of more than two independent groups arranged in a one-way design, the data are first ranked in order over all groups, disregarding group designation. The sum of the ranks for each group is then calculated. A statistic with an approximate $X^2$ distribution can be computed as shown in the following example, in which the hardness of tablets of three different formulations of the same drug are compared (Table 11.22). The replicates are randomly selected tablets. The $X^2$ test, with $C - 1 = 3 - 1 = 2$ df, is:

$$X^2_{C-1} = [12/N(N+1)]\Sigma R_i^2/N_i - 3(N+1)$$

where, $C$ is the number of treatments, $R_i$ is the sum of ranks in ith treatment, $N_i$ is the number of observations in $i$th treatment and $N$ is the total number of observations.

In this example:

$$X^2_2 = [12/(10)(11)][208.33 + 65.33 + 64.0] - (3)(11) = 3.84$$

For significance at the 5% level, a $X^2$ value with 2 degrees of freedom must exceed 5.99. Therefore, the differences among the three formulations are not significant at the 5% level.

**Table 11.22:** Hardness (rank) of tablets from three formulations

| Formula | 1 | 2 | 3 |
|---|---|---|---|
| | 8.3 (6) | 7.9 (4) | 8.4 (7) |
| | 10.0 (10) | 7.1 (2) | 8.0 (5) |
| | 9.7 (9) | 8.5 (8) | 6.5 (1) |
| | | | 7.3 (3) |
| $R_i = \Sigma$Ranks | 25 | 14 | 16 |
| $(\Sigma R_i^2)/N_i$ | 208.33 | 65.33 | 64.0 |

## *Friedman's Two-way Analysis*

Friedman's analysis for related samples with more than two groups is also based on ranking of data. This test is analogous to a two-way ANOVA.

Four batches of tablets were produced on three tablet presses (machines) and the average weight of the tablets was estimated as the mean of 20 randomly selected tablets. To test for machine differences, the machines are ranked *within each batch* as shown in Table 11.23. An approximate $X^2$ test can be used to assess significance using the following formula.

$$X^2_{C-1} = [12/(RC(C+1))]\Sigma C_j^2 - 3R(C+1)$$

where, $C$ is the number of groups (machines), $R$ is the number of blocks (batches) and $C_j$ is the sum of ranks in the *jth* group.

In this example:

$$X^2_2 = [12/(4 \cdot 3 \cdot 4)] \cdot 218 \cdot (12)(14) = 6.5$$

Since 6.5 is greater than the tabulated value of $X^2_2$ (5.99) at the 5% level, 5 the result is significant; the first machine shows the higher average weight. (Wilcoxon and Wilcox provide a discussion of comparisons that involve more than two groups).

The tests described in this section, as well as other nonparametric tests, are described in more detail by Siegel, Wilcoxon and Wilcox and Hollander and Wolfe. Tables for tests of significance are also provided by these sources.

**Table 11.23:** Data showing comparison of tablet weights from four machines using three batches with ranks

| | Machine | | | | | |
|---|---|---|---|---|---|---|
| | 1 | | 2 | | 3 | |
| Batch | Weight | Rank | Weight | Rank | Weight | Rank |
| 1 | 202 | (3) | 199 | (2) | 197 | (1) |
| 2 | 203 | (3) | 199 | (2) | 198 | (1) |
| 3 | 205 | (3) | 200 | (2) | 196 | (1) |
| 4 | 202 | (3) | 197 | (1) | 198 | (2) |

## OPTIMIZATION

The word “optimize” means ‘*to make as perfect, effective or functional as possible*’. Optimization is often used in pharmacy for formulations and their processing and practically, it may be considered as the search for a result that is satisfactory and at the same time the best possible within a limited field of search.

For optimization of a formulation, the formulator has two types of problems in front of him. These include constrained and unconstrained conditions. Constrained problems are those which have some physical restrictions placed on them where as unconstrained are independent. These can be better understood with the help of an example. Suppose, the formulator has to make a hardest tablet possible. This is unconstrained. Where as if it is being suggested that make hardest tablet possible which should disintegrate in less than 15 min., the restriction is placed and thus makes the problem more constrained type.

The development of a pharmaceutical formulation and their process involves several variables. Mathematically, they can be divided into two groups: independent and dependent variables. Independent variables are those which are under the control of the formulator and dependent variables (also called as the responses) are those which are the outcomes of the change in the independent variables. There is always a relationship between a given response and the independent variables which defines a response surface as shown in Fig. 11.13. It is this surface which must be evaluated to find the values of the independent variables $X_1$ and $X_2$ which give the most desir1able level of the response, $Y_1$.

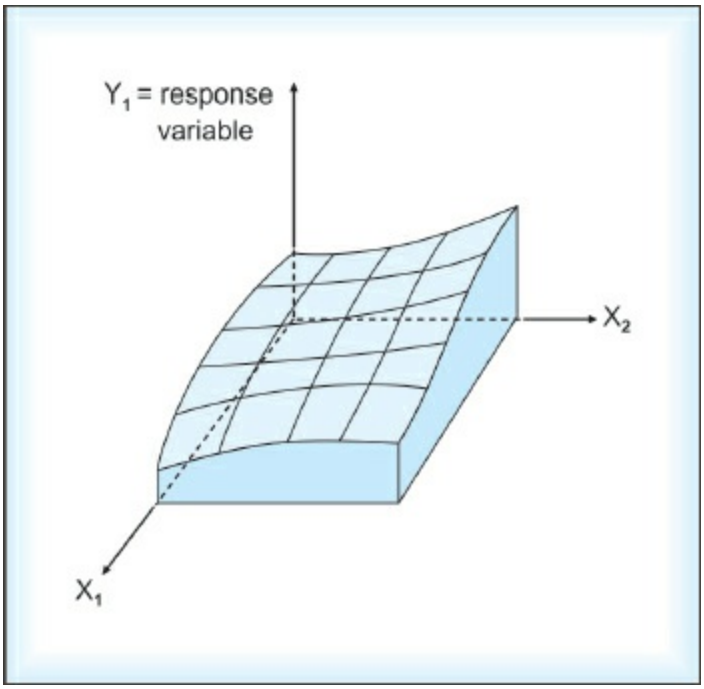



**Fig. 11.13:** Response surface representing the relationship between the independent variables $X_1$ and $X_2$ and the response $Y_1$

## Methods for Optimization

When optimizing a formulation or process, there are a number of different methods for tackling the problem and the resulting data may also be analysed in a number of different ways. The methods most widely used for optimization can be divided into two major groups: one in which experimentation continues as the optimization study proceeds (sequential simplex method and evolutionary operations), and the second one in which the experimentation is completed before the optimization takes place (classical mathematical methods, i.e. lagrangian method, canonical analysis and search methods).

The second group of methods necessitate that the relationship between dependent and independent variables be known, which can be possible by two approaches: the theoretical and the empirical. If the experimenter knows a priori the theoretical equation for the formulation properties of interest, no experimentation is necessary. To apply the empirical or experimental approach for a system with one or more independent variables, the experimenter experiments at several levels, measures the property of interest, and obtains a relationship, usually by simple or multiple regression analysis or by the least-squares method. Usually, least squares procedures are used to obtain an empirical polynomial equation from experimental data that adequately describes the system within the range of the test variables, the levels of which are fixed in advance. Predictions and optimization are then based on the polynomial equation. Any set of data can be fit exactly by a polynomial of sufficient degree. For example, an equation of the form $y = a + bx + cx^2$ can be fit exactly to three points ($x$, $y$ pairs). This does not mean that such an equation has physical meaning in describing the system or that it will accurately predict responses at extra-design points, i.e. combinations of factors not included in the experiment. These procedures usually result, however, in equations that closely approximate the response as a function of the variables being studied, and although the procedure may seem chancy, it has good predictive properties if used cautiously and intelligently.

### *Sequential Simplex Method*

This method has been used widely over the past 30 years, its success as much owing to its simplicity as to its efficiency. Unlike the other methods, it

assumes no mathematical model for the phenomenon or phenomena being studied. It is *sequential* because the experiments are analysed one by one, as each is carried out. The simplex which lends its name to this optimization method is a convex geometric figure of $k$ +1 non-planar vertices in $k$ dimensional space. For two dimensions it is a triangle, for three dimensions it is a tetrahedron.

This procedure may be used to determine the relative *proportion* of ingredients that optimizes a formulation with respect to a specified variable(s) or outcome. A common problem in pharmaceutics occurs when the components of a formulation are varied in an attempt to optimize its performance with respect to such variables as drug solubility, dissolution time and hardness. Application of a simplex design can be used to help solve this problem. The method is illustrated using data estimated from a publication by Fonner and co-workers, in which a different approach, a constrained optimization, is described. In the present example, three components of the formulation will be varied—stearic acid, starch, and dicalcium phosphate—with the restriction that the sum of their total weight must equal 350 mg. The active ingredient is kept constant at 50 mg; the total weight of the formulation is 400 mg. The formulation can be optimized for more than one attribute, but for the sake of simplicity, only one effect, dissolution rate, is considered.

The arrangement of the three variable ingredients in a simplex is shown in Fig. 11.14. Note that this simplex is represented by a triangle. With more than three ingredients, the representation of the simplex is more difficult. The simplex, in general, is represented by an equilateral figure, such as a triangle for the three-component mixture and a tetrahedron for a four-component system. Each vertex represents a formulation containing either (1) a pure component or (2) the maximum percentage of that component, with the other two components absent or at their minimum concentration. The choice of upper (maximum) and lower (minimum) concentrations of the variable ingredients is usually based on judgment, experience, or data from preliminary experiments, and represents concentrations within which a viable product can be manufactured. In this example, the vertices represent mixtures of all three components, with each vertex representing a formulation with one of the ingredients at its maximum concentration. The reason for not using pure components is that a formulation containing only one of the three

components (350 mg of only starch, for example) would result in an unacceptable product. Careful preliminary thought must be given to the choice of the upper and lower concentrations of the three ingredients, under the constraint that the total weight of the components is fixed. In this case, the lower and upper limits are stearic acid 20 to 180 mg (5.7 to 51.4%*); starch 4 to 164 mg (1.1 to 46.9%); dicalcium phosphate 166 to 326 mg (47.4 to 93.1%). Thus, as shown in Fig. 11.11, the vertex associated with the maximum percentage of starch would be represented by a formulation containing 164 mg of starch (46.9%), 166 mg of dicalcium phosphate (47.4%) and 20 mg of stearic acid (5.7%).

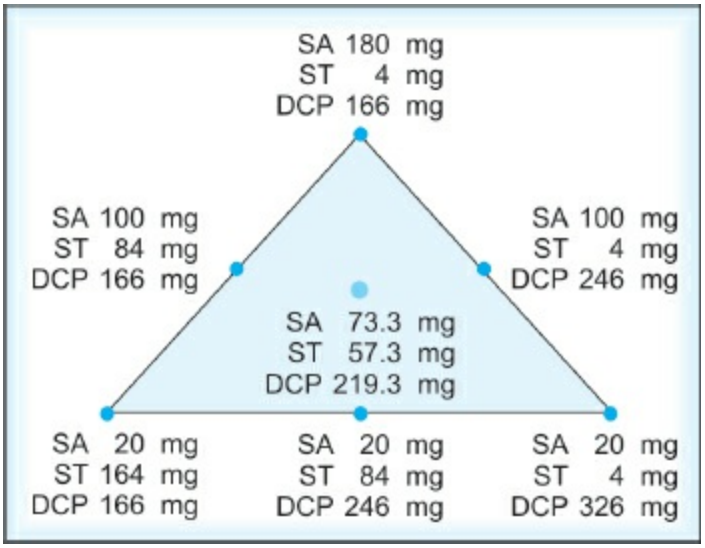



**Fig. 11.14:** Special cubic simplex design for a three-component mixture. Each point represents a different formulation. SA = stearic acid, ST = starch; DCP = dicalcium phosphate

Various formulations can be studied in this triangular simplex space. One basic simplex design includes formulations at each vertex, halfway between the vertices, and at one center point as shown in Fig. 11.11. Note that a formulation represented by a point halfway between two vertices contains the average of the minimum and maximum concentration of the two ingredients represented by the two vertices. The composition of these seven formulas in the present example, a three-component system, is shown in Table 11.24.

If the vertices in the design are *not* single pure substances (100%), as is the case in this example, the computation is made easier if a simple transformation is initially performed to convert the maximum percentage of a

component to 100%, and the minimum percentage to zero (0%), as follows:

$$\text{Transformed \%} = \frac{(\text{Actual \%} - \text{Minimum \%})}{\text{Maximum \%} - \text{Minimum \%}}$$

In the case of stearic acid, for example, the transformation is (Actual % – 5.7)/(51.4 – 5.7). An actual concentration of stearic acid of 28.6% is transformed to a concentration of 50%, using this formula. Figure 11.11 shows that the simplex points nicely cover the space in a symmetric fashion. This simplex arrangement allows easy construction of an equation that exactly fits the resulting data, a polynomial equation with seven terms.

**Table 11.24:** Seven formulas to be tested-actual and (transformed) values

| Stearic acid (%) | Starch (%) | DCP (%) | Response (min) | | ln |
|---|---|---|---|---|---|
| 51.4 (100) | 1.1 (0) | 47.4 (0) | 292 | Y(l) | 5.68 |
| 5.7 (0) | 46.9 (100) | 47.4 (0) | 5.6 | Y(2) | 1.72 |
| 5.7 (0) | 1.1 (0) | 93.1 (100) | 50.4 | Y(3) | 3.92 |
| 5.7 (0) | 24.0 (50) | 70.2 (50) | 15.6 | Y(2,3) | 2.75 |
| 28.6 (50) | 24.0 (50) | 47.4 (0) | 25.6 | Y(l,2) | 3.24 |
| 28.6 (50) | 1.1 (0) | 70.2 (50) | 124.5 | Y(l,3) | 4.82 |
| 20.9 (33) | 16.4 (33) | 62.6 (33) | 37 | Y(l,2,3) | 3.61 |

$$\text{Response} = b_1X_1 + b_2X_2 + b_3X_3 + b_{12}X_1X_2 + b_{13}X_1X_3 + b_{23}X_2X_3 + b_{123}X_1X_2X_3 \quad \ldots (9)$$

where, $X_1$, $X_2$ and $X_3$ represent transformed percentage concentrations of stearic acid, starch and dicalcium phosphate, respectively. This empirical equation represents the data in the confines of the simplex space, and the coefficients can be calculated as simple linear combinations of the responses as follows, using In $Y$, the response variable recommended by Fonner. As shown in Table 11.19, the responses are $Y(1)$, $Y(2)$ and so forth where, for example, $Y(1)$ is the response for the formulation with $X_1$ (stearic acid) at a maximum concentration.

$b_1 = Y(1) = 5.68$
$b_2 = Y(2) = 1.72$
$b3 = Y(3) = 3.92$
$b_{12} = 4Y(1,2) - 2Y(1) - 2Y(2) = -1.83$
$b_{13} = 4Y(1,3) - 2Y(1) - 2Y(3) = 0.10$

$b_{23} = 4Y(2,3) - 2Y(2) - 2Y(3) = -0.30$

$b_{123} = 27Y(1,2,3) - 12\,[Y(l,2) + Y(l,3) + Y(2,3)] + 3[Y(1) + Y(2) + Y(3)] = 1.71$

Substituting these values of the coefficients into Eq. (9), the following response equation is obtained.

Dissolution Time

$= Y$

$= 5.68X_1 + 1.72X_2 + 3.92X_3$

$-l.83X_1X_2 + 0.l0X_1X_3$

$-0.30X_2X_3 - 1.71X_1X_2X_3 \ldots$ (10)

This is an empirical equation that should represent the response surface in the simplex space. Its adequacy can be tested by running one or more experiments at other experimental points (different formulations from those in the simplex) and noting if the equation accurately predicts the results. In the analysis of their experiment, Fonner and co-workers included data from four experiments not covered by the simplex. Using equation (10), the prediction of the results for these four formulations is good, as can be seen in Table 11.25. The last point in Table 11.25 represents a formulation outside the region of the simplex and the prediction is good in this case. In general, predictions outside of the simplex space might not be reliable, and caution should be exerted when extrapolating into extra-design regions. When possible, replication is recommended in simplex experiments to estimate the variance, which can be used to assess the "fit" of the model by comparing the predicted and actual results from extra-design points.

**Table 11.25:** Prediction of dissolution results of extra-design points using equation derived from simplex experiment

| Extra-design formulations (transformed) | | | Observed dissolution | Dissoluton predicted from Eq (10) |
|---|---|---|---|---|
| Stearic acid (%) | Starch (%) | DCP (%) | | (antilog)* |
| 20.4 | 21.6 | 58.2 | 33.4 | 38.9 |
| 20.4 | 65.3 | 14.4 | 12.9 | 12.7 |
| 64.1 | 21.6 | 14.4 | 72.8 | 72.9 |
| 64.1 | 65.3 | −29.3 | 19.9 | 22.1 |

* Fonner and associates recommend use of ln (response) for the optimization using their method. This transformation is used here

Computer programs can be used to construct contour maps and identify optimal regions once the response equation has been established. A number of well-designed and tested statistical software packages are available commercially (Table 11.26). Several companies have adapted these experimental analysis techniques to computer software but have kept the programs in-house.

**Table 11.26:** Statistical software with their applications

| Statistical software | Applications |
|---|---|
| System for statistical analysis (SAS) | Conduct simulation studies with random number generators for many different distributions |
| Number Cruncher Statistical System (NCSS) | It includes documented statistical and plot procedures |
| Design Expert | It includes propagation of error for mixtures, crossed designs and transformed responses, as well as response surface methodology |
| Data Plot | Used to generate majority of sample outputs and graphics |
| Cadila Sytem (developed by Cadila Pharm. Ltd.) | Expert system for formulation of tablets for active ingredients based on their physical, chemical and biological properties |
| Sanofi System (Sanofi Research Division of Philadelphia) | Formulation of hard gelatin capsules based preformulation data on the active ingredient) |
| Capsugel System (School of Pharmacy, University of London) | Formulation of hard gelatin capsules, excipients used and their properties, analysis of marketed formulations from some countries |
| Zeneca System (Zeneca Galenical development system, Heidelberg (Department of Pharmaceutics and Biopharmaceutics and Department | Tablet formulations and their optimisation System for development of range of formulations aerosols, intravenous |

| | |
|---|---|
| of Medical Informatics at the University of Heidelberg, Germany) | injection solutions, capsules (hard shell powder) and tablets (direct compression), each incorporating information on all aspects of that dosage formulation (e.g. properties of the excipients to be added, compatibility, processing operations, packaging and containers and storage conditions) |

## Choosing an Optimization Method

The various methods are listed in Table 11.27, with a very brief and approximate summary of the circumstances in which they are used.

The choice of method should be dependant on the previous steps and probably on our ideas about how the project is likely to continue.

The criteria are:

- Are there one or several responses to optimize?
- Is there a known mathematical model which describes the response adequately within the domain?
- Do we expect to continue the optimization *outside* the experimental region that has been studied up to this point?
- Is it necessary to map the response surfaces about the optimum? Can the experimental conditions be changed slightly without unacceptable variation in the responses?

**Table 11.27:** Choosing an optimization method

| Sr. No. | Method | Circumstances for use | Examples |
|---|---|---|---|
| 1. | Sequential simplex method | No mathematical model<br>Direct optimization<br>Single or multiple response | Optimization of a tablet formulation with three independent variables, i.e. binder, disintegrant and lubricant levels. |
| 2. | Evolutionary operations | Industrial situations<br>Little variation possible | Optimization of a tablet formulation and inspection system for parenteral products |
| 3. | Lagrangian method | Mathematical model<br>Multiple response<br>Normally no more than 2 factors | Optimization of a tablet formulation with two independent variables, i.e. disintegrant and lubricant levels. |
| 4. | Conical analysis | Second order model<br>Optimum outside domain<br>Single response | Optimization of controlled drug release from pellet system prepared by extrusion-spheronization. |
| 5. | Search method | Mathematical model of any order multiple response from 2 to 6 factors | Optimization of a tablet formulation with five independent variables, i.e. diluent ratio, compression force, disintegrant, binder and lubricant levels. |

## Section III:

# Pharmaceutical Dosage Forms

# 12:
# Pharmaceutical Excipients and Polymers

## PHARMACEUTICAL EXCIPIENTS

Today's commercially available excipients provide a gamut of required functions, from processing aids that increase lubricity, enhance flow ability, and improve compressibility and compatibility to agents that impart a specific functional property to the final product (e.g. modifying drug release). The International Pharmaceutical Excipients Council (IPEC) defines an excipient as any substance other than the active drug or prodrug that is included in the manufacturing process or is contained in a finished pharmaceutical dosage form. Nonactive pharmaceutical excipients are chemicals with a wide range of molecular sizes, from small molecules to large polymers, and a large variety of unique physicochemical characteristics. Therefore, pharmaceutical excipients offer a wide range of properties to influence many characteristics of a pharmaceutical product, thereby achieving the optimal therapeutic efficacy.

The overall contribution of excipients in dosage form designing can be better appreciated from the fact that more than 70% of the formulations contain excipients at a concentration higher than the drug. In reality, no single excipient would satisfy all the criteria; therefore, a compromise of the different requirements has to be made. For example, although widely used in pharmaceutical tablet and capsule formulations as a diluent, lactose may not be suitable for patients who lack the intestinal enzyme lactase to break down the sugar, thus leading to the gastrointestinal tract symptoms such as cramps and diarrhoea. Excipients are not inactive and have substantial impact on the manufacture and quality, safety, and efficacy of the drug substance(s) in a dosage form. Further, variability in the performance of an excipient is a key determinant of dosage form performance. The role of excipients varies substantially depending on the individual dosage form. These include (i) modulating solubility and bioavailability of the drug, (ii) enhancing stability of the drug in its dosage forms, (iii) helping drug to maintain a suitable polymorphic form, (iv) maintaining pH and osmolarity of liquid products, (v) acting as antioxidants, suspending agent, emulsifier, aerosol propellants, base, tablet diluent, (vi) preventing aggregation or dissociation and (vii) modulating the immunogenic response of drug (e.g. adjuvants) and many others. In these various contexts, excipients and issues associated with them can be considered in the following different areas.

“Functionality”: An excipient interacts with the drug in the dosage form and/or provides a matrix that affects critical quality attributes of the drug, including solubility, stability and bioavailability. Limited understanding of excipient functionality can compromise product quality and process control. “Safety and efficacy”: Excipients, can be associated with adverse events, either by direct action or by formation of undesirable adducts. Excipients can improve the immunogenic properties of vaccines by acting as adjuvants. By modifying pharmacokinetic parameters such as absorption and distribution, excipients can change exposure patterns and thus influence both safety and efficacy outcomes. “Processability”: good understanding of the characteristics and functional contributions of excipients aid in the day-to-day manufacture of a dosage form. In addition to their functional performance, ideally, excipients should be chemically stable, nonreactive with the drug and other excipients, inert in the human body, have low equipment and process sensitivity, have pleasing organoleptic properties, and are well characterized and well accepted by the industry and regulatory agencies. Approximately 800 excipients are currently used in the marketed pharmaceutical products and this number is expected to grow with new therapeutic categories, such as gene therapy and cell therapy, and new drug delivery technologies.

## CLASSIFICATION OF PHARMACEUTICAL EXCIPIENTS

Excipients are categorized as compendial or noncompendial materials. Compendial excipients are the better characterized excipients and most likely to possess the desirable qualities previously stated. These materials are recognized as preferred excipients for pharmaceutical formulations. Noncompendial excipients might also be used in pharmaceutical formulations. The use of these noncompendial materials is supported by Type IV drug master files (DMFs) in regulatory dossiers (i.e. new drug applications, abbreviated new drug applications, and investigational new drug applications). There may be approved drug products containing noncompendial excipients, thereby demonstrating the acceptance of these excipients by the US Food and Drug Administration or other agencies in the major markets. Excipients are of various origin: animal (e.g. lactose, gelatin, stearic acid), plant (e.g. starches, sugars, cellulose, alginates), mineral (e.g. calcium phosphate, silica) and synthesis (e.g. PEGs, polysorbates, povidone, etc.). Based on their nature, the pharmaceutical excipients are classified under chemical categories as alcohols, esters, ethers, glycerides, waxes, carbohydrates, hydrocarbons, minerals and proteins. The most widely used classification, however, is based on their use and objective of their addition in various dosage form (Fig. 12.1). In order to carry out the numerous functions required, new classes of excipients have now become available, derived from old and new materials, alone or in combination, adapted to the manufacture of high-performance pharmaceutical dosage forms. *USP-NF* lists 40 functional categories of excipients for pharmaceuticals, and many more are expected as new and increasingly complex-drug-delivery systems emerge and evolve. Few of the excipients used in formulation of solid dosage forms such as tablets and capsules, liquid, semisolid and aerosol dosage forms are described in detail in individual chapters. This chapter details organoleptic excipients, stabilizers, excipients for novel delivery systems, drug targeting, protein and vaccines delivery.

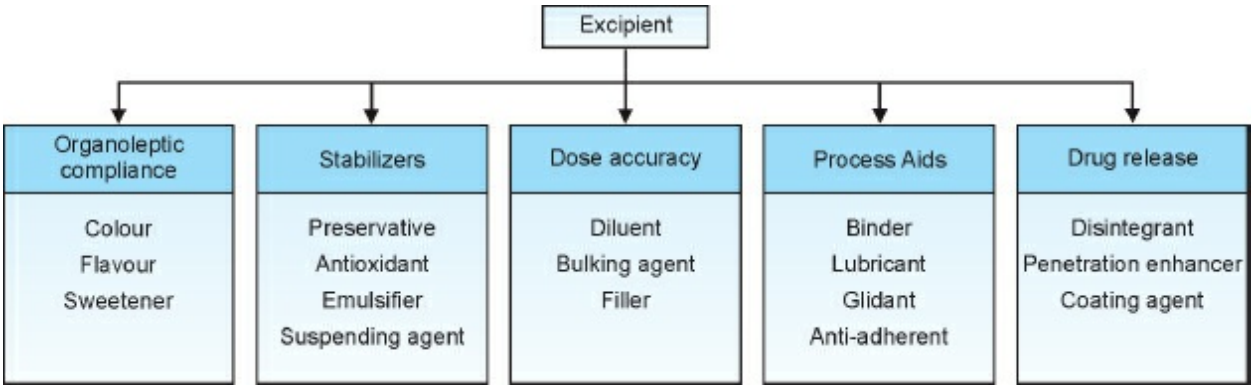



**Fig. 12.1:** Excipient classification based on objective of addition in dosage form

# ORGANOLEPTIC AGENTS

## Colours

Colouring agents may be defined as substances employed in pharmacy for the purpose of imparting colour. The use of colourants in pharmaceutical dosage forms produces no direct therapeutic benefit, the psychological effects of colour have long been recognized. The colouring of pharmaceuticals is extremely useful for product identication during manufacturing and distribution. Many patients rely on colour to recognize the prescribed drug and strength. Patient compliance of an unattractive medication can also be improved by the careful selection of colour. Historical accounts describe many colours derived from natural sources that were used for colouring foods, drugs, and cosmetics. Natural colouring principles are obtained from animal, plant, or mineral sources.

Animals have been the source of colouring principle from the earliest period. The dye 6,6'-dibromoindigo (tyrian purple) was prepared by air oxidation of a colourless secretion obtained from the glands of snail, *Murex brandaris*. Cochineal from the insect *Coccus cacti* contains carminic acid, a bright red colouring principle and a derivative of anthraquinone. Salmonella contamination precludes the use of this dye for pharmaceuticals. Many plants contain colouring principles that may be extracted and used as colourants. Flavones such as rutin, riboflavine, hesperidine and quercetin are yellow pigments. Yellow colour is also imparted by natural β carotene obtained from carrots and glycoside such as saffron. Alizarine is a reddish yellow dye obtained from madder plant. Anattenes are yellow to orange water soluble dyes obtained from annatto seeds and indigo plant is a source of blue pigment, indigo. Mineral colours frequently are termed pigments and are used to impart colour to preparations meant for external use. Titanium dioxide, red ferric oxide, yellow ferric oxide and carbon black are few examples.

Although colours from plant, animal, and mineral sources—at one time the only colouring agents available-remained in use early in this century, manufacturers had strong economic incentives to phase them out.

In contrast to natural colouring principles, chemically synthesized colours or synthetic colouring principles are less expensive, easier to produce,

superior in colouring properties and only small amounts are needed. They blend well and dont impart unwanted flavours to foods and drugs. However, use of synthetic colours was at times a threat to health because they were used without discrimination between those that were toxic and those that were safe. Increasing health concern led to early studies and regulations that produced various lists of colours found suitable for addition to foods and drugs. All synthetic colourants approved for use must meet the specications, uses, and restrictions as described in Title 21 of the CFR (Parts 74, 81, and 82). The Federal Food, Drug, and Cosmetic (FD&C) Act of 1938 made certication of the synthetic colourants mandatory.

The FD&C Act divided the synthetic colours into three categories: colours permitted for foods, drugs, and cosmetics (FD&C), colours permitted for drugs and cosmetics (D&C), and colours permitted for externally applied drugs and cosmetics (external D&C). The colours are lettered and numbered that separate their hues. These letter and number combinations make it easy to distinguish colours used in foods, drugs, or cosmetics from dyes made for textiles and other uses. Only FDA certied colouring agents can carry these special designations.

## *Certified Colourants*

## *Soluble Dyes*

Dyes are water-soluble synthetic organic molecules produced from highly puried intermediates derived from petrochemicals. Dyes exhibit their colour by transmitted light. Water-soluble dyes are offered in powder, granular, liquid, dispersion and paste form. The liquid form offers ease of handling, is dust-free, and is ready to use. However, it is probably the most expensive form of colouring. Dyes particularly in liquid form, is subject to light, heat and microbial stability problems. The powder form is least expensive but poses potential dust problems and can result in contamination. The granular dye is recommended because dust problems are reduced, although, in some cases, it may have slower dissolution rates compared to powder form. Different forms of dyes are customized for specific uses and are selected by the user for their particular application. The instability to these parameters and the solubility limits differ from colour to colour and should be considered when selecting dyes for use in various applications.

## *Lake Pigments*

Lakes are insoluble pigments manufactured through the adsorption of a metal salt of a dye on a base of alumina hydrate. Properties that have made lakes more suitable for colouring dosage forms are their relative opacity, stability with regard to light and heat, and ability to be used dry when colouring tablets made by direct compression. They are the only choice for colouring oils, fats and lipid formulations, because there are no oil-soluble colourants approved for this application. Only aluminum lakes are permitted for use in foods, drugs, and cosmetics. Some of the physical characteristics listed for lakes in the literature include shade, particle size, moisture content, bulk density, oil absorption, specific gravity, and pH stability. Out of these properties, the shade and particle size are the two most important attributes. The shade of the lake may be influenced by the quantity of dye adsorbed onto the alumina hydrate and the particle size distribution. The particle size determines the tinting properties (colouring power) of a particular pigment. Smaller particles result in increased surface area, which allows for an increase in reflected light and hence more colour. Tables 12.1 to 12.3 show list of permitted colour additives permanently listed (Table 12.1), provisionally listed (Table 12.2) and exempted from certification (Table 12.3).

**Table 12.1:** Permanently listed colour additives

| Colour | Common name | Colour index number |
|---|---|---|
| FD&C blue #1 | Brilliant blue FCF | 42090 |
| FD&C blue #2 | Indigotine | 73015 |
| D&C blue #4 | Alphazurine FG | 42090 |
| D&C blue #6 | Indigo | 73000 |
| D&C blue #9 | Indanthrene blue | 69825 |
| D&C brown #1 | Resorcin brown | 20170 |
| FD&C green #3 | Fast green FCF | 42053 |
| D&C green #5 | Alizarine cyanide green F | 61570 |
| D&C green #6 | Quinizarine green SS | 61565 |

| | | |
|---|---|---|
| D&C green #8 | Pyranine concentrated | 59040 |
| D&C orange #4 | Orange II | 15510 |
| D&C orange #5 | Dibromofluorescein | 45370:1 |
| D&C orange #10 | Diiodofluorescein | 45425:1 |
| D&C orange #11 | Erythrosine yellowish Na | 45425 |
| FD&C red #3 | Erythrosine | 45430 |
| FD&C red #4 | Ponceau SX | 14700 |
| D&C red #6 | Lithol rubin B | 15850 |
| D&C red #7 | Lithol rubin B Ca | 15850:1 |
| D&C red #17 | Toney red | 26100 |
| D&C red #21 | Tetrabromo fluorescein | 45380:2 |
| D&C red #22 | Eosine | 45380 |
| D&C red #27 | Tetrachlorotetra-bromofluorescein | 45410:1 |
| D&C red #28 | Phloxine B | 45410 |
| D&C red #30 | Helidone pink CN | 73360 |
| D&C red #31 | Brilliant lake red R | 15800:1 |
| D&C red #33 | Acid fuchsine | 17200 |
| D&C red #34 | Lake Bordeaux B | 15800:1 |
| D&C red #36 | Flaming red | 12085 |
| D&C red #39 | Alba red | 13058 |
| FD&C red #40 | Allura red AC | 16035 |
| FD&C red #40 lake | Allura red AC | 16035:1 |
| D&C violet #2 | Alizurol purple SS | 60725 |
| Ext.D&C red #2 | Alizarine violet | 60730 |
| FD&C yellow #5 | Tartrazine | 19140 |
| FD&C yellow #6 | Sunset yellow FCF | 15983 |
| D&C yellow #7 | Fluorescein | 45350:1 |
| Ext.D&C yellow #7 | Napthol yellow S | 10316 |
| D&C yellow #8 | Uranine | 45350 |

| | | |
|---|---|---|
| D&C yellow #10 | Quinoline yellow WS | 47005 |
| D&C yellow #11 | Quinoline yellow SS | 47000 |

**Table 12.2:** Provisionally listed colour additives

| Colour | Common name | Colour index number |
|---|---|---|
| FD&C blue #1 lake | Brilliant blue FCF | 42090:2 |
| FD&C blue #2 lake | Indigotine | 73015:1 |
| D&C blue #4 lake | Alphazurine FG | 42090 |
| FD&C green #3 lake | Fast green FCF | 42053 |
| D&C green #5 lake | Alizarine cyanide green F | 61575 |
| D&C green #6 lake | Quinizarine green SS | 61565 |
| D&C orange #4 lake | Orange II | 15510:2 |
| D&C orange #5 lake | Dibromofluorescein | 45370:2 |
| D&C orange #10 lake | Diiodofluorescein | 45425:2 |
| D&C orange #11 lake | Erythrosine | 14700 |
| FD&C red #4 lake | Ponceau SX | 15850:2 |
| D&C red #6 lake | Lithol rubin B | 15850:1 |
| D&C red #7 lake | Lithol rubin B Ca | 26100 |
| D&C red #17 lake | Toney red | 45380:3 |
| D&C red #21 lake | Tetrabromo fluorescein | 45380:3 |
| D&C red #22 lake | Eosine | 45410:2 |
| D&C red #27 lake | Tetrachlorotetra-bromofluorescein | 73360 |
| D&C red #28 lake | Phloxine B | 15800:1 |
| D&C red #30 lake | Helidone pink CN | 17200 |
| D&C red #31 lake | Brilliant lake red R | 15880:1 |
| D&C red #33 lake | Acid fuchsine | 12085 |

| | | |
|---|---|---|
| D&C red #34 lake | Lake Bordeaux B | 60725 |
| D&C red #36 | Flaming red | 19140:1 |
| D&C violet #2 lake | Alizurol purple SS | 15985:1 |
| FD&C yellow #5 lake | Tartrazine | 45350:1 |
| FD&C yellow #6 lake | Sunset yellow FCF | 10316 |
| D&C yellow #7 lake | Fluorescein | 15350:1 |
| Ext.D&C yellow #7 lake | Napthol yellow S | 10316 |
| D&C yellow #8 lake | Uranine | 45350 |
| D&C yellow #10 lake | Quinoline yellow WS | 47005:1 |

**Table 12.3:** Colour additives exempt from certification and permitted for use

| **Colour name** | **Colour index number** |
|---|---|
| Alumina | 77002 |
| Aluminium powder | 77000 |
| Annatto extract | 75120 |
| β APO-8-carotinol | 40820 |
| β carotene | 40800 |
| Bismuth oxychloride | 77163 |
| Bronze powder | 77440 |
| Calcium carbonate | 77220 |
| Canthaxanthin | 40850 |
| Carbazole violet | 51319 |
| Carmine | 75470 |
| Chlorophyllin copper complex | 75810 |
| Chromium-cobalt-aluminium oxide | 77343 |
| Chromium hydroxide green | 77289 |

| | |
|---|---|
| Chromium oxide green | 77288 |
| C.I. vat orange 1 | 59105 |
| Cochineal extract | 75470 |
| Copper powder | 77400 |
| 7,16-dichloro-6,15-dihydro-5,9,14,18-anthrazinetetrone | 69825 |
| Ferric ammonium ferrocyanide | 77510 |
| Ferric ferrocyanide | 77510 |
| Guanine | 75170 |
| Henna | 75480 |
| Iron oxides, synthetic | 77491(red), 77492 (yellow), 77499 (black) |
| Log wood extract | 75290 |
| Manganese violet | 77742 |
| Mica | 77019 |
| Potassium sodium copper chlorophyllin | 75180 |
| Phthalocyamine green | 74260 |
| Pyrogallol | 76515 |
| Pyrophyllite | 77004 |
| Saffron | 75100 |
| Silver | 77820 |
| Tagetes meal and extract | 75125 |
| Talc | 77019 |
| Titanium dioxide | 77891 |
| Turmeric | 75300 |
| Turmeric oleoresin | 75300 |
| Ultramarine blue | 77007 |
| Ultramarine green | 77013 |
| Ultramarine Pink | 77007 |
| Ultramarine Red | 77007 |

| | |
|---|---|
| Ultramarine violet | 77007 |
| Zinc oxide | 77947 |

## Flavours

Flavouring agents have been used to flavour foods and to make “medicines” palatable. The flavour of the drug is a sensation with multidimensional components involving subjective and objective perceptions of taste, feeling factors and odor. Taste is more narrowly defined, referring only to those sensations perceived through the stimulation of receptor cells enclosed within the taste buds on the tongue, namely, sweet, sour, bitter, salty and umami. Feeling factors, which include numbing, cooling and burning sensations, occurs when the free nerve endings in the mucus membranes are stimulated, in turn exciting the trigeminal nerve. In this way, feeling factors are more akin to a chemical irritation or tactile response. These mouth feel factors are important in improving the organoleptic qualities of pharmaceuticals. Odor component of flavour is perceived by the sense of the smell. The aroma of a product is caused by the perception of volatile components that stimulate the olfactory region of the nose. Aroma can reach the olfactory region of the nose through multiple pathways including orthonasal perception (sniffing) and retronasal perception (swallowing). It is not physiologically possible for aroma of a drug product to affects its bitterness since it is perceived by a completely different mechanism than its taste. However, aroma is important in palatable formulations, because it is the first attribute perceived. A well constructed flavour system is crucial to oral drug formulations because it influences patient acceptability and compliance. Although sweeteners, flavour potentiators and taste extenders are needed in a proper formulation, a flavoured product requires a identifying flavour. While only one piece of a flavouring system, the identifying flavour may be the one thing a patient most readily recognizes about the drug product.

Flavouring agents may be classied as natural, articial, or natural and articial (N&A) flavours. Just as a drug product is not only composed of the APIs, flavours are not only blends of aromatic molecules. Functional ingredients support the performance of the aromatic molecules. Aroma molecules typically only make up 10 % of the weight of the flavour. The largest component of the remaining 90% is called as carrier. The choice of carrier determines the physical form of the flavour-dry (e.g. crystalline vanillin, cinnamon powders, and dried lemon fluid extract) or liquid (e.g. essential oils, fluid extracts, tinctures, and distillates).

For liquid flavours, the carriers used are usually propylene glycol or ethyl alcohols. Ideally, these carriers serve as solvents that completely dissolve the aroma molecules. Less ideally, an emulsion may be required to deliver the aroma molecules. Emulsions may require that the bulk flavour be agitated before it is added to the final product. Liquid flavours are by far the most widely used because they diffuse readily into the substrate.

The manufacture of dry flavour begins with liquid flavours that are subsequently plated onto a dry carrier such as maltodextrins, modified starch, dextrose or gum arabic. Due to the volatile nature of flavour molecules, these processes contribute to evaporation of the aroma chemicals. Loss of volatiles is the primary reason liquid flavours are generally of higher flavour quality than dry flavours. The carrier also assures the uniform dispersion of flavours throughout the product, regulate the strength of flavour and inhibit chemical reaction from occurring.

## *Natural Flavours*

Natural compounds are the first source for flavour. In a natural flavour, the flavourant must be extracted from the source substance. The methods of extraction can involve extrusion, solvent extraction or distillation. Modern use of natural flavours in pharmaceuticals is limited, because they are often unstable and their quality is unpredictable from season to season. The most commonly used natural flavours are terpeneless citrus oils, which are stable if well protected from light and air. A variety of other natural flavours are used in the food and pharmaceutical industries; some of the more common flavours are described below:

## *Anise (Pimpinella anisum, Umbelliferae)*

The anise oil is obtained by steam distillation of dried fruits (seeds). The oil is a clear-to-pale yellowish oil which solidies at low temperatures and has a characteristic sweet licoricelike odor and flavour. Anise oil is used frequently at concentrations of up to approximately 3000 ppm in liquid preparations. Its main constituents include anethol (90%), methylchavicol, p-methoxyphenylacetone, and acetic aldehyde.

## *Cardamon (Elettaria cardamomum, Zingiberaceae)*

The essential oil is a greenish-yellow liquid with a warm, spicy, aromatic

odor and flavour, obtained by steam distillation of seeds. The main constituents of the oil are limonene, cineol, D-α-terpineol, and terpinyl acetate. Oil is generally used at concentrations of approxi-mately 5–50 ppm.

## *Wild Cherry (Prunus serotina, Rosaceae)*

Fluid extract and tincture are prepared by extraction of bark, small branches, and twigs.

It has a characteristic sweet, tart, cherry-like flavour. Glucoside prunasin, the main constituent of wild cherry extract, on enzymatic hydrolysis yields prussic acid, glucose, and benzaldehyde. Also present are coumarin, phytosterols, benzoic acid, and fatty acids (e.g. oleic, linoleic, and palmitic acids). Wild cherry bark extract is commonly used at concentrations of 50–800 ppm in foods and pharmaceuticals.

## *Lemon (Citrus limonum, Rutaceae)*

The essential oil of lemon is obtained by cold expression of fruits, and rind. The main constituents of the oil are limonene. Fluid extracts and tinctures are obtained from the dried peel. Lemon petitgrain is obtained by steam distillation of the leaves. For flavouring, it must be terpeneless. The main constituents are *d*-α-pinene, camphene, *d*-limonene, dipentene, nerol, Z-linalol, and citral. All lemon oil derivatives have the characteristic lemon odor and a slightly bitter flavour. Lemon petitgrain oil is used in a wide variety of applications in concentrations ranging from 1 to 35 ppm. The essential oil and extract of lemon are generally used at higher concentrations that may range up to 1000 and 10,000 ppm, respectively.

## *Orange, Bitter (Citrus aurantium, Rutaceae)*

The leaves and twigs produce essential petitgrain oil following steam distillation. Neroli bigarade essential oil is produced from the blossoms by steam distillation. The peel is expressed and steam distilled to produce essential oil of orange. The main constituent of orange oil is d-limonene. Essential oil of orange is widely used in pharmaceuticals at concentrations of up to 500 ppm.

## *Orange, Sweet (Citrus sinensis, var. aurantium dulcis, Rutaceae)*

Essential oil of sweet orange is obtained by expression. Its physical-chemical

properties vary according to origin. It has a characteristic odor and a mildly bitter, astringent flavour; it is generally used at concentrations similar to that of bitter orange. The oil contains more than 90% limonene, in addition to relatively high quantities of decylic, octylic, nonylic, and dodecylic aldehydes, and citral esters.

## *Peppermint (Mentha piperita, Labiatae)*

The peppermint oil is obtained by steam distillation of the flowering plant tops. It has a strong mint odor with a sweet balsam taste masked by a strong cooling effect. It is widely used in foods, as well as in liquid pharmaceuticals, to 8000 ppm. The main constituents of the essential oil are $\alpha$-and $\beta$-pinene, cineol, limonene, ethyl amylcarbinol, menthol, isomenthol, menthone, menthyl acetate, and piperitone.

## *Artificial Flavours*

Artificial flavours are mixtures of individual synthesized aroma chemicals that may be identical to natural flavours. The development of artificial flavours paralleled the development of instrumental analysis, in which active ingredients in natural flavours are identified and reconstructed synthetically with reasonable accuracy. Exact duplication of a natural flavour is, however, difficult because often minor components, the most important contributors to the overall flavour, are not easily identified. These flavours are combined together in ratios that successfully emulate a natural aroma. Unlike natural flavouring agents, artificial flavours are usually stable, have greater consistency and decreased impact from raw material changes.

## *Natural and Artificial Flavours*

In N&A flavour systems, natural flavours are combined with artificial flavours to enhance flavour balance and fullness. These flavours are generally classified according to type and taste sensation. Table 12.4 contains a list of N&A flavour components that elicit various sensory properties, all of which are commonly used in food and drug compounding. Many N&A flavours may be chemically and structurally similar, but vary significantly in taste and aroma. A relatively small change in chain length can have a profound impact on flavour type. Minor changes, such as the conversion of allyl benzoate to valerate transform a basic cherry flavour to pineapple. In situ conversion of essential semisynthetic flavour components from one molecular form to

another, as a result of ion pairing, is common in food and drug products. Therefore, the inadvertent conversion of flavours between types during drug formulation studies (e.g. effect of pH, salts, and temperature) can present a serious challenge in flavour-quality assessment. The fact that one and the same N&A flavour component can deliver several flavour and odor impressions imply that a blend of several flavour compounds would be preferable. Another advantage of N&A flavours is the broad spectrum of flavouring agents from which the formulator can develop an entirely new flavour system as exemplified by tutti-frutti flavour (Table 12.5).

**Table 12.4:** Taste and flavour characteristics of N&A flavour ingredients

| Ingredients | Primary taste | | | Flavour |
|---|---|---|---|---|
| | Bitter | Bitter-sweet | Sweet | |
| Allyl benzoate | | √ | | Cherry |
| Allyl caproate | | √ | | Pineapple |
| Allyl cyclohexylbutyrate | | √ | | Pineapple |
| Allyl cyclohexylcaproate | | √ | | Peach/Apricot |
| Allyl cyclohexylvalerate | | √ | | Apple |
| Allyl phenoxyacetate | | √ | | Honey/Pineapple |
| Anethol | | | √ | Anise |
| Anisyle alcohol | | | √ | Peach |
| Anisyle formate | | | √ | Strawberry |
| Benzyle isobutyrate | | | √ | Strawberry |
| Benzyle salicylate | | | √ | Raspberry |
| Cinnamaldehyde | | √ | | Cinnamon/Melon |
| Cinnamyl anthranilate | | √ | | Grape |
| Citral | | √ | | Lemon |
| Citronellyl formate | | √ | | Plum |
| Decyl formate | | | √ | Grape |
| Diacetyl | | | √ | Butter |
| Diphenyle ether | | | √ | Black currant |
| Ethyl valerate | | √ | | Banana/Apple |
| Eugenol | √ | | | Clove buds |
| Geraniol | √ | | | Rose like |
| α ionone | | | √ | Raspberry |
| Isobutyl salisylate | | √ | | Strawberry |
| Isobutyle anthranilate | | √ | | orange/Strawberry |
| Isopropyl valerate | | √ | | Apple |
| Linalyl anthranilate | | | √ | Orange |
| Methyl ionone | | | √ | Raspberry/Currant |
| Methyl propionate | | | √ | Black currant |
| Methyl undecyl ketone | | √ | | Coconut |
| Musk ambrette | | | √ | Peach |
| Nerol | √ | | | Rose like |
| Neryl acetate | | | √ | Raspberry |
| Neryl butyrate | | | √ | Cocoa |
| Propenyl guaethol | | | √ | Vanilla |
| Propyl isobutyrate | | | √ | Pineapple |
| Rhodinol | √ | | | Rose |
| Santalyl acetate | | √ | | Apricot |
| Terpenyl butyrate | | √ | | Plum |
| Tetrahydrofurfuryl proprionate | | √ | | Chocolate/Apricot |
| Vanillylidene acetone | | √ | | Vanilla |
| Yara yara | | | √ | Strawberry |

| **Table 12.5:** Formulation and composition of tutti-frutti flavour | |
|---|---|
| **Ingredient** | **Parts by weight** |
| Amyl acetate | 300.0 |
| Amyl butyrate | 48.0 |
| Ethyl butyrate | 36.0 |
| A-ionone | 120.0 |
| Jasmine absolute (10% in alcohol) | 0.1 |
| Lemon essential oil | 1.0 |
| Orris resinoid | 80.0 |
| Imitation rose (10% in alcohol) | 28.0 |
| Rum ether | 100.0 |
| γ-Undecalactone | 18.0 |
| Vanillin | 11.0 |
| Alcohol (solvent) | 257.0 |
| **Total** | **1,000.0** |

## *Flavour Enhancers and Potentiators*

Sugar, carboxylic acids (e.g. citric, malic, and tartaric), common salt (NaCl), amino acids, amino acid derivatives (e.g. mono-sodium glutamate-MSG), and spices (e.g. peppers) are most often employed to enhance flavour. Citric acid is most frequently used to enhance taste performance of both liquid and solid pharmaceutical products. Other acidic agents, such as maleic and tartaric acids, are also used. Although extremely effective with proteins and vegetables, MSG has limited use in pharmaceuticals because flushing, headache, and chest pain have been ascribed to its presence, albeit after food intake rather than medication. This is the background to the so-called Chinese restaurant syndrome. Monoammonium glycyrrhizinate has a lingering sweet aftertaste, which can be exploited for taste-masking products with a mildly bitter aftertaste. It is also effective in enhancing chocolate flavour.

## *Selection of Flavour*

The number of known flavouring agents includes thousands of molecular

compounds. There are three major classes of compounds in an individual flavour:

## *Impact Compounds*

These compounds represent most of the organoleptic potency of the flavour type. When smelled alone, these compounds are reminiscent of the named flavour. They are necessary, characteristic and essential for the flavour. Methyl anthranilate is the impact compound of Concord grape.

## *Contributory Compounds*

They enhance the complexity and identity of the named flavour. Contributory compounds in conjugation with impact compound, bring aroma closer to that of named flavour.

## *Differential Compounds*

They impart a distinctive characteristic to an individual flavour.

The compound employed as flavours vary considerably in their chemical structure, ranging from simple esters (methyl salicylate), alcohols (glycerine), and aldehyde (vanillin) to carbohydrates (honey) and the complex volatile oil (lemon oil). There is a close relation between chemical structure and taste. Sour taste is caused by hydrogen ions and is proportional to hydrogen ion concentration and lipid solubility of the compound. It is characteristic of acids, phenols, alums and tannins. Saltiness is due to presence of anions and cations. Bitter taste is due to high molecular weight salts. Polyhydroxy, poly-halogenated aliphatic compounds and amino acids produce sweet taste. Typically, for a sweet and fruity flavour, aroma molecules have ester bonds in them. Sweetness increases with number of hydroxyl groups, possibly because of increased solubility. Meaty, earthy and savory aromas consist of thiol groups. Presence of unsaturation bestows a sharp, biting odor and taste on compounds.

The difficulty involved in formulating a palatable drug product typically stems from the taste of the active pharmaceutical ingredient. The key to formulating a palatable drug product is to build a complex flavour system that covers the negative sensory effects of the API as well as of excipients. Different flavour concentrations produce highly subjective sensations. Specic requirements for balance and fullness are dependent, in part, on the drug

substance and the physical form of the product. For this reason, when selecting a flavour system, the compounding pharmacist must take into account several variables upon which a desired response would depend. Some of these are product texture (e.g. viscosity of formulation, solid or liquid), water content, base vehicle or substrate, and taste of the subject drug. Flavours, in compounding prescriptions, can be divided into four categories according to the type of taste that is to be masked, as follows:

***Salty taste*** **–** for salty drugs such as sodium salicylate, ammonium chloride and ferric ammonium citrate, cinnamon syrup has been found to be the best vehicle. Orange syrup, citric acid syrup, cherry syrup, cocoa syrup, wild cherry syrup, raspberry syrup and glycyrrhiza syrup are useful vehicles in descending order.

***Bitter taste*** **–** bitter taste of anti-malarials could be efficiently masked by cocoa syrup. Followed in descending order of usefulness are raspberry syrup, cherry syrup, cinnamon syrup, citric acid syrup, licorice syrup, orange syrup and wild cherry syrup.

***Sour or acrid taste*** **–** raspberry syrup and other fruit syrups are efficient in masking taste of sour substances such as hydrochloric acid. Acacia syrup and other mucilaginous substances are best for disguising acrid taste like that of capsicum.

***Oily taste*** **–** cod liver oil is made palatable by adding wintergreen oil or peppermint oil. Lemon, orange and anise or their combinations are also useful. Castor oil is made palatable by emulsifying with an equal volume of aromatic rhubarb syrup.

Most importantly, selection of an appropriate flavour requires knowledge of the physical and chemical properties of the product, input from marketing and a system for appropriate evaluation.

## Sweeteners

Sweeteners are indispensable components of many liquid oral dosage forms and chewable medications, especially those containing bitter or other unacceptable tastes. Sweeteners are often classied as either caloric or non-caloric. Non-caloric sweetening agents are preferred for diabetic patients, as ingestion does cause increases in systemic glucose concentrations. The commonly used sweeteners in the pharmaceutical industry are sucrose (sugar), glucose, fructose, sorbitol, and glycerin. Sucrose, glycerin, glucose and sorbitol have limited use in solid dosage forms (e.g. tablets) because the materials are hygroscopic. Mannitol is used more often in tablet manufacture. Besides being less hygroscopic, it has a negative heat of solution. For this reason, chewable tablets containing mannitol have a pleasant cooling sweet taste, which complements flavour quality. Liquid glucose is an extremely viscid substance that imparts both body and sweetness to liquid formulations. It is obtained by the incomplete hydrolysis of starch and consists chiey of dextrose, dextrins, maltose, and water. It imparts a characteristic odor and flavour to the formulation in similar fashion to honey and molasses, but to a lesser degree. The polyols are sometimes called sugar replacers, sugar-free sweeteners, sugar alcohols, or novel sugars. Polyols occur naturally in plants but can also be produced commercially. They include such compounds as sorbitol, mannitol, xylitol, and hydrogenated starch hydrolysates.

Saccharin (Sweet'N Low) is a non-nutritive artificial sweetening agent widely used in foods and pharmaceuticals. Saccharin is a sucrose substitute for diabetics, the obese, and others who do not wish to ingest sucrose. It has approximately 500 times the sweetening power of sucrose, depending in extent on the strength of the solution. It is sweet at very low concentrations (equivalent to about 5–10% sugar) but bitter at higher concentrations. Sodium salt form of saccharin is more palatable and comparatively free of unpleasant aftertaste. Saccharine is suspected as being a carcinogen due to cyclohexylamine formation, possibly by gut ora, on ingestion. It was banned as a food additive by the FDA in 1977, but has remained available consequent to regular congressional moratoria on the proposed ban.

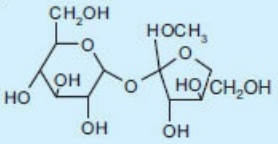
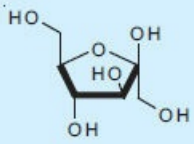
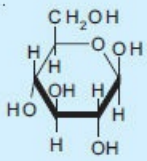
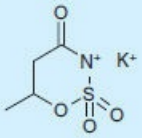
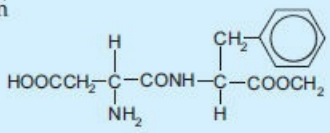
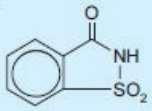
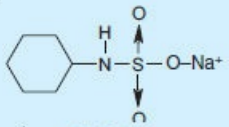
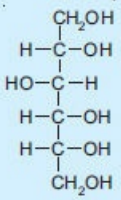
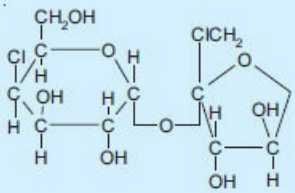
Sodium cyclamate is another artificial sweetening agent that is approximately 30 times as sweet as sugar. However, its use as an articial sweetener is banned in the USA. because of the carcinogenic concerns of its

metabolite cyclohexylamine.

Aspartame (NutraSweet, Equal), N-L-a-Aspartyl-L-Phenylalanine methyl ester, is 200 times sweeter than sucrose and, unlike saccharin, has no aftertaste. It hydrolyzes to aspartylphenylalanine and diketopiperazine, with a loss in sweetness by aspartame synergistic with saccharin, sucrose, glucose, and cyclamate. It is contraindicated in patients suffering from phenylketone urea, as hydrolysis can lead to formation of phenylalanine. Aspartame may also cause angiodema and urticaria.

Sucralose (Splenda) is a new, non-caloric sweetening agent, approximately 600 times sweeter than sucrose and differs from sucrose by the substitution of three chlorines for hydroxyl groups. Sucralose is heat-stable and stable over a wide pH range affording its utility in formulations prepared at high temperatures. Acesulphame-K (Sunett, Sweet One) is approximately 200 times sweeter than sucrose and is commonly used concomitantly with aspartame to synergistically enhance overall sweetening. This sweetener is also heat stable. Furthermore, Monoammonium glycyrrhizinate has even been used in liquid oral preparations. Table 12.6 lists some commonly used sweetening agents along with their structure and relative sweetening intensity.

**Table 12.6:** List of commonly used sweetening agents

| Sweetener | Brand name | Sweetening intensity | Comments | Chemical structure |
|---|---|---|---|---|
| **Caloric** | | | | |
| Sucrose | | 1 | Commonly used, known as table sugar |  |
| Fructose | | ~1.2 | Rapid onset of apparent sweetness |  |
| D-Glucose | | ~0.5-0.9 | Commonly referred to as dextrose |  |
| **Non-caloric** | | | | |
| Acesulfame-k | Sunnet®, Sweet One® | ~200 | Heat stable |  |
| Aspartame | NutraSweet®, Equal® | ~200 | Unstable in solution |  |
| Saccharine | Sweet'N Low® | ~500 | Unpleasant aftertaste |  |
| Sodium cyclamate | | ~30 | Carcinogenic concern |  |
| Sorbitol | | ~0.6 | Pleasant taste |  |
| Sucralose | Splenda® | ~600 | Heat stable and stable over a broad pH range |  |

# STABILIZERS

## Preservatives

The preservatives are added to prevent contamination, deterioration, and spoilage by bacteria and fungi, since many of the components in the pharmaceutical preparations serve as substrates for these microorganisms. Several terms are used to describe microbial organisms associated with pharmaceutical and cosmetic products: “harmful,” “objectionable,” and “opportunistic.”

The USP XX uses the term “harmful” to refer to microbial organisms or their toxins that are responsible for human disease or infection. Examples of organisms that must not be present in a product are given, namely, *Salmonella* species, *Escherichia coli*, certain species of *Pseudomonas*, including *P. aeruginosa*, and *Staphylococcus aureus*. An “objectionable” organism can cause disease, or its presence may interrupt the function of the drug or lead to the deterioration of the product. Organisms are defined as “opportunistic” pathogens if they produce disease or infection under special environmental situations, as in the newborn or the debilitated person.

The following objectionable organisms should not be present in a pharmaceutical or cosmetic product: *P. putida, P. multivorans, P. maltophilia, Proteus mirabilis, Serratia marcescens, Klebsiella* sp., *Acinetobacter anitratus (Bacterium anitratum), and Candida* sp.

Preservatives are not added to enhance stability to the formulation, but rather to give exibility in the use of the drug product. With single-use formulations (those lacking a preservative), there is a requirement that each drug container be entered only once. Exposure of the drug product in the container to air may result in the introduction of bacteria to the product. This is not a problem if the drug is used immediately because there is not suffcient time for the bacteria to colonize. However, on storage, it is possible for a signicant number of bacteria to grow. For this reason, a multidose drug product must contain a preservative to prevent bacterial growth. A list of preservatives that have been used in pharmaceutical formulations is shown in Table 12.7. The success or failure of a preservative in protecting a formulation against microbial spoilage depends upon many factors. The interaction of the preservative with surfactants, active substances, other

components of the vehicle, sorption by polymeric packaging materials, and product storage temperature may change the concentration of the unbound or free preservative in the aqueous phase.

Perfumes, high concentrations of glycerine, and electrolytes make the environment less favorable to microbial growth, thus enhancing the effectiveness of the preservatives. Preservative action appears to depend on the concentration of the free preservative in the aqueous phase. Surfactant solubilized preservative may be bound within the micelles and there inactivated, or on the contrary, the micelles may act as reservoirs of preservative in an actively preserved system.

Preservatives are typically small hydrophobic compounds that may interact with hydrophobic regions of the protein, leading to a disruption in protein structure. Acidic preservatives, such as the p-Hydroxybenzoic acid esters and salts of benzoic acid, are the most widely used for oral preparations. These are adequately soluble in aqueous systems and possess both antifungal and antibacterial properties.

The p-hydroxybenzoate esters are used in combination with one another because of their synergistic action. In general, they are employed at a concentration level approaching their maximum solubility in water. The propyl or butyl ester is usually dissolved in the fat phase and should be increased for vehicles with a high fat content. Satisfactory protection of the emulsion against microbial growth may possibly be attained with sorbic acid, in which the p-hydroxybenzoate esters prove to be ineffective. The paraben esters of p-hydroxybenzoic acid are popular as preservatives because their toxicity is low, they are odorless, they do not discolour, and they are nonirritating to the skin. On the negative side, the parabens have a low solubility in water and are less effective against gram-negative bacteria than molds and yeasts. Combining the parabens with phenoxyedianol, or with imidazolidinyl urea (Germall II), improves their activity against bacteria, yeast, and molds. The supplier claims that the combination system retains activity against yeast and mold even when paraben activity has been diminished by interaction with nonionics or other substances in the formulation, or has migrated into the oil phase. Germall II is used in concentrations of 0.1 to 0.5% alone or in combination with the parabens. It should be added to the product below 60°C. The solid parabens may be difficult to incorporate into some formulations because of their low water

solubility. A 50% by weight oil-in-water emulsion (Liqua Par) has been marketed. The oil phase is a mixture of p-hydroxybenzoic acid esters: n-butyl, isobutyl and isopropyl. The aqueous portion contains water with emulsion stabilizers. The solubility of the active ingredients in water at 25°C is 0.06 g/100 g and is freely irascible with propylene glycol. The preservative should be added to the aqueous phase at a temperature not exceeding 70 to 75°C and stirred until thoroughly dissolved before the preparation of the emulsion. Paraben hydrolysis may occur if the temperatures exceed 80°C.

**Table 12.7:** Preservatives used in pharmaceutical systems

| Preservative | Effective concentration (%) | Administration route | Incompatibility |
|---|---|---|---|
| Benzalkonium chloride | 0.02 | IM, inhalation, nasal, ophthalmic, otic, topical | Citrate, methylcellulose |
| Benzithonium chloride | 0.01 | IM, IV, ophthalmic, otic | Anionic surfactant |
| Benzoic acid | 0.1–0.3 | IM, IV, irrigation, oral rectal, topical, vaginal | Alkalis, heavy metal, kaolin |
| Benzyl alcohol | 0.75–10 | Injections, oral, topical, vaginal | Oxidizing agents, nonionic surfactant, methyl cellulose |
| Bronopol | – | Topical | Sulfhydryl compounds, aluminium |
| Butylparaben | 0.001–0.2 | Injection, oral, rectal, topical | Nonionic surfactant, some plastics |
| Cetrimide | 0.01–0.02 | Topical, ophthalmic | Ionic and nonionic surfactant, metals |
| Chlorohexidine | – | Topical, ophthalmic | Anionic material, $Ca^{2+}$, $Mg^{2+}$, viscous material |
| Chlorobutanol | 0.25–0.5 | IM, SC, IV, inhalation, nasal, otic, topical, ophthalmic | Some plastic, rubber, carboxymethylcellulose, polysorbate 80 |
| Chlorocresol | 0.05–0.1 | Topical | Some plastic, rubber, methylcellulose, nonionic surfactant |
| Cresol | 0.1–0.35 | Topical, intradermal, IM, SC | Nonionic surfactant |
| Ethylparaben | 0.001–0.2 | Topical, oral | Nonionic surfactant, some plastic, silica |
| Imidurea | | Topical | – |
| Methylparaben | 0.05–0.18 | IM, IV, SC, ophthalmic, oral, otic, rectal, topical, vaginal | Nonionic surfactant, sorbitol, some plastic |
| Phenol | 0.2–0.5 | Injections | Albumin, gelatin |
| Phenoxyethanol | 0.50 | Topical | Nonionic surfactant, PVC, cellulose derivative |
| Phenylethyl alcohol | 0.2–1.0 | Nasal, ophthalmic,otic | Oxidizing agent, proteins, polysorbates |
| Phenyl mercuric acetate/borate | 0.001 | Ophthalmic | Halides, amino acids, some plastic, rubber |
| Phenyl mercuric nitrate | 0.001 | Topical, ophthalmic, IM | Halides, amino acids, some plastic, rubber, aluminium |
| Propylparaben | 0.01–0.1 | IM, IV, SC, inhalation, topical, ophthalmic, oral, otic, rectal, vaginal | Nonionic surfactant, some plastic |
| Sodium benzoate | 0.1–0.3 | Dermal, IM, IV, oral, rectal, topical | Quaternary compounds, gelatin, calcium salts, nonionic surfactant |
| Thimerosal | 0.003–0.012 | IM, IV, SC, ophthalmic, otic, topical | Sodium chloride solution, proteins, some plastic, rubber |

The other three classes of preservatives have been widely used in ophthalmic, nasal, and parenteral products, but not frequently in oral liquid preparations. The neutral preservatives are volatile alcohols; their volatility introduces problems of odor and loss of preservative on aging in multidose preparations. The mercurials and quaternary ammonium compounds are excellent preservatives but are subject to incompatibilities.

Mercurials are readily reduced to free mercury, and the quaternary compounds are inactivated by anionic substances. Thimerosal is also common, especially in vaccines, even though some individuals are sensitive to mercurics.

Another preservative that is available is Dowicil 200, which is described as a broad-spectrum antimicrobial effective against bacteria, yeast, and molds at concentrations of 0.02 to 0.3% weight. It is not inactivated by nonionic, anionic, or cationic formulation ingredients. The substance is extremely soluble in water but is virtually insoluble in oils and organic solvents. Chemically, it is the cis isomer l-(3-chloroallyl)-3,5,7-triaza-l-azoniaadamantane chloride. The preservative should not be heated above 50°C and is unstable in solution below pH 4 and above pH 10. Discolouring of this material may occur, but can be prevented by the addition of sodium sulfite. Strong oxidizing or reducing agents should be avoided since these may adversely affect the antimicrobial efficacy.

Newer preservatives are being marketed, but all of these substances must be thoroughly evaluated for their effectiveness in the product, and their effect on the physicochemical stability of the product.

## *Preservative Efficacy*

In screening preservatives for use in formulations, it is necessary to determine what levels of preservatives are efcacious at preventing bacterial growth in the particular formulation. The minimum inhibitory concentration of preservative necessary to prevent microbial spoilage may be estimated by (1) the use of experimentally determined physicochemical parameters such as the oil/water partition coefficient, concentration of surfactant, the number of independent binding sites on the surfactant, oil/water phase ratio, and concentration of free pre-servative in the aqueous phase, (2) an ultracentrifuge technique and (3) direct dialysis. These techniques provide an

approximate value for the minimum preservative concentration required for a formulation, but to ensure quality, the product must be tested for its ability to withstand accidental and deliberate microbial contamination.

Preservative efficacy in a formulation is determined by the addition of pure or mixed cultures of microbial organisms to the finished preparation. The number of microorganisms initially present in the inoculated material is determined by plating aliquots of suitable dilutions. Table 12.8 gives the USP XX procedure and the investigational FDA procedure for topicals, including the organisms used, the levels of inoculum, sampling periods, and the measure of effectiveness. Various neutralizes for the preservative are added to the culture media to recover a maximum number of organisms. A TAT broth consisting of tryptone (2%), azolectin (0.5%), and polysorbate 20 (4%) has been found to be a suitable medium for topical products. Azolectin is a neutralizing agent for quaternary ammonium compounds and polysorbate 20 inactivates parabens. The samples should be tested at intervals for both slow-growing and rapidly proliferating organisms.

**Table 12.8:** Preservative efficacy (high-level inocula challenge) tests

A. USP XX procedure

1. Organisms used: *C. albicans, A. niger, E. coli, S. aureus, P. aeruginosa*.
2. Inoculum: 0.1ml/20 ml; 100,000 to 1,000,000 cells/ml.
3. Sampling at 7, 14, 21, and 28 days following inoculation
4. Effectiveness: Vegetative cells not more than of 0.1% of initial concentrations by 14 th day; concentration of viable yeasts and molds at or below initial concentration after 14 days; concentration of each test organism remains at or below these levels after 28 days.

B. Investigational FDA procedure for topicals

1. Organisms used: Same as USP XX plus *P. putida, P. multivorans, Klebsiella* sp., *S. marcescens*.

2. Inoculum: 0.2 ml/20 ml; 0.8–1.2 × $10^6$ cells ml.
3. Sampling: Weekly observations.
4. Effectiveness: Vegetative cells <0.01% survival by 28 days; *C. albicans* <1% survival; *A. niger* <10% survival.
5. Re-inoculate: Vegetative cells: 1–2 × $10^5$ cells/ml; 0.1% survival in 28 days.

The USP XX has procedures for determining the microbial content of raw materials and finished products. Suitable limits on the number and types of microorganisms have not been officially specified, however. All materials must be free of the harmful microorganisms listed in the USP XX. Manufacturers have set up their own microbiologic specifications suitable to their raw materials and finished products. A typical manufacturer's microbiologic specification may read as follows: (1) The material must be free of viable organisms restricted by the USP XX. (2) The total aerobic count must not be more than 5000 microorganisms per gram; (3) not more than 100 molds per gram; (4) not more than (nmt)100 yeasts per gram; and (5) not more than 90 coli-forms per gram.

Microbiologic quality guidelines have been established by The Cosmetic, Toiletry and Fragrance Association, Inc. These have been grouped according to product type:

1. Baby products-nmt 500 microorganisms per gram or milliliter.
2. Products used about the eye-nmt 500 microorganisms per gram or milliliter.
3. Oral products-nmt 1000 microorganisms per gram or milliliter.
4. All other products-nmt 1000 microorganisms per gram or milliliter.

The use of multiple preservatives in a formulation can be extremely advantageous. Individual preservatives are able to interact synergistically or antagonistically with one another. Synergy in combination preservative systems is thought to occur by increasing accessibility of a preservative to the target, inhibition of preservative inactivation or by acting at different molecular targets. The advantages of achieving synergy in the selection of a

preservative system can include increased antimicrobial activity over a broader range of microorganisms, enhanced efcacy in microbial killing concurrent with lower preservative concentrations (can mitigate potential toxicity of single preservatives in higher concentrations), and a longer duration of antimicrobial action.

## Antioxidants

Oxidation is defined as a loss of electrons from a compound that results in a change in the oxidation state of the molecule. Such reactions are mediated by free radicals or molecular oxygen, and are often catalyzed by metal ions. Drugs possessing favorable oxidation potential are especially vulnerable to degradation.

Antioxidants are used to prevent the oxidation of active substances and excipients in the finished product. There are three main types of antioxidants:

1. True Antioxidants (water insoluble): They act by a chain-termination mechanism by reacting with free radicals, e.g. butylated hydroxytoluene.
2. Reducing Agents (water soluble): They have a lower redox potential than the drug and get preferentially oxidized, e.g. ascorbic acid. Thus, they can be consumed during the shelf-life of the product.
3. Antioxidant Synergists (chelating agents): These enhance the effect of antioxidants, e.g. EDTA.

Antioxidants are added to pharmaceutical formulations as redox systems possessing higher oxidative potential than the drug that they are designed to protect, or as chain inhibitors of radical induced decomposition. In general, the effect of antioxidants is to break up the chains formed during the propagation process by providing a hydrogen atom or an electron to the free radical and receiving the excess energy possessed by the activated molecule.

Although the selection of an antioxidant can be made on sound theoretic grounds based on the difference in redox potential between the drug and antioxidant, electrometric measurements only rarely predict the efficiency of antioxidants in complex pharmaceutical systems. The effectiveness of an antioxidant or the comparative value of various antioxidants for a particular pharmaceutical preparation is best accomplished by subjecting the pharmaceutical system with the antioxidant to standard oxidative conditions and periodically assaying the formulation for both drug and antioxidant. Although this method may require maximum effort, it yields the most useful information.

It should be remembered that because of the complexity of free radical oxidative processes and their sensitivity to trace amounts of impurities, attempts to compare the effectiveness of antioxidants among different

pharmaceutical systems are of limited validity. Antioxidant materials used in pharmaceutical systems are listed in Tables 12.9 and 12.10.

**Table 12.9:** Antioxidants commonly used foraqueous systems

| | |
|---|---|
| Sodium sulfite | Ascorbic acid |
| Sodium metabisulfite | Isoascorbic acid |
| Sodium bisulfite | Thioglycerol |
| Sodium thiosulfate | Thioglycolic acid |
| Sodium formald\ehyde sulfoxylate | Cysteine hydrochloride |
| Sulfur dioxide | Acetylcysteine |

**Table 12.10:** Antioxidants commonlyused for oil systems

| | |
|---|---|
| Ascorbyl palmitate | Butylated hydroxy toluene |
| Hydroquinone | Butylated hydroxy anisole |
| Propyl gallate | α-tocopherol |
| Nordihydroguaiaretic acid | Lecithin |

Water-soluble antioxidants act by preferentially undergoing oxidation in place of the drug. Oil-soluble antioxidants serve as free radical acceptors and inhibit the free radical chain process. Sulfurous acid salts consume molecular oxygen present in solution. Structurally, other antioxidants have the property of losing a hydrogen free radical and/or an electron. The effectiveness of these antioxidants can depend on the concentration used, whether they are used singularly or in combination, the solution pH, and the package integrity and nonreactivity.

Although sodium metabisulfite has been used extensively in the past and is still used to a considerable extent as an effective antioxidant, recent reports have indicated that the antioxidant activity of this substance is inhibited by a number of compounds, that it actually undergoes degradation itself. The effectiveness of bisulfite as an antioxidant in typical pharmaceutical systems depends on the ease with which this compound is oxidized in comparison with the drug it is to protect. Substances that inhibit bisulfite oxidation may exert important effects on the overall stability of the product by decreasing

the antioxidant effect of bisulfite. Typical substances that can inhibit the oxidation of bisulfite are mannitol, phenols, inorganic anions, aldehydes, ketones, and alkaloids.

## Chelating Agents

The effectiveness of antioxidants can be enhanced through the use of synergists such as chelating agents. Chelating agents tend to form complexes with the trace amounts of heavy metal ions inactivating their catalytic activity in the oxidation of medicaments. Examples of some chelating agents are ethylenediamine tetraacetic acid derivatives and salts, dihydroxyethyl glycine, citric acid, and tartaric acid.

## EXCIPIENTS FOR SOLID DOSAGE FORMS

Solid dosage forms include different types of compressed tablets, granules, troches, lozenges, coated dosage forms, and hard and soft gelatin capsules. In order to produce accurate and reproducible dosage form it is essential that each component is uniformly dispersed within the mixture and the segregation of components should be minimum. The processing operation also demands that the mixture should be free-flowing and cohesive when compressed. The basis of selection of a particular raw material for solid dosage form is based on properties such as particle size, particle shape, porosity, density, moisture content, crystalline form, strength and friability, flow properties, compatibility, disintegrating potential, lubricating potential, etc.

Frequently the single dose of the API is small and an inert substance is added to increase the bulk in order to make the tablet a practical size for compression, improved cohesion, and permit use of direct compression manufacturing or to promote flow. Dilutents used for this purpose are both organic and inorganic materials (lactose, sucrose, mannitol, sorbitol, cellulose, calcium carbonate, Dicalcium phosphate, kaolin, sodium chloride, etc.) (Table 12.11). Binders or granulators impart cohesiveness to the tablet formulation which ensures the tablet remaining intact after compression, as well as improving the free flowing qualities by the formulation of granules of desired hardness and size. Most commonly used binders are gelatin, povidone, starch, carboxymethyl cellulose, hydroxypropyl methylcellulose, (HPMC), hydroxypropyl cellulose (HPC), and maltodextrin. Polyethylene glycols, fatty acids, fatty alcohols, waxes, and glycerides are some examples of meltable binders. The surface tension and the viscosity of a binder play important roles in granulation because these properties influence the liquid bridges between the particles, as well as the distribution of the binder during the wet massing stage. Disintegrating agents are routinely included in tablet formulations and in some hard shell capsule formulations to promote moisture penetration and dispersion of the matrix of the dosage form in dissolution fluids. The mechanism of action of disintegrants includes water wicking, swelling, deformation recovery, repulsion, and heat of wetting. It seems likely that no single mechanism can explain the complex behavior of the disintegrants. The choice of a suitable type and an optimal amount of

disintegrants is paramount for ensuring a rapid disintegration rate. While there are some tablet fillers (e.g. starch and microcrystalline cellulose) which aid in disintegration, there are more effective agents referred to as superdisintegrants. Sodium Starch Glycolate and Croscarmellose are used in capsules because of their greater swelling and wicking capability. Glidants and lubricants are materials that are added to tablet formulations to improve flow properties of the granulation and to prevent the adherence of granules to the punch and die faces of the tablet press, reduce interparticle friction and facilitate the ejection of the tablets from the die cavity.

**Table 12.11:** Properties of some commonly used diluents

| Diluent Maximum | Marketed products | Particle aspect ratio | Characteristics | Mean volumetric diameter (µm) | True density (g/cm$^3$) | Maximum compression stress (MPa) | Tensile strength (MPa) |
|---|---|---|---|---|---|---|---|
| Micro-crystalline cellulose | Avicel PH 102 | 1.5–3 | Equant,column | 123 | 1.58 | 397 | 14.3 High |
| | Avicel PH 105 | 2–3 | Flake | 21 | 1.55 | 443 | 13.6 High |
| | Avicel PH 302 | 2–4 | Equant, lath | 127 | 1.54 | 496 | 10.0 High |
| Lactose | Direct tableting lactose | 1–2 | Equant, plate | 153 | 1.50 | 700 | 10.7 High |
| | Lactose–310 | 1.5–2 | Plate, column | 79 | 1.54 | 700 | 7.2 High |
| | Lactose-316 fast flo | 1–1.5 | Equant | 109 | 1.54 | 564 | 7.7 High |
| Calcium Phosphate Dibasic | Emcompress | 1 | Equant | 173 | 2.26 | 700 | 8.2 High |
| | CD anhydrous | 1 | Equant | 13 | 2.83 | | |
| | A-TAB | 1 | Equant | 184 | 2.78 | 700 | 9.6 High |
| Mannitol | D-(-) Mannitol | 2–4 | Column, lath | 97 | 1.44 | 365 | 2.0 Low |
| | Mannogem 2080 | 1–2 | Equant | 457 | 1.45 | 333 | 2.2 Low |
| | Mannogem EZ | 1–2 | Equant | 112 | 1.45 | 582 | 4.1 Moderate |

## Multifunctional Excipients

The obvious advantages of solid dosage forms and changing technological requirements keep alive the search for newer excipients. The newer excipients are required to be compatible not only with the latest technologies and production machineries, but also with the innovative active principles such as those originating from biotechnology. Developments in the field of excipients and manufacturing machinery have helped in establishing traditional inert excipients as functional components. With the increasing number of new drug moieties of varying physicochemical and stability properties being pushed into the development pipeline, there is a growing pressure on formulators to search for newer excipients to achieve the desired set of functionalities. The development of new excipient involves relatively high cost. This leaves modification of physicochemical property of the existing excipient the most successful strategy in the development of new excipients. Multifunctional excipients or high-functionality excipients are those that contribute at least two functions to formulations through a single ingredient, require no complex processing and improve the performance of the formulations for better drug products. Excipient performance has been improved by modifying already approved compounds. The concept of altering excipient functionality by retaining the favorable attributes and supplementing with newer ones, by processing the parent excipient with another excipient is termed as coprocessing. These developed tailor-made designer excipients help in faster product development. Excipients undergoing these approaches may be advantageous in their formulation, manufacture, and marketing. In formulation, these excipients may decrease strain rate sensitivity, increase rework potential, increase dilution potential, decrease lubricant sensitivity, enhance flow properties, enhance the blending process, optimize content uniformity, increase compression ratio, facilitate material handling, require smaller quantities, decrease environmental concerns, and improve stability. These formulation benefits can lead to manufacturing advantages such as enable direct compaction to avoid time-consuming wet granulation, increase production capacity using excipients with enhanced flow and compaction behavior, reduce tablet tooling and machine wear, and eliminate the facility need of solvent recovery. Benefits such as rapid formulation development, smaller tablet size, better quality products, and no solvent residues may be possible by using these excipients

with proven functionality. For example, Avicel CE-15 (FMC BioPolymer, Newark, Delaware, USA), a coprocessed excipient of MCC and guar gum, designed for providing chewable tablets with reduced grittiness and tooth packing, minimal chalkiness, better mouth feel, and improved overall palatability. Mannitol with sorbitol when coprocessed resulted in interlocked crystals with stronger binding capacity and eased the dispensing of orally dissolving tablet formulations in conventional bottles, eliminating the need for specialized packaging, and thus providing signicant cost savings. Low-substituted hydroxypropyl cellulose can facilitate disintegration and prevent capping during tableting. Hydrogenated vegetable oil (e.g. Lubritab, Serotex), distilled glyceryl monostearate (e.g. Myvaplex 600P), glyceryl behenate (Compritol 888 ATO), and glyceryl palmitostearate (Precirol ATO5) are promoted as tablet and capsule lubricants and sustained-release agents. Microcrystalline cellulose can be a bulking agent and compression aid to impart high compactibility-compressibility, good flow behavior, improve blending, and possibly enhance disintegration to drug formulations. Nonsoluble, high-swell, pregelatinized starch has been promoted as a carrier for hygroscopic ingredients, a stabilizer for moisture-sensitive drugs, and a granulation aid for high yield, fast disintegration, and dissolution enhancement.

**Table 12.12:** List of some modified excipients and the process involved

| Examples of marketed excipients | Modification |
|---|---|
| Emcompress and Emocel, Di-tab, Avicel, Flowlac, Lactopress, Pharmatose, Xylitab, Maltrin, Parteck M, Dipac, Advantose, Mannogem, Carbopol 71G | Modification of particle-size distribution via milling, sieving, agglomerating, or roller compacting |
| Dipac and emdex α-lactose anhydrous and β-lactose monohydrate, glucose monohydrate | Cocrystallization Selection of a particular crystal polymorphic form |
| Avicel, Ceolus, Emcocel | Isolation of crystalline portion of cellulose fibre chain |
| Avicel PH 103, 112, 113, Starch 1500 LM, Emcocel LM50 and XLM90 | Drying to reduce moisture content |

| | |
|---|---|
| Avicel PH 301, 302, Prosolv HD 90, Ceolus PH 301, 302 | Densified materials |
| Surelease and Acryl-EZE, Aquacoat ECD, Eudragit RS30D, NE30D | Nano-sized water dispersion/redispersible form |

These coprocessed multifunctional excipients can be considered to retain the generally regarded as safe (GRAS) status if the parent excipients are also GRAS certied by the regulatory agencies; if any chemical change is absent during processing. Table 12.13 lists some of the marketed coprocessed excipients.

**Table 12.13:** List of some marketed co-processed excipients

| Trade name | Coprocessed excipient | Added advantages |
|---|---|---|
| Ludipress | Lactose monohydrate (93%), Kollidon 30 (3.5%), and Kollidon CL (3.5%) | Lower hygroscopicity, good ow ability, tablet hardness independent of machine speed |
| Cellactose 80 | *α*-Lactose monohydrate (75%) and cellulose powder (25%) | Highly compressible, good mouth feel, better tableting at low cost |
| StarLac | *α*-Lactose monohydrate (85%) and maize starch (15%) | Good ow, optimized disintegration, excellent tablet hardness |
| Pharmatose DCL14 | Anhydrous P-lactose (95%) and lactitol (5%) | High compactibility, superior owing properties, low lubricant sensitivity |
| Avicel CE-15 | MCC and guar gum | Less grittiness, reduced tooth packing, minimal chalkiness, creamier mouth feel, improved |

| | | |
|---|---|---|
| | | overall palatability |
| Vitacel VE-650 | MCC (65%) and calcium carbonate (35%) | Direct compression, encapsulation |
| Formaxx $CaCO_3$70 | Calcium carbonate (70%) and sorbitol (30%) | High compressibility, excellent taste masking, free ow, superior content uniformity, controlled particle size distribution |
| Di-Pac | Sucrose (97%) and dextrin (3%) | Directly compressible, Low hygroscopicity |
| Advantose FS 95 Fructose | Fructose (95%) and Starch (5%) | Excellent ow, good compressibility, tablets hold shape well, but are very chewable |
| Captisol | Modified cyclodextrin | New chemical entity, improved water sulfobutylether-β-cyclodextrin solubility |
| Kollidon CL, CL-F, CL-M | Crosslinked water-insoluble polyvinyl pyrrolidone | Size modified according to application for disintegration and solubility enhancement |
| Xylitab 100 | Xylitol and polydextrose | Directly compressed sugar with improved mouth-feel |
| Eudragit RL and RS LustreClear | Methacrylic acid polymers MCC and Carageenan | Modified for sustain release Efficient tablet coating with short hydration time prior to coating and fast drying time |

## Excipients for Liquid Dosage Forms

Liquid dosage forms have typically been targeted for use in geriatric and pediatric patients and include solutions, syrups, suspensions, emulsions, elixirs, or concentrates. In aqueous based formulations, solubilizers are used to modify the polarity of water to allow an increase in the solubility of a nonpolar drug (propylene glycol (PG), alcohols such as ethanol, sugars such as sorbitol, or polyethylene glycols such as PEG-400). On addition of a solubilizer polarity, density, surface tension, viscosity, boiling point, and specific heat of solution, all are affected in various ways. Solubility of a drug in solution may also be enhanced by using a complexing agent such as a cyclodextrins. Wetting agents are used to create a homogenous dispersion of solid particles in a liquid vehicle and allow removal of adsorbed air and easy penetration of the liquid vehicle into pores of the particle in a short period of time. For an aqueous vehicle, alcohol, glycerin, and PG are frequently used to facilitate the removal of adsorbed air from the surface of particles whereas for a non-aqueous liquid vehicle, mineral oil is commonly used as a wetting agent. Hydrophobic particles are not easily wetted even after the removal of adsorbed air. Hence, it is necessary to reduce the interfacial tension between the particles and the liquid vehicle by using a surface-active agent. Hydrophilic particles, however, do not require the use of such surface-active agents for their solubilization. In addition to the concentration of surfactant, the location of the drug or excipient in the micelle structure can influence its stability. Suspending agents function in pharmaceutical systems to impart viscosity, and as such retard particle sedimentation. A number of factors must be considered in the selection of the appropriate agent including desired rheological property, suspending ability in the system, chemical compatibility with other excipients, pH stability, length of time to hydrate, batch-to-batch reproducibility, and cost. Suspending agents can be classied into cellulose derivatives, clays, natural gums, and synthetic gums. In many cases, these excipients are used in combination.

## Excipients for Coating

Site-specific release formulations usually use an external coating to allow the release of drug in a specific region of the gastrointestinal tract. For example, pH sensitive polymers are used to facilitate release in the small intestine where the pH is much higher than that in the stomach. Similarly, colon-targeting formulations use excipients that are susceptible to particular enzymes present in the colon. Enteric coatings are useful for protection of drugs that are labile to acidic environment in the stomach and from presystemic destruction. Enteric coating materials usually dissolve at pH higher than 5 and readily dissolve at pH 7. Shellac (puried lac) dissolves at pH greater than 7 and can be used alone or in combination with other materials. Cellulose acetate phthalate (CAP) dissolves at pH higher than 6; polyvinyl acetate phthalate (PVAP), hydroxypropyl methyl-cellulose acetate succinate (HPMCAS) and hydroxypropyl methylcellulose phthalate (HPMCP) dissolves at pH greater than 5.

Sustained-release in the small intestine can be achieved using a film coating or a sustained-release matrix or sustained-release drug-loaded granules. Sustained release film coatings can be applied to tablets, granules, or beads. Various materials can be used as films for sustained-release of drugs; film coatings can form two types of membranes: permeable and semipermeable. The permeable membrane allows the intestinal fluid to enter the dosage forms to dissolve the drug as well as allow the drug to permeate out of the dosage form through the membrane whereas semipermeable membranes are permeable only to the intestinal fluid but impermeable to drug molecules dissolved. Materials that form a permeable membrane include fats, bee wax, carnauba wax, cetyl alcohol, cetylsteryl alcohol, zein, acrylic esters, silicone elastomers, and ethylcellulose. Aqueous dispersions of water-insoluble polymers are commonly used for sustained-release film coatings. Examples of commercially available aqueous polymer dispersions include Surelease-containing ethylcellulose, Aquacoat-containing ethylcellulose, Eudragit RS 30 D-containing poly(ethylacrylate-methylmethacrylate) triethylammonioethyl methacrylate chloride 1:2:0.1, Eudragit RL 30 D-containing poly (ethylacrylate-methylmethacrylate) triethy-lammonioethyl methacrylate chloride 1:2:0.2, and Eudragit NE 30 D-containing poly (ethylacrylate-methylmethacrylate) 2:1.

Methacrylic acid copolymer coatings (for example, Eudragit RL and Eudragit RS) are insoluble but permeable throughout the gastrointestinal tract. The detail of commercially available Eudragit grades is shown in Table 12.14. Plasticizers used with Eudragit to reduce the glass transition temperatures of Eudrugit films include polyethylene glycol (PEG), propylene glycol, diethylphthalate, dibutylphthalate, and triacetin. In osmotic pumps, semipermeable membrane allows water to enter the table matrix through the driving force of osmotic pressure while preventing the permeation of drug molecules across the membranes. Excipients used for the semipermeable membranes include polyvinyl alcohol, cellulose acetate, and ethylcellulose. Polyacrylic acid is a commonly used matrix for sustained-release formulations and is available as Carbomer 910, 934, 934P, 940, 941, 971P, and 974P. Carbopol (Noveon, Cleveland, Ohio, USA) polymers are polymers of acrylic acid cross-linked with polyalkenyl ethers or divinyl glycol. Upon exposure to intestinal fluid, Carbopols swell to form hydrogel-like matrices through which drug molecules could be released at a controlled rate. In addition to Carbopol, there are other materials used to form sustained-release matrix, including methylcellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose (CMEC), carnauba wax, and glyceryl palmitostearate.

Another strategy to control drug release is through formation of coprecipitates with pharmaceutical excipients. Ibuprofen is a drug with adequate solubility and membrane permeability for complete oral absorption. Coprecipitates of anionic, cationic, or zwitterionic Eudragits (methacrylate polymers and copolymers) with ibuprofen deterred the release rates of ibuprofen. Although no signicant interactions were observed between ibuprofen and any Eudragit, and the crystalline state of ibuprofen was not altered, the release of ibuprofen was slowed down by the swelling and slow dissolution of Eudragits.

**Table 12.14:** Commercially available Eudragit grades and their functions

| EUDRAGIT | Application | Availability | Function ability | Dissolution Properties | Advantages |
|---|---|---|---|---|---|
| **pH Dependent grades** EUDRAGIT® L100-55 | Drug delivery in duodenum | Powder | | Dissolution above pH 5.5 | Effective and stable enteric coating with a fast dissolution in the upper bowel |
| EUDRAGIT®L 30 D-55 | Drug delivery in duodenum | 30% aqueous dispersions | | | |
| EUDRAGIT® L 100 | Drug delivery in jejunum | Powder | Anionic polymers with meth-acrylic acid as a functional group | Dissolution above pH 6.0 | Granulation of drug substance in powder form for controlled release |
| EUDRAGIT®S 100 | Drug delivery in ileum | Powder | | | Site specific drug delivery in intestine by combination of EUDRAGIT®L/S grades |
| EUDRAGIT®FS 30 D | Colon delivery | 30% aqueous dispersions | | Dissolution above pH 7.0 | Variable release profile |
| EUDRAGIT®E 100 | Taste masking, odour masking | Granules | Cataionic polymer with diethyl aminoethyl-methacrylate as a functional group | Soluble in gastric fluid upto pH 5.0 | Low viscosity, high pigment binding capacity good adhesion, low polymer weight gain |
| EUDRAGIT®E PO | Insulating coatings | Powder | | Swellable and permeable above pH 5.0 | |
| **pH Independent grades** | | | | | |
| EUDRAGIT®RL 30 D | Sustain release formulations | 30% aqueous dispersions | | Insoluble | Customised release profile by combination of RL and RS in different ratio |
| EUDRAGIT®RL PO | Sustain release formulations | Powder | | High permeability | |
| EUDRAGIT®RL 100 | Sustain release formulations | Granules | Meth-/ acrylates copolymers with trimethyl ammonio-ethylmetha-crylate as a functional group | pH independent swelling | |
| EUDRAGIT®RS 30 D | Sustain release formulations | 30% aqueous dispersions | | Insoluble | |
| EUDRAGIT®RS PO | Sustain release formulations | Powder | | Low permeability | Suitable for matrix structures |
| EUDRAGIT®RS 100 | Sustain release formulations | Granules | | pH independent swelling | |
| EUDRAGIT®NE 30 D | Sustain release formulations | 30% aqueous dispersions | | Insoluble | No plastic required, highly flexible, |
| EUDRAGIT®NE 40 D | Suitable for matrix structures and miscible with other EUDRAGIT® grades | 40% aqueous dispersions | Natural polymer of meth-/ acrylates | Low permeability<br>pH independent swelling | Suitable for matrix structures |

## EXCIPIENTS FOR NDDS AND DRUG TARGETING

Carriers for biopharmaceutical therapeutic agents range from well-established excipients of natural origin to custom-made synthetic materials with putatively enhanced protective or targeting features. Natural or semisynthetic materials predominate however. Sources as diverse as primitive marine plants (chitosans and alginates), plant or animal phospholipids (egg and seed lecithin), and mammalian collagens (gelatin) are being mined for useful delivery or targeting aids, as well as for components of complex formulations such as microemulsions or liposomes. The wide use of biological materials may reflect the need for long and expensive safety evaluation of novel synthetic materials prior to use in man. This hinders timely evaluation other than in vitro or animal models. More esoteric materials that confer target specificity include glycoproteins, recombinant proteins, or monoclonal antibodies. To date, clinical performance of such carrier systems has been disappointing. Further refinement of concepts and materials may be necessary before the performance matches the promise. Attenuated adenoviruses have been used as vectors where delivery to cell nuclei is required (e.g. in gene medicine). It is a moot point whether these or other targeting or carrier materials are “excipients” part of a prodrug or something in-between. The boundaries between “active” and “inactive” materials are much less clear in such cases. The traditional approach of evaluating a novel entity in its own right in animal safety programs and then formulating with “inert” materials is inappropriate with sophisticated delivery systems because of the important effect of the adjuvant on disposition and kinetics of the active ingredient.

## EXCIPIENTS FOR PROTEIN DELIVERY

With the beginning of recombinant DNA technology, protein-based drugs have become continually and increasingly common place in the repertoire of medicines available to medical practitioners for the treatment of a wide range of diseases from cancer to rheumatoid arthritis. Among the first recombinant protein drugs to be approved were naturally occurring growth factors, hormones, blood factors, and cytokines. These were followed by monoclonal antibody (mAb)-based therapies. The initial antibody-based drugs were met with limited success because they were of nonhuman origin (e.g. murine) and were often prone to immunological reactions. However, as technologies to express and produce recombinant humanized and fully human monoclonal antibodies have matured, there has been a tremendous resurgence in the success and development of mAb-based drugs. As of the year 2005, there were more than 150 approved protein-based pharmaceuticals on the market, and this number is expected to rise dramatically in the coming years. Although stable formulations for small-molecule drugs can be developed by minimizing chemical degradation pathways, protein formulation development can be far more challenging. Proteins are complex molecules with defined primary, secondary, tertiary, and in some cases quaternary structures, all of which are essential for highly specific biological functions. In addition to chemical degradation, protein drugs are susceptible to the physical degradation process of irreversible aggregation. Protein aggregation is of particular interest because it often results in diminished bioactivity that affects drug potency, and can also elicit serious immunological or antigenic reactions in patients. Chemical degradation of a protein drug has also been implicated in increasing its immunogenic potential. Thus, stable protein formulations require that both physical and chemical degradation pathways of the drug be minimized.

Excipients are additives that are included in a formulation, because they either impart or enhance the stability, delivery, and manufacturability of a drug product. Regardless of the reason for their inclusion, excipients are an integral component of a drug product and therefore need to be safe and well tolerated by patients. For protein drugs, the choice of excipients is particularly important because they can affect both efficacy and immunogenicity of the drug. Hence, protein formulations need to be

developed with appropriate selection of excipients that afford suitable stability, safety, and marketability. Protein-based drugs have been formulated mainly as stable liquids or in cases where liquid stability is limiting as lyophilized dosage forms to be reconstituted with a suitable diluent prior to injection. This is because their delivery has been limited primarily to the parenteral routes of intravenous (IV), subcutaneous (SC), or intramuscular (IM) administration. There are a few drugs that have been developed for pulmonary delivery, such as DNAase (Pulmozyme®) and an inhalable formulation of insulin (e.g. Exubra®). However, even such drugs have been formulated as either liquid or lyophilized or spray-dried powders.

The principal challenge in developing formulations for therapeutic proteins is stabilizing the product against the stresses of manufacturing, shipping, and storage. The role of formulation excipients is to provide stabilization against these stresses. Excipients may also be employed to reduce viscosity of highly concentrated protein formulations to enable their delivery and enhance patient convenience. A liquid formulation is usually comprised of a buffering agent, a stabilizer (which may also serve as a tonicity agent), a surfactant, and an anti-oxidant when protein oxidation is signicant. Chelating agents are employed when metal ion catalyzed reactions predominate. A preservative may be included when a multi-dose formulation is desired.

A lyophilized formulation is usually comprised of a buffer, a bulking agent, and a stabilizer. The utility of a surfactant may be evaluated and selected in cases where aggregation during the lyophilization step or during reconstitution becomes an issue. An appropriate buffering agent is included to maintain the formulation within stable zones of pH during lyophilization.

Bulking agents are typically used in lyophilized formulations to enhance product elegance and to prevent blowout. Conditions in the formulation are generally designed so that the bulking agent crystallizes out of the frozen amorphous phase (either during freezing or annealing above the $T_g$) giving the cake structure and bulk. Mannitol and glycine are examples of commonly used bulking agents. Stabilizers include a class of compounds that can serve as cryoprotectants, lyoprotectants, and glass forming agents. Cryoprotectants act to stabilize proteins either during freezing or in the frozen state at low temperatures lyoprotectants stabilize proteins in the freeze-dried solid dosage form by preserving the native-like conformational properties of the protein

during dehydration stages of freeze drying.

The stability of a protein drug is usually observed to be maximal in a narrow pH range. This pH range of optimal stability needs to be identied early during preformulation studies.

Buffers for lyophilized formulations need additional consideration. Some buffers like sodium phosphate can crystallize out of the protein amorphous phase during freezing resulting in rather large shifts in pH. Other common buffers such as acetate and imidazole should be avoided since they may sublime or evaporate during the lyophilization process, thereby shifting the pH of formulation during lyophilization or after reconstitution.

Salts are often added to protein parenterals to increase the ionic strength of the formulation, which can be important for protein solubility, physical stability, and isotonicity. Salts can affect the physical stability of proteins in a variety of ways. Ions can stabilize the native state of proteins by binding to charged residues on the protein's surface. Alternatively, they can stabilize the denatured state by binding to the peptide groups along the protein backbone (–CONH-). Salts can also stabilize the protein native conformation by shielding repulsive electrostatic interactions between residues within a protein molecule. Electrolytes in protein formulations can also shield attractive electrostatic interactions between protein molecules that can lead to protein aggregation and insolubility.

**Table 12.15:** Commonly used buffering agents

| S.No. | Buffers | Example of drug product |
|---|---|---|
| 1 | Acetate | Neupogen®, Neulasta® |
| 2 | Succinate | Actimmune® |
| 3 | Citrate | Humira® |
| 4 | Histidine | Xolair® |
| 5 | Phophate | Enbrel® |
| 6 | Tris | Leukine® |

Amino acids have found versatile use in protein formulations as buffers, bulking agents, stabilizers, and antioxidants. Histidine and glutamic acid are

employed to buffer protein formulations in the pH range of 5.5 to 6.5 and 4.0 to 5.5, respectively. Glutamic acid is particularly useful in such cases (e.g. Stemgen®). Histidine is commonly found in marketed protein formulations (e.g. Xolair®, Herceptin®, Recombinate®). It provides a good alternative to citrate, a buffer known to sting upon injection. Interestingly, histidine has also been reported to have a stabilizing effect on ABX-IL8 (an IgG2 antibody) with respect to aggregation when used at high concentrations in both liquid and lyophilized presentations. The amino acids glycine, proline, serine, and alanine have been shown to stabilize proteins by the mechanism of preferential exclusion. Glycine is also a commonly used bulking agent in lyophilized formulations (e.g. Neumega®, Genotropin®, and Humatrope®). It crystallizes out of the frozen amorphous phase, giving the cake structure and bulk. Arginine has been shown to be an effective agent in inhibiting aggregation and has been used in both liquid and lyophilized formulations (e.g. Activase®, Avonex®, and Enbrel® liquid). Furthermore, the enhanced efciency of refolding of certain proteins in the presence of arginine has been attributed to its suppression of the competing aggregation reaction during refolding.

Polyols encompass a class of excipients that includes sugars (e.g. mannitol, sucrose, and sorbitol), and other polyhydric alcohols (e.g. glycerol and propylene glycol). We have included the polymer polyethylene glycol (PEG) in this category for ease of discussion. Polyols are commonly used as stabilizing excipients and/or isotonicity agents in both liquid and lyophilized parenteral protein formulations. Oxidation of protein residues arises from a number of different sources. Beyond the addition of specic antioxidants, the prevention of oxidative protein damage involves the careful control of a number of factors throughout the manufacturing process and storage of the product, such as atmospheric oxygen, temperature, light exposure, and chemical contamination. The most commonly used pharmaceutical antioxidants are reducing agents, oxygen/free-radical scavengers, or chelating agents.

In general, transition metal ions are undesired in protein formulations because they can catalyze physical and chemical degradation reactions in proteins. However, specic metal ions are included in formulations when they are cofactors to proteins and in suspension formulations of proteins where they form coordination complexes (e.g. zinc suspension of insulin). Recently,

the use of magnesium ions (10–120 mM) has been proposed to inhibit the isomerization of aspartic acid to isoaspartic acid.

Preservatives are necessary when developing multidose parenteral formulations that involve more than one extraction from the same container Although preservatives have a long history of use with mall-molecule parenterals, the development of protein formulations that include preservatives can be challenging. Preservative caves almost always have a destabilizing effect (aggregation) on proteins, and this has become a major factor in limiting their use in multidose protein formulations. To date, most protein drugs have been formulated for single-use only. However, when multidose formulations are possible, they have the added advantage of enabling patient convenience and increased marketability. A good example is that of human growth hormone (hGH) where the development of preserved formulations has led to commercialization of more convenient, multidose injection pen presentations. At least four such pen devices containing preserved formulations of hGH are currently available on the market. Norditropin® (liquid, Novo Nordisk), Nutropin AQ® (liquid, Genentech), and Genotropin (lyophilized—dual chamber cartridge, Pharmacia & Upjohn) contain phenol, whereas Somatrope® (Eli Lilly) is formulated with m-cresol.

## EXCIPIENTS FOR VACCINE DELIVERY

Vaccines existing in the market contain various types of antigens, adjuvants, and additives, which in combination provide maximum protection against several infectious diseases (Table 12.16). These vaccines might contain live or killed viruses, puried viral proteins, inactivated bacterial toxins, and polysaccharides or puried subunit recombinant proteins. The diverse nature of these antigens requires different excipients be used to stabilize them for use within their designated shelf life. The excipients used in vaccines for parenteral administration should be evaluated to meet higher purity and safety standards because these are injected into the human body. Because most commonly used vaccines are administered parenterally, excipients must comply with strict guidelines set forth by the US Food and Drug Administration (FDA) for any parenteral dosage form.

**Table 12.16:** Excipients used in vaccines

| Vaccine | Formulation excipients |
|---|---|
| Anthrax (BioThrax) | Aluminum hydroxide, amino acids, benzethonium chloride, formaldehyde or formalin, inorganic salts and sugars, vitamins |
| BCG (Tice) | Asparagine, citric acid, lactose, glycerin, iron ammonium citrate, magnesium mulfate, potassium phosphate |
| DTaP (Daptacel) | Aluminum phosphate, ammonium sulfate, casamino acid, dimethyl-betacyclodextrin, formaldehyde or formalin, glutaraldehyde, 2-phenoxyethanol |
| DTaP (Infanrix) | Aluminum hydroxide, bovine extract, formaldehyde or formalin, glutaraldhyde, 2-phenoxyethanol, polysorbate 80 |
| DTaP (Tripedia) | Aluminum potassium sulfate, ammonium sulfate, bovine extract, formaldehyde or formalin, gelatin, polysorbate 80, sodium phosphate, thimerosal |
| DTaP/Hib (TriHIBit) | Aluminum potassium sulfate, ammonium sulfate, bovine extract, formaldehyde or formalin, gelatin, |

| | |
|---|---|
| | polysorbate 80, sucrose, thimerosal |
| DTaP-IPV (Kinrix) | Aluminum hydroxide, bovine extract, formaldehyde, lactalbumin hydrolysate, monkey kidney tissue, neomycin sulfate, polymyxin B, polysorbate 80 |
| DTaP-HepB-IPV (Pediarix) | Aluminum hydroxide, aluminum phosphate, bovine protein, lactalbumin hydrolysate, formaldehyde or formalin, glutaraldhyde, monkey kidney tissue, neomycin, 2-phenoxyethanol, polymyxin B, polysorbate 80, yeast protein |
| DtaP-IPV/Hib (Pentacel) | Aluminum phosphate, bovine serum albumin, formaldehyde, glutaraldhyde, MRC-5 DNA and cellular protein, neomycin, polymyxin B sulfate, polysorbate 80, 2-phenoxyethanol, |
| Hib/Hep B (Comvax) | Amino acids, aluminum hydroxyphosphate sulfate, dextrose, formaldehyde or formalin, mineral salts, sodium borate, soy peptone, yeast protein |

# DRUG-EXCIPIENT, EXCIPIENT-EXCIPIENT INTERACTIONS

In order to develop and manufacture a medicine, three main components are considered: drug, excipients and the manufacturing process. For some types of product the primary packaging may also need to be considered. It is important to understand the properties, advantage, limitations and interaction of these three components. For the transformation of a drug into a medicinal product, the interactions between excipients and the other two components (the drug and the manufacturing process), and/or between two or more excipients, are fundamental criteria. Excipient interactions are a large part of why medicines work (and sometimes why they do not work in development). They can be either beneficial or detrimental and depends on the particular application, however it must be reemphasized that excipient interactions are not always detrimental. For example, what may be beneficial for a prolonged release product may be detrimental in an immediate release product, and vice versa. Sometimes they can be used to our advantage, particularly in the areas of product manufacture and drug delivery systems. Interactions occur more frequently between excipient and active principle than between excipient and excipient. Excipient interactions can have implications for drug stability, product manufacture, drug release (dissolution; both in vitro and in vivo), therapeutic activity, and side effect profile. Interactions between drugs and excipients can occur by means of several possible mechanisms, including adsorption, complexation, chemical interaction, pH effects, and eutectic formation, resulting in drug products with desired or undesired properties.

Excipient interactions can be classied simply as: *Physical*—physical interactions do not involve chemical change. The components retain their molecular structure, *Chemical*—chemical interactions involve chemical reactions; i.e. a different molecule(s) is created, and *Physiological/Biopharmaceutical*—Physiological interactions are the interactions between the excipient(s) and the body fluids. In reality, they are also physical interactions, but since they are so important, and because they occur after the medicine has been administered to the patient, they can be considered separately.

## Physical Interactions

Physical interactions involving excipients are quite common and are frequently used in pharmaceutical science to aid processing and to aid or modify drug dissolution (such as oral modied release) or distribution in the body (such as with the use of a parenteral modied release product). Some of these interactions are deliberately invoked to produce a certain effect. Others are unintended, and it is these interactions that usually cause the problems. Silicied microcrystalline cellulose is an end product of a beneficial interaction between excipients that can improve performance (functionality) under certain circumstances. It is thought that the fumed silica particles prevent the structural collapse of the microcrystalline cellulose that can occur on drying after wet massing during wet granulation. The benefits of the silicication were not seen when attempts were made to prepare the material in situ using conventional pharmaceutical processing. Another example is the one between xanthan gum and locust bean gum (carob gum or ceratonia) in the presence of water, the interaction creates a much more viscous gel system than can be created using either component alone and is used in the formulation of controlled release oral solid dosage forms. Phenolic preservatives, e.g. parabens esters, are inactivated in the presence of nonionic surfactants and this detrimental interaction can have serious consequences for preservation of the product such as pharmaceutical creams, as these are usually oil-in-water emulsions stabilized using a surfactant. The skin penetrationenhancing properties of laurocapram are lost when laurocapram and mineral oil (liquid paraffin) are included in the same formulation.

A balance between beneficial and detrimental effects of magnesium stearate during processing should be made. Magnesium stearate is used as a lubricant and has a polar head and a fatty acid tail and it is believed that the polar head of the magnesium stearate is oriented toward the die wall or tablet punch face. In these ways it is able to reduce the ejection force (the force required to eject the tablet from the die after compaction) and prevent sticking to the punch faces. But magnesium stearate is also recognized for causing problems such as reduced tablet “hardness” and dissolution from tablets and capsules.

Physical interaction between primary amine drugs and microcrystalline cellulose is an example of drug-excipient interaction. Small percentage of the

drug may bound to the microcrystalline cellulose and is not released when dissolution is carried out in water. It is not a major issue for high-dose drugs but for low-dose drugs it can lead to dissolution failures. The problem is sorted by carrying out the dissolution using a weak electrolyte solution for the dissolution medium (e.g. 0.05 M HCl), adsorption onto the microcrystalline cellulose is very much reduced and complete dissolution may be achieved even for low-dose APIs.

The techniques used for identifying physical interactions include differential scanning calorimetry (DSC), isothermal microcalorimetry, thermogravimetric analysis, high performance liquid chromatography, or thin layer chromatography and differential thermal analysis.

## Chemical Interactions

One notable exception to the detrimental nature of most chemical interactions is the beneficial interaction of the effervescent couple whereby sodium bicarbonate reacts with an organic acid, typically citric acid, in the presence of water to generate carbon dioxide, thereby disintegrating the tablet, and forming a solution or suspension of the drug in water depending on its solubility, that can be administered orally. The presence of bicarbonate and citric acid in the same tablet or granule requires that precautions be taken such as manufacturing and packing at very low RH (< 20%) to prevent premature activation of the couple. The packaging also needs to be impermeable to moisture for the same reason.

Interaction may occur between the carbo-nilic groups of a widely-used excipient like polyvinylpyrrolidone, and pharmaceuticals containing donor groups of hydrogen, like famotidine and atenolol, causing problems of incompatibility. Silicon dioxide ($SiO_2$), in anhydrous conditions, behaves like a Lewis acid, giving rise to reactions such as hydrolisis, epimerisation, trans-esterification, etc. One example is the hydrolysis of the imino nitrogen link of nitrazepam with consequent disactivation of the drug.

These reactions can involve excipients, either as reactants or as catalysts. However, these reactions do not always occur. In some instances there may be steric factors in the drug molecule that restrict access to the reactive group and the reaction does not occur, or occurs at a much-reduced rate. For almost all chemical interactions, a key component is presence of “free” (unbound) water. In the absence of a sufficient amount of “free” water, the reactions do not proceed. The “free” water serves to dissolve sufcient of the drug and the excipient, or form bridges between particles, such that the components/reactants come into sufficiently close contact for the reaction to occur. The most common type of chemical reactions are summarized in Table 12.17.

**Table 12.17:** Types of chemical interactions

| Chemical interaction | Effect |
|---|---|
| Maillard reaction: Primary amines with reducing sugars | The glycosidic hydroxyl group of the reducing sugar interacts with the |

| | |
|---|---|
| | primary amine to form an imine (Schiff's base) which further breaks down to form Amidori compounds (responsible for the yellow-brown coloration, e.g. chlorpheniramine and dextrose). |
| Secondary amines with reducing sugars | The reaction cascade does not proceed beyond the formation of the imine, and thus no coloration develops. |
| Primary amines with double bonds | The reaction is analogous to a Michael addition reaction. For example, amine group of fluvoxamine maleate, can interact with the double bond in the maleic acid. |
| Acid or base hydrolysis | In the presence of acid, i.e. anion and hydrogen ion, the reaction is at equilibrium. However, in the presence of base and the associated cations, the reaction is driven to completion (e.g. acetyl salicylic acid and sodium and magnesium salts). |
| Reaction of hetero atoms and active hydrogen. | Lactone formation takes place, e.g. benazepril. |
| Oxidation in presence of metal oxides | Oxidation of atorvastatin and cytidinenucleoside analogues in presence of fumed metal oxides (e.g. fumed silica, fumed titania, and fumed zirconia) |

## Physiological/Biopharmaceutical Interactions

These are physical interactions; the major distinctions being that the interaction is between the medicine (including excipients) and the body fluids, having the potential to influence the rate of absorption of the drug and vary depending on the route of administration. Certain excipients are included in a formulation specically because they interact with the physiological fluids in a certain way. For example, disintegrants are added in immediate release tablet and capsule formulations. When they encounter the aqueous environment of the stomach they will cause the formulation to disintegrate and thereby aid dissolution. Hydrophilic colloid matrices swell and create a diffusion barrier when comes in contact with the aqueous environment of the GIT and thus slows the rate of dissolution of the dissolved drug leading to prolonged release drug delivery system.

A classic biopharmaceutical incompatibility is the interaction between tetracycline antibiotics and calcium and magnesium ions. A complex is formed that is not absorbed from the GIT. This is a well-known interaction, and tetracycline antibiotics usually carry a warning against taking them with certain types of food.

One physiological interaction that can potentially cause serious problems for the patient is the interaction between enteric coatings and antacids. Certain products may be enteric coated to protect the drug from degradation in the stomach, e.g. pro-drugs. Other drugs are enteric coated to protect the stomach from them, e.g. nonsteroidal anti-inammatory drugs (NSAIDs). The entericcoating polymers, e.g. cellulose acetate phthalate and hydroxypropyl cellulose acetate phthalate, rely on their pH-solubility profile for their function; they are soluble at a more basic pH, but insoluble at acid pH. Antacids raise the pH of the stomach contents and thus cause the enteric coating to begin to dissolve in the stomach. The enteric coating thus begins to break-down allowing the premature release of the API in the stomach. For the pro-drugs, this might mean that more of the drug is degraded than is desirable and the patient would receive a suboptimal dose. For the NSAIDs, the premature breakdown of the enteric coat may cause unwanted side effects, such as gastric bleeding.

Some drugs, such as aspirin, appear to be well absorbed along the length of the lower GIT (the ileum and colon). Certain other drugs, e.g. metoprolol,

have a limited absorption in small section of GIT. For these drugs, it is clear that the speed with which the drug passes down the GIT (gastrointestinal motility) will influence absorption of the drug. Certain excipients such as polyols (e.g. sorbitol and xylitol) can increase gastrointestinal motility, and thereby reduce the time available at the site of absorption for drugs such as metoprolol. PEG 400 has also been reported to influence the absorption of ranitidine in a similar fashion.

There are certain mechanisms whereby drugs are actively secreted back into the lumen of the GIT, known as efflux mechanisms; and a major efflux system concerns p-glycoprotein. The signicance of p-glycoprotein is that if a drug is a p-glycoprotein substrate (or a substrate for any other efflux mechanism), it may not matter how well absorbed the drug is, the efflux mechanism is likely to pump the drug back out into the GIT. In the past it has been assumed that certain drugs were just not well absorbed, and that may have been the case, but there is now another possible explanation—that they are substrates for an efflux mechanism. There are recent reports in the literature that at least one excipient, a-tocopheryl PEG 1000 succinate, appears to inhibit the p-glycoprotein efflux pathway. If conrmed, generally, this potentially beneficial biopharmaceutical interaction may have important implications for the oral delivery of certain drugs.

## EXCIPIENT TOXICITY

Pharmaceutical excipients have a vital role in drug formulations, a role that has tended to be neglected as evidenced by the lack of mechanisms to assess excipient safety outside a new drug application process. Currently, it is assumed that an excipient is “approved” when the new drug formulation, of which it is a constituent, receives regulatory acceptance. Excipients can be broadly divided into three categories: established (“approved”), new (novel), and essentially new excipients. Established excipients are well-known materials with long use in pharmaceutical preparations. Recent additions to the ranks of established excipients include materials giving “sugar-free” status to medical preparations, the cyclodextrins, and the hydrofluoroalkane (HFA) inhalation propellants. A novel excipient is a compound which has not been previously used or permitted for use in a pharmaceutical preparation. Essentially new excipients form an intermediate category and include substances resulting from a structural modification of an “approved” excipient, a recognized food additive (or cosmetic ingredient), a structurally modified food additive, or a constituent of an over the counter (OTC) medicine. There is an expectation for relevant quality and preclinical data for excipients together with evidence of no excipient-induced adverse effects in the final formulated clinical drug substance. The specific safety data that may be needed will vary depending upon the clinical situation, including such factors as the duration, level, and route of exposure, but may include acute, repeat-dose, reproductive, and genetic toxicity data, carcinogenicity data, and specialized toxicology information, such as sensitization or local irritation data. Much guidance exist to aid in the development of pharmaceuticals, including the International Conference on Harmonization (ICH) documents and various Food and Drug Administration/Center for Drug Evaluation and Research (FDA/CDER) pharmacology and toxicology guidances. The FDA/CDER has recently adopted a new guidance for industry, “Nonclinical Studies for Development of Pharmaceutical Excipients,” which focuses on issues associated with development of safety databases that will support clinical use of excipients in drug products. All pivotal toxicology studies are performed in accordance with state-of-the-art protocols and good laboratory practice regulations. All potential new excipients be appropriately evaluated for pharmacological activity using a battery of standard tests.These evaluations can be performed during the course of toxicology studies or as

independent safety pharmacology studies. It is useful for these data to be obtained at an early point during the safety evaluation of an excipient, since, if the excipient is found to be pharmacologically active, this information can influence subsequent development. Appropriate regulatory guidance can be given by the responsible review division.

Preclinical evaluation of excipients is not only necessary because of regulatory expectation but because excipients are not inert (as traditionally viewed) and can show adverse toxicological findings by themselves or in drug formulations. Regulators are fully aware that a number of clinically manifested adverse reactions are caused by established excipients, although they are of low occurrence and uncommon when compared to the overall prevalence of adverse drug reactions. Such reactions are commonly of a hypersensitivity, allergic, or anaphylactic nature and the area is well covered in the literature. Examples of known excipient-induced toxicities include renal failure and death from diethylene glycol, osmotic diarrhea caused by ingested mannitol, hypersensitivity reactions from lanolin, and cardiotoxicity induced by propylene glycol. The excipients that have been authorised to be used as food additives have been evaluated as regards toxicology by the JECFA (Joint Expert Committee on Food Additives), which handles the evaluation of the risk from consuming additives or contaminants with food. In the case of additives, their use is voluntary and has a technological reason, exactly as in the case of pharmaceutical excipients, whereas contaminants are substances that can be vehicled by the food chain, given the ubiquitousness of the distribution of pollutants in the environment. The JECFA usually terminates its toxicological evaluations with the publication of an admissible daily intake (ADI), which represents the dose that does not carry risks to the population if taken every day for a life-time. This dose is expressed in mg per kg of daily weight per day. To establish the total daily dose, we should multiply this number by the body weight (usually reckoned as 60 kg). The value of the ADI is extrapolated from studies conducted in laboratory animals, dividing the highest dose without toxic effects in the animal by a safety factor (generally 100).

Certain toxic effects concern only some consumers; but this does not mean that these effects are less important. The groups at risk to be considered with attention are people with allergies and intolerances. The former involve the immunitary system, whereas the latter are determined by genetically-

transmitted metabolic anomalies (phenylketonuria, galactosemia, etc.) or family predisposition (celiachia, diabetes, etc.). For example, Gluten is a proteic component of wheat to which bearers of celiac illness are sensitive. The classic manifestation of these patients' intolerance is an enteropathology (malabsorption syndrome), but atypical symptoms such as herpetiform dermatitis, iron deficiency anrnmia, alopecia and osteoporosis are also encountered. The pathology is linked mainly to gliadine, composed of different proteic fractions, the most toxic of which is the alpha fraction. Lactose, a common pharmaceutical excipient, can also induce intolerance due to the physiological diminution of lactase activity in adults. Many patients experience difficulty in tolerating lactose (in Italy, about 70% of the population), even though symptoms may appear after intake of widely varying amounts. Clinical symptoms: for the more unfortunate even traces of lactose can unleash severe intestinal crises. Consequently, lactose should not be included in antispastic medicaments. The discovery of BSE (bovine spongiform encephalopathy) and its trans-missibility to man have recently brought about restrictive regulations as regards the use not only of lactose but also of bovine gelatin and derivates of tallow.

Overall, the current regulatory situation of having to wait to see if an excipient is "approved" by virtue of regulatory acceptance of the new drug formulation in which it is a constituent is not helpful or acceptable. Nor is the view that new excipients should be treated as new chemical entities and, by inference, undergo a full preclinical testing package. Many other areas of drug development have recently benefited from new or updated regulatory guidance, including advice on necessary preclinical investigations. Thus, specific guidance to assist companies in the development of their excipients is urgently needed and/or an excipient testing strategy would be an excellent topic for inclusion for International Conference on Harmonisation (ICH) consideration. Such guidance/discussion would complement the current advances in pharmacopoeial standardization of excipient quality. As a consequence, it may be possible to have excipients reviewed by a Committee of an International Pharmacopoeia with the safety data assessed by elected experts and published; these data would then be acceptable to international regulatory bodies.

**Table 12.18:** Toxicity of some commonly used excipients

| Excipient | Toxicological effects |
|---|---|
| Carbolated vaseline Parabens | Irritant, dermatitis Contact dermatitis, hypersensitivity |
| BHT, BHA | Animal toxicity data (lesions to hepatic cells) |
| PG | Methaemoglobinemia, sensitization, pruritus |
| Thiomersal | Hypersensitivity, mercurial toxicity |
| Freons | Cardiac effects at high dose, sudden death |
| Aluminium salts | Local tissue reaction at dose site in animals (dogs and guinea pig) |
| Talc | Lung tumor in female rats and adrenal gland neoplasms in rats |
| Chlorbutol | Drop pressure, CNS effects, somnolence |
| Halotane | Cardiac effects, bradycardia and arrhythmias in dogs |
| Sunset yellow Tartrazina | Utricaria exacerbation Headaches, GI disturbances, exacerbation of asthma, dangerous in aspirin intolrent individuals |
| Limonene | Hyaline droplet formation in male rat kidney |
| Low molecular weight PEG | Teratogenic in mouse with fetal loss, decreased body weight and malformations |
| Saccharine | Proliferative changes in male rat bladder epithelium |
| Mannitol | Induced renal and cardiac injury, apoptosis in kidney and heart |

For materials in which toxicity is a possible concern, formulators can gain information about the excipient's regulatory acceptance and allowable amount by consulting with excipient manufacturers and toxicology experts. This information also may be found in the *Food Chemicals Codex, Code of Federal Regulations* (*CFR*), *FDA Inactive Ingredients Guide* (FDA IIG), and

other references. In addition, 21 *CFR* parts 182 and 184 list generally regarded as safe (GRAS) food ingredients.

## EXCIPIENT SELECTION

The nature and properties of the active ingredient dictate the choice of an excipient, the dosage form to be elaborated, and the process by which it is manufactured. It is also important to know the patient group and clinical condition. The mode of use of the medication and the envisaged dose must also be considered. Candidate excipients should then be evaluated to demonstrate that they function in the manner intended (do what they are meant to do) and do not adversely interact with the drug, or with other excipients. Moreover an increased complexity of the formulation because of the number of used excipients increases the risk of error in manufacturing of the formulation that can result in product failure. Obviously, they should not have any pharmacological effect and should not otherwise compromise safety or tolerance. It is also necessary to consider the regulatory status of excipients and any country-specific requirements or constraints. Caution should be taken when consulting GRAS list as it refers to compounds that have been administered orally and specifically to food additives. The safety of excipients administered by other delivery routes may be quite different than those administered orally.

Magnesium and calcium salts are quite common excipients, in terms of both the range of formulations containing them and the level of inclusion in formulations. Interaction between tetracycline antibiotics and calcium and magnesium ions is well known. But how many of us would think about not including magnesium stearate in a formulation of medicine intended to be an adjunct therapy to treatment with a tetracycline? The point is that the design of a formulation cannot be undertaken in isolation, and the possibility of excipient interactions on final administration to the patient must be considered, not only with the formulation being developed, but also with other medicines administered concomitantly.

Excipient residues may also compromise safety, efficacy or tolerance. Residues in excipients can also affect quality and performance by interacting with the drug or other key components. Reducing sugar impurities in mannitol were responsible for the oxidative degradation of a cyclic heptapeptide. Monomers or metal catalysts used during a polymerization process are toxic and can also destabilize the drug product if present in trace amounts. Due to safety concerns, the limit of vinyl chloride (monomer) in

polyvinyl pyrrolidone is not more than 10 ppm, and for hydrazine (a side product of polymerization reaction) not more than 1 ppm.

Excipients manufactured by fermentation processes, such as dextrose, citric acid, mannitol, and trehalose, should be specially controlled for endotoxin levels. Mycotoxin (highly toxic metabolic products of certain fungi species) contamination of an excipient derived from natural material has not been specifically addressed by regulatory authorities. Wool fat or lanolin derived from sheep wool may contain low levels of insecticides from sheep treated for parasites. These insecticide levels are probably too low to cause direct toxicity, but may cause allergic reactions when lanolin in cosmetics or topical medicaments is applied to the skin.

Heavy metal contamination of excipients is a concern, especially for sugars, phosphate, and citrate. Several rules have been proposed or established. For example, the EP sets a limit of nmt 1 ppm of nickel in polyols. Similarly, the FDA has proposed a guideline that would limit the aluminum content for all LVPs used in TPN therapy to 25 mg/L. Furthermore, it requires that the maximum level of aluminum in SVPs intended to be added to LVPs and pharmacy bulk packages, at expiration date, be stated on the immediate container label.

**Table 12.19:** Excipient residues in some common excipients

| Excipient | Residue |
|---|---|
| Lactose | Aldehydes, reducing sugars |
| Polyethylene glycol Magnesium stearate | Aldehydes, peroxides, organic acids Antioxidants |
| Micorcrystalline cellulose | Lignin, hemicelluloses |
| Benzyl alcohol Fixed oils, paraffins PVP, Polysorbates Talc | Benzaldehyde Antioxidants Peroxides Heavy metals |
| Stearate lubricants | Alkaline residues |
| Hydroxyproply methyl/ethyl celluloses | Glyoxal |
| Starch | Formaldehyde |

## REGULATION OF EXCIPIENTS-FDA PERSPECTIVE

The US Food and Drug Administration (FDA) does not have a formal regulatory denition of “excipient”; however, recent guidance on nonclinical safety studies for excipients provides some indication of FDA acceptance of excipients as more than llers. The background section of the guidance states:

*In this guidance, the phrase new excipients means any inactive ingredients that are intentionally added to therapeutic and diagnostic products, but that (i) we believe are not intended to exert therapeutic effects at the intended dosage, although they may act to “improve product delivery” (e.g. enhance absorption or control release of the drug substance) and (ii) are not fully qualied by existing safety data with respect to the currently proposed level of exposure, duration of exposure, or route of administration. Examples of excipients include fillers, extenders, diluents, wetting agents, solvents, emulsiers, preservatives, flavours, absorption enhancers, sustained-release matrices, and colouring agents*.

FDA has published listings in the Code of Federal Regulations (CFR) for GRAS substances that are generally recognized as safe. Over the years, the agency also maintains a list entitled Inactive Ingredient Guide (IIG) for excipients that have been approved and incorporated in the marketed products and provides the database of allowed excipients with the maximum dosage level by route of administration or dosage form for each excipient. Both GRAS listings and IIG information can be used by industry as an aid in developing drug products. For new drug development purposes, once an inactive ingredient has appeared in an approved drug product for a particular route of administration, the inactive ingredient is not considered new and may require a less extensive review the next time it is included in a new drug product. For example, if a particular inactive ingredient has been approved in a certain dosage form with certain potency, a sponsor could consider it safe for use in a similar manner for a similar type of product. In general, nonclinical and clinical studies are required to demonstrate the safety of a new excipient before use. In this context, the US FDA has recently published a guidance document for industry on the conduct of nonclinical studies for the safety evaluation of new pharmaceutical excipients. This guidance not only provides the types of toxicity data to be used in determining whether a potential new excipient is safe, but also describes the safety evaluations for

excipients proposed for use in over-the-counter and generic drug products. The document also depicts testing strategies for pharmaceuticals proposed for short-term, intermediate, and long-term use. More importantly, this guidance highlights the importance of performing risk-benefit assessments on proposed new excipients in the drug products while establishing permissible and safe limits for the excipients. As illustrated, with proper planning, it is often possible to assess the toxicology of an excipient in a relatively efficient manner. Existing human data for some excipients can substitute for certain nonclinical safety data. In addition, an excipient with documented prior human exposure under circumstances relevant to the proposed use may not require evaluation with a full battery of toxicology studies.

There is no process or mechanism currently in place within the FDA to independently evaluate the safety of an excipient. Instead, for a drug or biological product subject to premarketing approval, their excipients are reviewed and approved as ‘components’ of the drug or biological product in the application. From a scientific standpoint, the regulatory process is appropriate since excipients play an integral part to the formulation and cannot be reviewed separately from the drug product. A number of proposals have been made over the years to attempt to provide some mechanism for independent review or approval of inactive ingredients, but such proposals have been largely unsuccessful for a variety of reasons.

## POLYMERS

Polymers have very large molecular weights made up of repeating units (or monomers) through-out their chains. Examples of their synthetic classication are condensation polymers and chain (addition) polymers. In condensation polymers, monomers consisting of two functional groups react with each other to form covalent bonds. For example, monomers with OH groups may react with monomers with −COOH groups, thus resulting in a polymer that has an ester linkage. As the polymerization proceeds, the number of repeating units, n, increases with time. The higher the degree of polymerization is, the higher the molecular weight will be. When monomers with bifunctional groups AA and BB are in exact proportion, the polymer molecules have infinite molecular weight.

–AA–BB–AA–BB–AA–BB–AA–BB

In chain polymerization, monomers containing an unsaturated bond polymerize in the presence of an initiator, which generates an active site at the end of the chain. Most pharmaceutical polymers are made via a free radical polymerization process. In copolymerization, polymers are formed with more than two monomers in the polymerization process. The properties of the copolymers are influenced by the sequence of monomers and monomer type. Random copolymers have the repeating units randomly arranged on the molecule chain. In alternating copolymers, two repeating monomer units are arranged alternatively and in an orderly fashion along the chain. Block copolymers have a chain with long sequences of each repeating unit along the main polymer chain. In grafted copolymers, the long sequences of one repeating unit are grafted onto the backbone of the other polymer (Fig. 12.2).

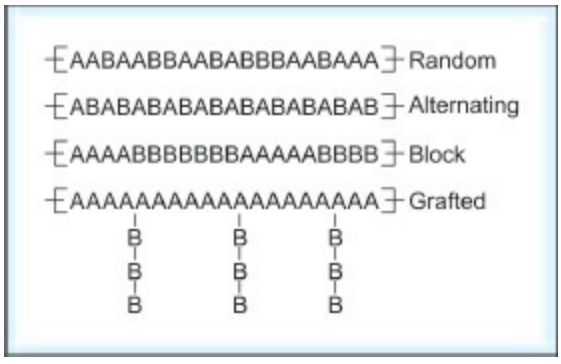



**Fig. 12.2:** Different types of copolymers

Polymers are classied according to their synthesis, mechanical behavior, processing characteristics, and morphology. In the medical and pharmaceutical fields, natural, semisynthetic, and synthetic polymers have made significant contributions to the improvement of human health. The functions of some of the polymers of natural, semisynthetic, and synthetic are origin tabulated in Table 12.20.

Polymers are capable of providing sustained release of an encapsulated drug, within its therapeutic window. This leads to reduced peaks and valleys typically associated with immediate release dosage forms. Typically natural polymers or their derivatives (such as cellulose and methyl cellulose) as well as synthetic nondegradable polymers [such as poly (vinyl pyrrolidone) and polymethacrylates] are used for oral controlled release (CR) applications. Due to this generic status of polymers, the emphasis in novel oral CR systems is geared more toward the mixing and matching of polymers and fabrication of the solid dosage form, rather than design of novel polymers. On the other hand, use of polymers for injectable and/or implantable CR systems is relatively new and more infrequent. In such situations, the polymer must not only permit CR of the drug, but also be biocompatible and nontoxic. Several drug delivery applications also require the polymer to be biodegradable—degrading into by-products that are safe and can be cleared from the body. Thus polymers serve as key excipients in oral and parenteral CR formulations.

**Table 12.20:** Classification of polymers used in drug delivery

| **Polymers** | **Major functions** |
|---|---|
| **Natural polymers** | |
| Gelatin | Binder, coacervation |
| Alginic acid, Na | Encapsulation |
| Xanthan gum, | Matrix, binder |
| Arabic gum | |
| Chitosan | Matrix, membrane |
| Methyl cellulose | Binder, coating |
| Ethyle cellulose | Matrix, coating |
| Hydroxyethyle cellulose | Binder, coating |

| | |
|---|---|
| Hydroxypropyle cellulose | Binder, coating |
| Hydroxyethyl-methyl cellulose | Binder, coating |
| Hydroxypropyl methyl cellulose | Matrix, coating |
| Carboxymethyl cellulose sodium | Binder, disintegrant |
| Cellulose acetate | Membrane |
| Cellulose acetate butyrate | Membrane |
| Cellulose acetate propionate | Membrane |
| Cellulose acetate phthalate | Enteric |
| Hydroxypropyl methyl cellulose phthalate | Enteric |
| **Synthetic polymers** | |
| Polyacrylic acid (carbopol) | Matrix, bioadhesive |
| Poly (MMA/MAA) | Enteric |
| Poly (MMA/DEAMA) | Matrix, membrane |
| Poly (MMA/EA) | Membrane |
| Poly(vinyl acetate phthalate) | Enteric |
| Poly(vinyl alcohol) | Matrix |
| Poly(vinylpyrrolidone) | Binder |
| Poly(lactic acid) | Biodegradable |
| Poly(glycolic acid) | Biodegradable |
| Poly(lactic/glycolic acid) | Biodegradable |
| Polyethylene glycol | Binder |
| Polyethylene oxide | Matrix, binder |
| Poly(dimethyl silicone) | Matrix, membrane |
| Poly(hydroxyethyl methacrylate) | Matrix, membrane |
| Poly(ethylene/vinyl acetate) | Matrix, membrane |
| Poly(ethylene/vinyl alcohol) | Matrix, membrane |
| Polybutadiene | Adhesive/matrix |
| Poly(anhydride) | Bioerodible |

| | |
|---|---|
| Poly(orthoester) | Biodegradable |
| Poly(glutamic acid) | Biodegradable |

## PHARMACEUTICAL POLYMERS

An ideal polymer for pharmaceutical applications should serve the following requirements:

1. It must be biocompatible and degradable (i.e. it should degrade in vivo into smaller fragments, which can then be excreted from the body).
2. The degradation products should be nontoxic and should not create an inflammatory response.
3. Degradation should occur within a reasonable period of time as required by the application (this may vary from days to months).
4. Based on the needs of certain application, the polymer should demonstrate versatile mechanical properties [e.g. stent coatings require polymers to be elastomeric and microsphere processing require them to have high glass transition temperature ($T_g$)].

No single polymer can match all of the above criteria. This has led companies to develop application-specific polymers and/or series of polymers that may have the structure-property variability to encompass all potential applications. Hence novel polymers that do not have prior history in biological applications undergo extensive regulatory scrutiny. As with any novel excipient, signicant physicochemical and biological testing is required to prove the safety of the polymer. This section describes some of the leading polymers that are being developed toward commercialization (Fig. 12.3).

## Poly (α-hydroxy acid)s

The poly (α-hydroxy acid) s series of polymers includes PLA and PLGA as well as other polymers such as polycaprolactone and poly (butyric acid) (Fig. 12.3). Most of the drug delivery research has been carried out using PLA and PLGA polymers due to the fact that they can have degradation times in the less than one-year time frame. Because the lactide monomer possesses two chiral carbons, polylactide can be obtained using D-lactide (the D-,D-cyclic dimer), L-lactide (the L-,L-cyclic dimer), and mesolactide (the D-, L-cyclic dimer) or DL-lactide (a racemic mixture of D-and L-lactide). Most of the PLLA applications involve long-term therapies in tissue engineering. Poly (DL-lactide) on the other hand is an amorphous polymer that is brittle and leads to degradation time frames that are suitable for drug delivery applications. Adding increasing proportions of glycolide into PLA lowers $T_g$ and generally increases polymer hydrophilicity. PLGA copolymers generally remain amorphous as long as the glycolide content remains within the range of about 0% to 70% (molar fraction). In contrast, poly (L-lactide-co-glycolide) is amorphous when the glycolide content is 25% to 70%. The most rapid degradation rate (i.e. two months) is observed in PLGA copolymers containing 50% glycolide. The degradation products of PLGA polymers, lactic acid and glycolic acid, are components of the Kreb's cycle and hence are well tolerated by the body. Furthermore, the degradation rates of PLGA copolymer can be varied from two weeks to greater than a year, simply by adjusting the ratio of lactide to glycolide. Glycolide homopolymers are highly crystalline and poorly soluble in most organic solvents. On the other hand, lactide homopolymers could be crystalline in case of L-lactide or amorphous for DL-lactide. Copolymerization of DL-lactide with glycolide monomer leads to a polymer that is amorphous as well as possesses increased hydrophilicity as compared to the DL-lactide homopolymer. The PLGA copolymer also possesses more chain defects due to the randomness of copolymerization. Due to these aspects, PLGA 50:50 (50% DL-lactide and 50% glycolide) is the fastest degrading polymer in the series, with degradation time in the order of week. On the contrary, polymers made from L-lactide demonstrate degradation times that are in the order of years. This wide acids, the side-chain methyl group is replaced by other alkyl groups range of degradation kinetics facilitates the use of these polymers for a variety of drug delivery applications. Polyesters typically allow water

penetration and hence undergo bulk erosion (degradation in the bulk of the matrix). This may be viewed as a limitation in certain applications, where maintenance of the matrix shape is desired, as in the case of medical device applications. For other poly(hydroxyl), besides lactic acid-and glycolic acid-based polyesters, polymers based on other ester-based cyclic monomers such as trimethylene carbonate and e-caprolactone have also been tested for applications in drug delivery (Table 12.21).

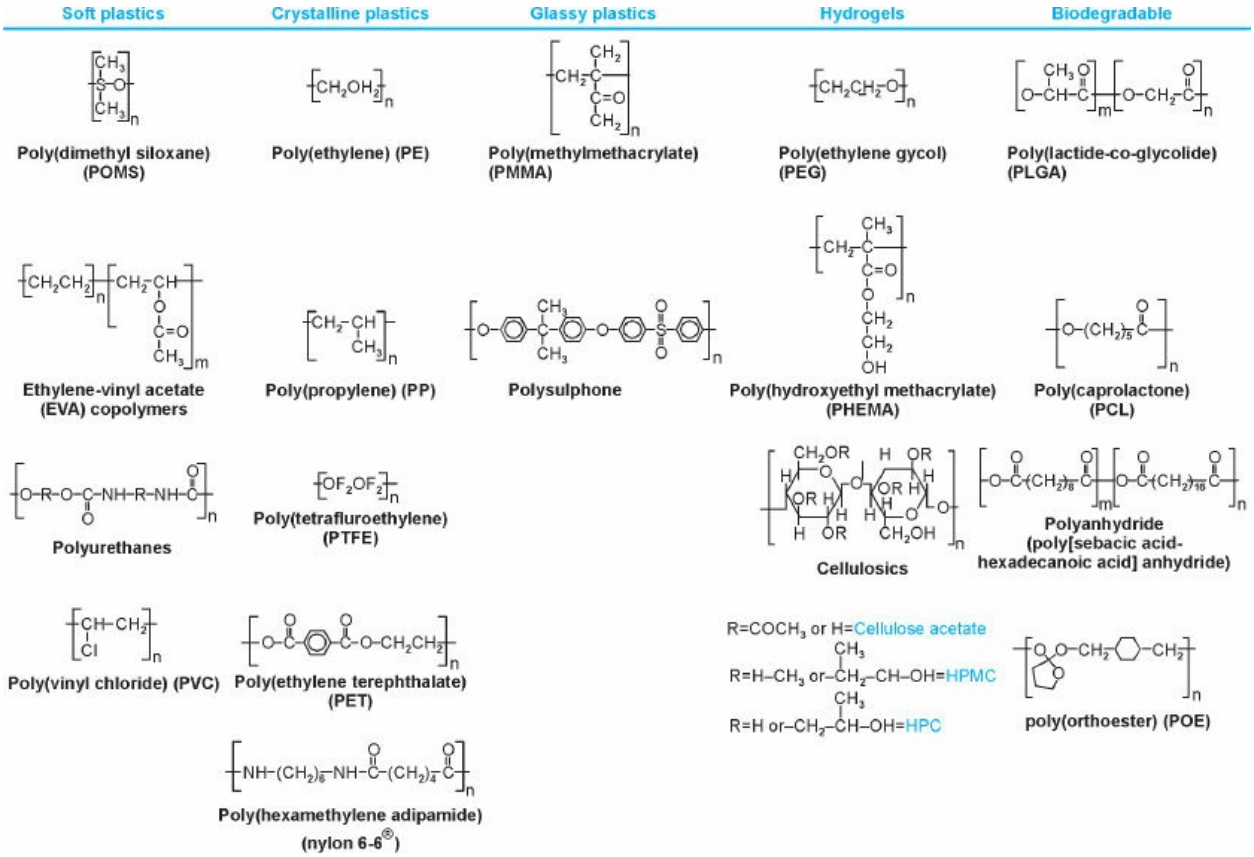



**Fig. 12.3:** Structure of various pharmaceutical polymers

**Table 12.21:** List of marketed drug delivery products utilizing polylactic acid or poly (lactic-co-glycolic) acid polymers

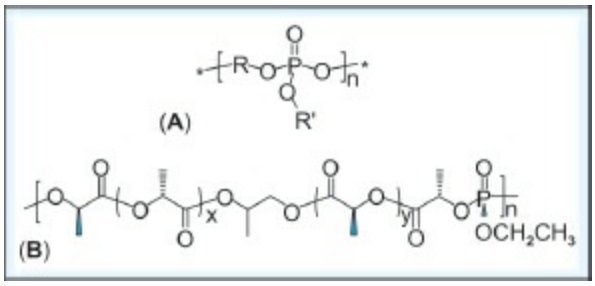
| Product name | Active ingredient | Company | Indication |
|---|---|---|---|
| Decapeptyl® | Triptorelin | Debiopharm | LHRH agonist |
| Trelstar™ Depot | Triptorelin pamoate | Pfizer | Prostate cancer |
| Sandostatin LAR® | Octreotide acetate Depot | Novartis | Growth hormone suppression |
| Eligard® | Leuprolide acetate | Sanofi-Synthelabo | Prostate cancer |
| Lupron Depot® | Leuprolide acetate | Takeda-Abott | Prostate cancer |
| Suprecur® MP | Buserelin acetate | Aventis | Prostate cancer |
| Profect® Depot | Buserelin acetate | Aventis | Prostate cancer |
| Zoladex® | Goserelin acetate | Astrazeneca | Prostate cancer |
| Posilac® | Recombinant bovine somatropin | Monsanto | To increase milk production in cattle |
| Nitropin Depot® | Recombinant human growth hormone | Genentech-Alkermes | Growth hormone deficiency |
| Somatuline® LA | Lanreotide | Ipsen | Acromegaly |

## Polyanhydrides

Polyanhydrides were first developed by Carothers and coworkers in the early $20^{th}$ century for applications in the textile industry. The interest in these polymers waned soon thereafter because of their instability. However, poor hydrolytic stability made these polymers an attractive candidate for drug delivery applications. Presently two series of polyanhydrides have undergone active clinical development. The first series involve sebacic acid (SA, hydrophilic component) and carboxyphenoxypropane (CPP, hydrophobic component). By varying the ratio of SA to CPP, the degradation rates of the copolymer could be significantly varied. The commercially developed polymer in this series is *P* (CPP: SA 20:80), developed by Guilford Pharmaceuticals as a matrix for Gliadel, a commercial product for the sustained release of carmustine, for the treatment of brain tumors. A second family of polyanhydrides that has also been developed for applications in drug delivery is the copolymer of fatty acid dimer and SA. In this case, the fatty acid dimer acts as the more hydrophobic, slower degrading moiety. A poly (erucic acid dimer: SA 50:50) copolymer has been used for the sustained release of gentamicin, an antibiotic used for the treatment of osteomyelitis. The property that makes polyanhydrides unique is their surface hydrophobicity. Due to this high hydrophobicity, polyanhydride matrices do not facilitate water absorption. Consequently, hydrolytic degradation is restricted to the surface a property that is termed as surface erosion. This type of degradation allows for zero-order release of drugs, i.e. the drug release profile is independent of the residual concentration of the drug in the matrix.

## Polyphosphoesters

Polyphosphoesters are a novel class of polymers being studied for applications in drug delivery and tissue engineering (Fig. 12.4). The individual components can be modified to obtain polymers with a wide range of properties and useful morphologies including injectable gels, elastomeric films, and amorphous solids. The phosphate groups impart several unique characteristics to this polymer series. It makes the polymer more soluble in common organic solvents. It also acts as an internal plasticizer, making the polymer more flexible. Finally the phosphate groups impart hydrophilicity to the polymer, thus giving the surface lower fouling characteristics via reduced protein adsorption.




**Figs 12.4A and B:** (A) General structure of polyphosphoesters. R and R can be varied to obtain polymers with varied physico-chemical properties; (B) Structure of polylactide-co-ethylphosphate copolymer

## Polyphosphazenes

As seen in (Fig. 12.5), polyphosphazenes contain alternating phosphorus-nitrogen double and single bonds and side-chain functionalities that can be varied to obtain various series of polymers. Polyphosphazenes are synthesized by the reaction of poly (dichloro phosphazene) with organic nucleophiles such as alkoxides, aryloxides, or amines. The side-chain functionalities can be modified to obtain a wide range of properties including water solubility and degradability. Water-soluble poly phosphazenes have attracted special attention due to the possibility of encapsulating biopharmaceutical drugs via a completely aqueous process. Water-soluble polyphosphazenes gel in calcium chloride solutions enabling microencapsulation via a completely aqueous process. Such aqueous processing of polyphosphazenes allows their utility for the encapsulation of sensitive drugs such as proteins and vaccines.

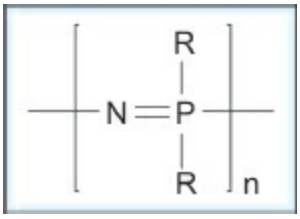



**Fig. 12.5:** General structure of polyphosphazenes

## Poly(orthoester)s

Poly(orthoester)s (POEs) developed in the 1970s were the first series of polymers developed specifically for applications in CR technologies. However, due to the high hydrophobicity of POEs, the early generations demonstrated very slow degradation rates. To overcome the slow degradation, encapsulation of acidic entities, which could catalyze the degradation process, was proposed. However, such acidic components would leach out, and hence this concept did not meet significant success. The POE-IV was created with latent acid groups (such as lactic or glycolic acid) present on the polymer backbone. As the lactic/glycolic acid segments degrade, more acidic functionalities are generated, which autocatalyze the overall degradation of the polymer. The degradation rates could be varied by changing the amount of the internal autocatalysis moiety on the polymer backbone. POEs have also demonstrated a variety of morphological characteristics. For example POE with cyclohexyl-dimethanol as a monomer unit is present at room temperature as a viscous paste rather than an amorphous solid.

## Pseudopoly(amino acid)s

Pseudopoly(amino acid)s, have been investigated as implantable, degradable materials for medical applications. Inherent to all pseudopoly(amino acid)s are nonamide bonds incorporated into the backbone of amino acid-derived polymers. Contrary to conventional poly(amino acid)s, pseudopoly(amino acid)s are readily soluble in organic solvents and processable by conventional melt-fabrication techniques. The ease of fabricating pseudopoly(amino acid)s into fibers, films, rods, microspheres, porous foams, or other configurations appropriate for medical devices is one of their major advantages over conventional poly(amino acid)s. Thus, these materials combine the inherent nontoxicity and biocompatibility of amino acids with outstanding material properties usually only found in industrial engineering plastics. Particularly noteworthy are the tyrosine-derived polycarbonates, a family of polymers based on alkyl esters of desaminotyrosyl-tyrosine. The lead polymer in this family is poly[desamino-tyrosyl-tyrosine ethyl ester (DTE) carbonate], a polymer derived from desaminotyrosyl-tyrosine ethyl ester. Other polymers in this series of tyrosine-derived polycarbonates are poly[desaminotyrosyl-tyrosine butyl ester (DTB) carbonate], poly[desaminotyrosyl-tyrosine hexyl ester (DTH) carbonate], and poly [desaminotyrosyl-tyrosine octyl ester (DTO) carbonate], where the letters B, H, and O indicate the presence of butyl, hexyl, or octyl ester pendent chains, respectively. Controlled intracranial release of dopamine, a drug for the treatment of Parkinson’s disease, from a poly (DTH carbonate) matrix was one of the first medical applications investigated for tyrosine-derived polycarbonates. Poly (DTH carbonate) has a relatively low processing temperature and its backbone is structurally related to dopamine, which seemed to improve dopamine incorporation into the polymer. Although this release rate is within the therapeutically useful range, no sub-sequent studies of this release system in vivo were reported. Tyrosine-derived polyarylates represent the first combinatorially designed library of biomaterials. Overall, 112 individual polymers were recently synthesized in such a way that the polymers exhibited continuous, incremental gradients of chemical, material, and biological properties. Tyrosine-derived polyarylates are strictly alternating copolymers of a diacid component and a diphenol component. Compared to tyrosine-derived polycarbonates, the polyarylates degrade faster, are more flexible, and encompass a wider range of physicomechanical and biological

properties. The major advantage of their combinatorial design is that structure-property correlations established at the outset may be effectively used to help tailor material properties to specific applications.

## POLYMERS FOR ORAL DRUG DELIVERY

Typically, oral controlled release (CR) formulations utilize natural or synthetic polymers that provide CR of the therapeutic agent via diffusion and/or erosion mechanism. Generally, the polymers used in oral CR systems are nonabsorbable biopolymers such as cellulose derivatives or hydrophilic gums. Alternatively, the polymer could be insoluble, and the drug release is governed by diffusion or osmosis. The diffusion could be Fickian or zero-order depending on the type of polymers and excipients used. In reality, except for osmotically driven systems, most oral CR formulations utilize a combination of dissolution and diffusion. Although the drug transport is via diffusion, the rate-controlling step for drug release is the dissolution of the polymer. An example of this concept is the TIMERx® (Penwest Pharmaceuticals) formulation used for CR of nifedipine. This formulation consists of a blend of locust bean gum and xanthan gum, which form aggregates via noncovalent interactions. The composition of these two polymers can be varied to obtain the desired drug release profile. The SODAS (Elan) formulation consists of the drug covered by layers of water-soluble and insoluble polymers in 1 to 2 mm beads. The drug is released via diffusion, wherein the rate of diffusion is controlled by the types of polymers and excipients used. Examples of commercial products utilizing this technology are Ritalin LA®, a sustained release formulation of methylphenidate hydrochloride, and Cardizem SR® that encapsulates diltiazem as the active ingredient.

Polymers can also be utilized to obtain targeted delivery into certain regions of the GI tract. Targeting to the GI tract can be achieved by several approaches. Enteric coating polymers are the simplest of examples, wherein the polymer dissolves in a specific portion of the GI tract (as a function of pH) resulting in targeted delivery. However, these formulations are typically pH dependant and thus depend on the fed or fasted state of the patient. More sophisticated pH-independent systems are also being developed. For example, drugs used for the treatment of ulcerative colitis, a form of inflammatory bowel disease, which affects the colon, are most effectively delivered directly to the colon. A novel formulation of prednisolone consisting of a combination of ethyl cellulose and a starch derivative to encapsulate the drug is being tested in the clinic. The polymer coating is

susceptible to degradation specifically by enzymes present in the colon, thus leading to colon-targeted delivery.

Azopolymers are examples of excipients, which are also useful for colon targeting. The drug is attached to these polymers via an azobond, which is degraded specifically by enzymes in the colon, thus leading to targeted delivery. Another approach to targeted delivery is delaying the release of the drug until it reaches the region of interest within the GI tract. This concept has been employed in a commercially available formulation of aminosalicylate for colon-targeted delivery. The most successful formulation utilizing the delayed release technology is AstraZeneca's Prilosec, which is a delayed-release formulation of omeprazole. Another novel use of polymers in oral drug delivery is for gastric retention formulations, utilizing superporous hydrogels. Such hydrogels are prepared by the addition of sodium bicarbonate during polymerization reaction. This leads to the generation of carbon dioxide, which leads to bubbles that escape leaving a porous structure. These formulations swell rapidly (within 20 minutes) to a size that is bigger than the pylorus in the GI tract. This swelling prevents the oral solid dosage form from being removed from the GI tract during a turnover cycle. Poly (vinyl pyrrolidone) is one of the polymers being explored for this application. Other hydrogels possessing bioadhesive and muco-adhesive properties also find applications in oral drug delivery due to their ability to increase residence time of drugs in the GI tract. List of polymers commonly used as excipients in oral drug delivery formulations is shown in Table 12.22.

**Table 12.22:** List of polymer-based technologies used for oral drug delivery

| Technology | Company | Polymers | Approach | Product example |
|---|---|---|---|---|
| SODAS | Elan | Ammoniomethacrylate copolymers, povidone | Combination of immediate and sustained release | Avinza |
| Andrx proprietary technology | Andrx | Candellila wax, Methacrylic acid copolymer, hypromellose | Extended release | Altoprev |
| Gastric retention | Depomed | Povidone, PEO | Extended release | Proquin XR |
| TIMERx | Penwest Pharmaceuticals | Xanthum and Locust gum | Extended release | Procardia XL |
| Microtrol | Shire Labs | Povidone, PEG, microcrystalline celllulose | Extended release/ Immediate release | Carbatrol |
| Contramid | Labopharm | High-amylase starch | Sustained release | Tramadol |
| Softgel | Cardinal | Gelatin | Microemulsion | Neoral |
| Diffucaps | Eurand | Povidone, HPMC, ethyl cellulose | Combination of immediate and sustained release | Metadate CD |
| Ion exchange | Celltech | Sulfonated styrene-divinylbenzene copolymer | Controlled release | Delsym |

## *Biodegradable Polymers*

Polylactic (PLA) and its copolymers with glycolide (PLGA) have received wide attention in recent years as excipients for sustained release of parenteral drugs. The use of biodegradable polymers, especially polylactic acid (PLA), in oral solid dosage forms has been reported in the literature. PLA has been used as a matrix for phenobarbital tablets. Similarly, the use of polylactide as a matrix for oral dosage form of naproxen has also been reported. A novel use of biodegrabale polymers in oral drug delivery was demonstrated by Mathiowitz and coworkers, who showed that fumaric acid-based polyanhydrides had bioadhesive properties useful for increased gastroretention. Poly (fumaric anhydride) is a rapidly degrading polymer that degrades to a component of the Kreb's cycle, fumaric acid. The rapid surface-erosion type degradation leads to a high concentration of carboxyl groups on the surface of the dosage form. It is hypothesized that these highly acidic groups enhance bioadhesiveness of the matrix, thus increasing its residence time in the GI tract.

## POLYMERS FOR PARENTERAL DRUG DELIVERY

Early development of polymers in injectable drug delivery primarily involved PLA and poly (lactic-co-glycolic) acid (PLGA) due to the prior use of these polymers in biomedical applications as sutures. Besides the safe and biocompatible nature of these polymers, their ease of availability made them ideal first candidates for screening parenteral CR formulations. Some of the early biodegradable polymer-based products for injectable sustained release used these polymers. However because the field has expanded to other areas such as medical devices, interest in novel polymers with unique functionalities has increased. Polymers for parenteral drug administration can be used as matrices that are injected or implanted. Ease of administration requires the polymer to be in an injectable form, such as microspheres, nanoparticles, or reversible gels.

## POLYMERS FOR IMPLANTABLE DRUG DELIVERY

While nonsurgical injections are desirable from a patient compliance standpoint, implantable drug delivery systems can be utilized especially when surgery is inevitable. Nondegradable polymers are stable and hence do not lead to pH changes in the formulation microenvironment (as opposed to degradable polymers where polymer degradation leads to drop in pH due to formation of acidic fragments). Another advantage of nondegradable polymers is that formulations based on such systems typically have drug release exclusively by single mechanisms (diffusion and osmosis). Consequently drug release profiles can be accurately predicted and modeled. Furthermore, by fabricating the polymer implant to specific geometry, the release profile can be tailored. For example, it was shown that hemispherical implants made out of nondegradable polymers provide a zero-order drug release. On the other hand, the nondegradability of polymers limits their use to applications involving extended release or applications where the presence of an external entity is not considered a major issue. A commercial example of very long-term sustained release via nondegradable polymer implants is the Norplant implant, designed for sustained delivery of levonorgestrel for up to five years. Nondegradable polymers are also useful as matrices for ocular implants. This application requires the polymer to be hydrophilic, to minimize local tissue irritation. Need for ocular implants stems from the challenges posed to conventional ocular medicines (i.e. eye drops) such as rapid dilution, tear washout, poor patient compliance, and limited bioavailability. Ocular implants from hydrophilic polymer matrices that provide localized sustained release may overcome the above limitations. The first polymeric sustained release product to reach the market was Ocusert®, a pilocarpin sustained release ocular implant developed by Alza. Ocusert® has the drug reservoir as a thin disc of pilocarpine-alginate complex sandwiched between two transparent discs of microporous membrane fabricated from ethylene-vinyl, acetate copolymer. The microporous membranes permit the tear fluid to penetrate into the drug reservoir compartment to dissolve pilocarpine from the complex. Pilocarpine molecules are then released at a constant rate of 20 or 40 mg/hr for a four-to seven-day management of glaucoma.

## Sol to Gel Polymers

Certain types of polymers possess an ability to transform from flowable liquid form to viscous gels as a function of increasing temperature. Such polymers typically possess an alternating block copolymer structure, AB or ABA type, wherein A and B represent hydrophobic and hydrophilic segments within the polymer backbone. On increasing the temperature the polymers undergo micellization, leading to phase transition from a flowable liquid state to a gel state. For instance, copolymers of polyethylene glycol (PEG) and PLGA have been found to possess thermoreversible properties. The copolymer formulations are soluble in aqueous systems and form a flowable solution under cold conditions. As the temperature is raised to approximately 37°C, the polymer solution undergoes rapid gelation to form a semi-solid depot. Another approach to obtain injectable gels is using concepts of solvent-induced gelation. Common biodegradable polymers such as PLA and PLGA are soluble in a range of pharmaceutical grade organic solvents. When an organic solution of such polymers is injected, the organic solvent diffuses away, and the water-insoluble polymer precipitates to form a semisolid gel. This concept has been applied in a product Eligard®, used for sustained release of leuprolide acetate. The product utilizes PLGA polymer dissolved in N-methyl pyrrolidone as the matrix. By modifying the ratio of solvent: drug: polymer, a one-month, three-month, and four-month sustained release leuprolide acetate products have been developed.

## POLYMER SELECTION

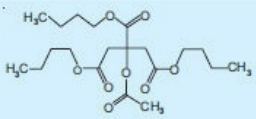
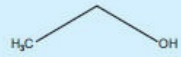
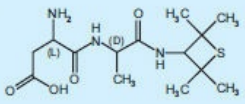
There are a variety of polymer properties or attributes to be considered when selecting a polymer. If an application requires rapid development and commercialization, then the polymer selection will most likely be made from among those polyesters that have already received regulatory approval. Another factor to consider is whether to use homopolymers consisting of a single monomeric repeat unit or copolymers containing multiple monomer species. If copolymers are to be employed, then the relative ratio of the different monomers may be manipulated to change polymer properties. Polymer composition would directly dictate many of the polymer physicochemical properties including bulk hydrophilicity, morphology, structure, and the extent of drug-polymer interactions. Ultimately, these properties will all influence the performance of the drug delivery system via changes to the relative rates of mass transport (e.g. water in and drug out of the system) and the degradation rate of both the polymer and the device.

The thermal attributes of the polymer, described as the glass transition temperature (Tg) and melting temperature (Tm), can also affect the mass transport rates through the polymer as well as the polymer processing characteristics and the stability of the dosage form. Below the glass transition temperature, the polymer will exist in an amorphous, glassy state. When exposed to temperatures above Tg, the polymer will experience an increase in free volume that permits greater local segmental chain mobility along the polymer backbone. Consequently, mass transport through the polymer is faster at temperatures above Tg. Often, polymer processing, such as extrusion or high-shear mixing, is performed above Tg. On the other hand, the greatest stability during storage of a polymer device may be obtained at temperatures below Tg, where solute diffusion is much slower and more subtle changes in polymer properties (e.g. tackiness) are reduced.

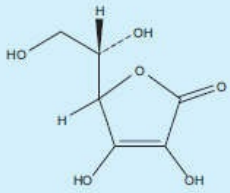
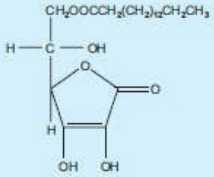
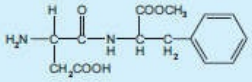
The presence of plasticizers, such as residual solvent or dissolved solutes including the drug or other additives, will tend to lower polymer Tg. Conversely, features that hinder segmental motion along the polymer, such as greater chain rigidity, bulky side groups, and ring structures, would tend to increase Tg. Finally, the presence of charged groups on a polymer can also influence drug release from the device. The number and density of ionized groups along the polymer backbone, on the side-chain groups, or at the

terminal end-groups of the polymer chains can all vary the extent of polymer-polymer and polymer-drug interactions. In this manner, ionizable groups can affect drug solubility in the polymer and, correspondingly, the release rate from the polymer. A number of reviews which describe in detail of the relationship between polymer properties and performance in drug delivery applications have been published.

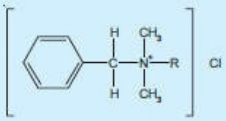
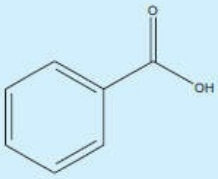
**Some commonly used excipients in pharmaceutical formulations (Appendix I)**

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| **Acetyltributyl citrate** | 1,2,3-Propanetri-carboxylic acid, 2-acetyloxy, tributyl ester [77-90-7] |  | Included in FDA IIG (oral capsules and tablets). Included in nonparenteral medicines licensed in the UK. Approved in the USA for direct food contact in food films | Acetyltributyl citrate is used to plasticize polymers in formulated pharmaceutical coatings | Acetyltributyl citrate is a clear, odorless, practically colorless, oily liquid |
| **Alcohol** | Ethanol [64-17-5] |  | Included in the FDA IIG (dental preparations; inhalations; IM, IV, and SC injections; nasal and ophthalmic preparations; oral capsules, solutions, suspensions, syrups, and tablets; rectal, topical, and transdermal preparations) | Antimicrobial preservative; disinfectant; skin penetrant; solvent | Alcohol is a clear, colorless, mobile, and volatile liquid with a slight, characteristic odor and burning taste |
| **Alitame** | L-α-Aspartyl-*N*-(2, 2,4,4-tetramethyl-3 -thietanyl)-D-alaninamide anhydrous [80863-62-3]L-α-Aspartyl-*N*-(2,2,4,4tetramethyl-3-thietanyl)-D-alaninamide hydrate [99016-42-9] |  | Alitame is an intense sweetening agent | Alitame is approved for use in food applications in a number of countries world-wideincluding Australia, Chile, China, Mexico, and New Zealand | Alitame is a white nonhygroscopic crystalline powder; odorless or having a slight characteristic odor |

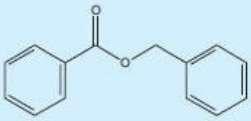
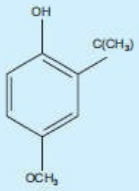
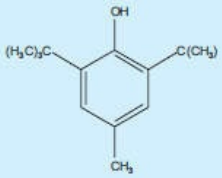
## Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| **Aluminum hydroxide adjuvant** | Aluminum oxyhydroxide [21645-51-2] | AlO(OH) octahedral sheets, wherehydroxyl groups are exposed at the surface | GRAS listed. Accepted for use in human and veterinary parenteral vaccines in Europe and the USA. The limits for use in human vaccines are 0.85 mg aluminum/ dose (FDA) and1.25 mg aluminum/dose (WHO) | Aluminum hydroxide adjuvant is used in parenteral human and veterinary vaccines | Aluminum hydroxide adjuvant is a white hydrogel that sediments slowly and forms a clear supernatant |
| **Ascorbic acid** | L-(+)-Ascorbic acid [50-81-7] |  | GRAS listed. Accepted for use as a food additive in Europe. Included in the FDA IIG (inhalations, injections, oral capsules, suspensions, tablets, topicals and suppositories). Included in medicines licensed in the UK | Antioxidant; therapeutic agent | Ascorbic acid occurs as a white to light-yellow-colored, nonhygroscopic, odorless, crystalline powder |
| **Ascorbyl palmitate** | L-Ascorbic acid 6-hexadecanoate [137-66-6] |  | GRAS listed. Accepted for use as a food additive in Europe. Included in the FDA IIG (oral, rectal, topical preparations) | Antioxidant | Ascorbyl palmitate is a practically odorless, white to yellowish powder |
| **Aspartame** | *N*-α-Aspartyl-L-phenylalanine 1-methyl ester [22839-47-0] |  | GRAS listed. Accepted for use as a food additive in Europe. Included in the FDA IIG (oral powder for reconstitution, buccal patch, granules, film-coated tablets) | Sweetening agent | Aspartame occurs as an off white, almost odorless crystalline powder with an intensely sweet taste |

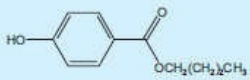
## Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| **Bentonite** | Bentonite [1302-78-9] | The USPNF 23 describes bentonite, purified benonite, and bentonite magma in three separate monographs. Bentonite is described as a native, colloidal, hydrated aluminum silicate; and purified bentonite is described as a colloidal montmorillonite that has been processed to remove grit and nonswellable ore compounds | Accepted in Europe as a food additive in certain applications. Included in the FDA IIG (oral capsules, tablets and suspensions, topical suspensions, controlled release transdermal films and vaginal suppositories) | Adsorbent; stabilizing agent; suspending agent; viscosity-increasing agent | Bentonite is a crystalline, claylike available as an odorless, pale buff, orcream to grayish-colored fine powder, which is free from grit |
| **Benzalkonium chloride** | Alkyldimethyl (phenylmethyl) ammonium chloride [8001-54-5] |  | Included in the FDA IIG (inhalations, IM injections, nasal, ophthalmic, otic, and topical preparations) | Antimicrobial preservative; antiseptic; disinfectant; solubilizing agent; wetting agent | Benzalkonium chloride occurs as a white or yellowish-white amorphous powder, a thick gel,or gelatinous flakes |
| **Benzoic acid** | Benzoic acid [65-85-0] |  | GRAS listed. Accepted as a food additive in Europe. Included in the FDA IIG (IM and IV injections, irrigation solutions, oral solutions, suspensions, syrups and tablets, rectal, topical, and vaginal preparations) | Antimicrobial preservative; therapeutic agent | Benzoic acid occurs as feathery, light, white or colorless crystals or powder |

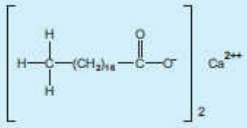
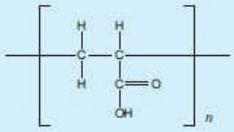
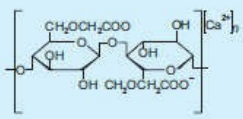
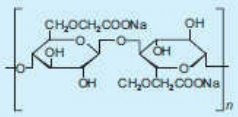
# Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| **Benzyl benzoate** | Benzoic acid phenylmethyl ester [120-51-4] |  | Included in the FDA IIG (IM injections and oral capsules). Included, as an active ingredient, in nonparenteral medicines licensed in the UK | Plasticizer; solubilizing agent; solvent; therapeutic agent | Benzyl benzoate is a clear, colorless, oily liquid with a slightly aromatic odor. It produces a sharp, burning sensation on the tongue |
| **Boric acid** | Orthoboric acid [10043-35-3] Metaboric acid [13460-50-9] | $H_3BO_3$ | Accepted for use as a food additive in Europe. Included in the FDA IIG (IV injections; ophthalmic preparations; otic solutions; topical preparations) | Antimicrobial preservative | Boric acid occurs as a hygroscopic, white crystalline powder, colorless shiny plates, or white crystals |
| **Butylated hydroxy-anisole** | 2-*tert*-Butyl-4-methoxyphenol [25013-16-5] |  | GRAS listed. Accepted as a food additive in Europe. Included in the FDA IIG (IM and IV injections, nasal sprays, oral capsules and tablets, and sublingual, rectal, topical, and vaginal preparations) | Antioxidant | Butylated hydroxyanisole occurs as a white or almost white crystalline powder or a yellowish-white waxy solid with a faint, characteristic aromatic odor |
| **Butylated Hydroxy-toluene** | 2,6-Di-*tert*-butyl-4-methylphenol [128-37-0] |  | GRAS listed. Accepted as a food additive in Europe. Included in the FDA IIG (IM and IV injections, nasal sprays, oral capsules and tablets, rectal, topical, and vaginal preparations) | Antioxidant | Butylated hydroxytoluene occurs as a white or pale yellow crystalline solid or powder with a faint characteristic odor |

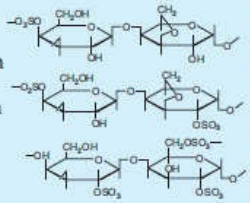
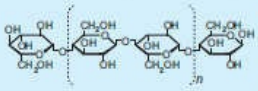
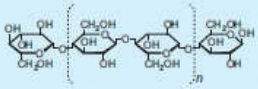
# Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| **Butylparaben** | Butyl-4-hydroxybenzoate [94-26-8] |  | Included in the FDA IIG (injections, oral capsules, solutions, suspensions, syrups and tablets, rectal, and topical preparations) | Antimicrobial preservative | Butylparaben occurs as colorless crystals or a white, crystalline, odorless or almost odorless, tasteless powder |
| **Calcium alginate** | Calcium alginate [9005-35-0] | $[(C_6H_7O_6)_2Ca]n$ | GRAS listed. Accepted for use as a food additive in Europe. Included in the FDA IIG (oral tablets). Included in non-parenteral medicines licensed in the UK | Emulsifiying agent; stabilizing agent; tablet disintegrant; thickener | Calcium alginate is an odorless or almost odorless, tasteless, white to pale yellowish-brown powder or fibers |
| **Calcium carbonate** | Carbonic acid, calcium salt (1 : 1) [471-34-1] | $CaCO_3$ | GRAS listed. Accepted for use as a food additive in Europe. Included in the FDA IIG (oral capsules and tablets; otic solutions) | Buffering agent; coating agent; opacifier; tablet and capsule diluent; therapeutic agent | Calcium carbonate occurs as an odorless and tasteless white powder or crystals |
| **Calcium phosphate, dibasic anhydrous** | Dibasic calcium phosphate [7757-93-9] | $CaHPO_4$ | GRAS listed. Accepted as a food additive in Europe. Included in the FDA IIG (oral capsules and tablets) | Tablet and capsule diluent | Anhydrous dibasic calcium phosphate is a white, odorless, tasteless powder or crystalline solid. It occurs as triclinic crystals |
| **Calcium phosphate, dibasic dihydrate** | Dibasic calcium phosphate dihydrate [7789-77-7] | $CaHPO_4 \cdot 2H_2O$ | GRAS listed. Accepted as a food additive in Europe. Included in the FDA IIG (oral capsules and tablets) | Tablet and capsule diluent | Dibasic calcium phosphate dihydrate is a white, odorless, tasteless powder or crystalline solid. It occurs as monoclinic crystals |

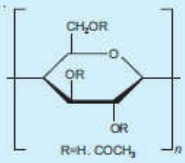
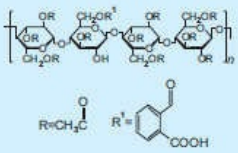
## Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| **Calcium Stearate** | Octadecanoic acid calcium salt [1592-23-0] |  | GRAS listed. Included in the FDA IIG (oral capsules and tablets). Included in non-parenteral medicines licensed in the UK | Tablet and capsule lubricant | Calcium stearate occurs as a fine, white to yellowish-white, bulky powder having a slight, characteristic odor. It is unctuous and free from grittiness |
| **Carbomer** | Carbomer [9003-01-4] |  | Included in the FDA IIG (oral suspensions, tablets; ophthalmic, rectal, and topical preparations, transdermal preparations, vaginal suppositories) | Bioadhesive; emulsifying agent; release-modifying agent; suspending agent; tablet binder; viscosity-increasing agent | Carbomers are white-colored, 'fluffy' acidic, hygroscopic powders with a slight characteristic odor |
| **Carboxymethylcellulose calcium** | Cellulose, carboxymethyl ether, calcium salt [9050-04-8] |  | Accepted for use as a food additive in Japan at concentrations up to 2% w/w. Included in the FDA IIG (oral, capsules and tablets) | Stabilizing agent; suspending agent; tablet and capsule disintegrant; viscosity-increasing agent; water-absorbing agent | Carboxymethylcellulose calcium occurs as a lose white to yellowish-white, hygroscopic, fine powder |
| **Carboxymethylcellulose sodium** | Cellulose, carboxymethyl ether, sodium salt [9004-32-4] |  | GRAS listed. Accepted as a food additive in Europe. Included in the FDA IIG (dental preparations; inhalations; intra-articular, intrabursal, intradermal, | Coating agent; stabilizing agent; suspending agent; tablet and capsule disintegrant; tablet binder; | Carboxymethylcellulose sodium occurs as a white to almost white, odorless, granular powder |

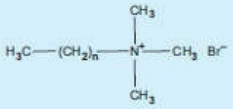
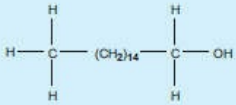
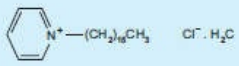
## Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| | | | intralesional, IM, intrasynovial and SC injections; oral capsules, drops, solutions, suspensions, syrups and tablets; topical and vaginal preparations) | viscosity-increasing agent; water-absorbing agent | |
| **Carrageenan** | Carrageenan [9000-07-1]<br>κ-Carrageenan [11114-20-8]<br>γ-Carrageenan [9064-57-7] |  | GRAS listed. Accepted as a food additive in Europe. Included in the FDA IIG (dental; oral granules, powders and syrups, topical; and transdermal preparations) | Emulsifying agent; gel base; stabilizing agent; suspending agent; sustained release tablet matrix; viscosity-increasing agent | Carrageenan, when extracted from the appropriate seaweed source, is a yellow-brown to white colored, coarse to fine powder that is odorless and tasteless |
| **Cellulose, microcrystalline** | Cellulose [9004-34-6] |  | GRAS listed. Accepted for use as a food additive in Europe. Included in the FDA IIG (inhalations; oral capsules, powders, suspensions, syrups, and tablets; topical and vaginal preparations). | Adsorbent; suspending agent; tablet and capsule diluent; tablet disintegrant | Microcrystalline cellulose is a purified, partially depolymerized cellulose that occurs as a white, odorless, tasteless, crystalline powder composed of porous particles |
| **Cellulose, powdered** | Cellulose [9004-34-6] |  | GRAS listed. Accepted for use as a food additive in Europe. Included in nonparenteral medicines licensed in the UK. Included in the Canadian List of Acceptable Non-medicinal Ingredients | Adsorbent; glidant; suspending agent; tablet and capsule diluent; tablet disintegrant | Powdered cellulose occurs as a white or almost white, odorless and tasteless powder of various particle sizes, ranging from a free-flowing fine or granular dense powder, to a coarse, fluffy, nonflowing material |

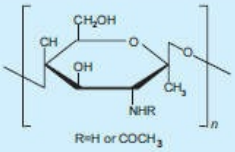
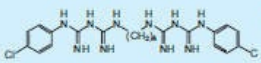
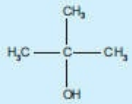
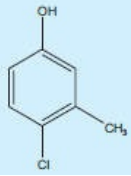
# Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| **Cellulose acetate** | Cellulose acetate [9004-35-7] |  | Included in the FDA IIG (oral tablets). Included in the Canadian List of Acceptable Non-medicinal Ingredients | Coating agent; extended release agent; tablet and capsule diluent [9035-69-2] Cellulose | Cellulose acetate occurs as a white to off-white powder, free-flowing pellets, or flakes. It istasteless and odorless, or may have a slight odor of acetic acid |
| **Cellulose acetate phthalate** | Cellulose, acetate, 1,2-benzenedicar-boxylate [9004-38-0] |  | Included in the FDA IIG oral tablets). ( Included in nonparenteral medicines licensed in the UK. Included in the Canadian List of Acceptable Non-medicinal Ingredients | Coating agent | Cellulose acetate phthalate is a hygro-scopic, white to off-white, free-flowing powder, granule, or flake. It is tasteless and odorless, or might have a slight odor of acetic acid |
| **Ceratonia** | Carob gum [9000-40-2] | Ceratonia is a naturally occurring plant material that consists chiefly of a high molecular weight hydrocolloidal polysaccharide, composed of D-galactose and D-mannose units combined through glycosidic linkages, which may be described chemically galactomannan. | GRAS listed. Accepted for use in Europe as a food additive. In Europe and the USA, ceratonia has been used in oral tablet formulations. | Controlled-release vehicle; stabilizing agent; suspending agent; tablet binder; viscosity increasing agent | Ceratonia occurs as a yellow-green or white colored powder. Although odorless and taste less in the dry powder form, ceratonia acquires a leguminous taste when boiled in water |
| **Cetostearyl alcohol** | Cetostearyl alcohol [67762-27-0] and [8005-44-5] | Cetostearyl alcohol is a mixture of solid aliphatic alcohols consisting mainly of stearyl($C_{18}H_{38}O$) and cetyl ($C_{16}H_{34}O$) alcohols | Accepted as an indirect food additive and as an adhesive and a component of packaging coatings of in the USA. Included in | Emollient; emulsifying agent; viscosity-increasing agent | Cetostearyl alcohol occurs as white or cream-colored unctuous masses, or almost white flakes or |

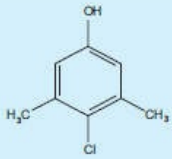
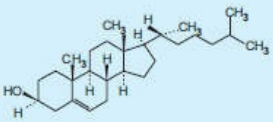
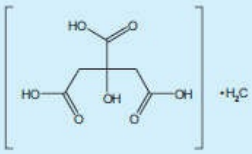
## Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| | | | the FDA IIG (oral tablets and topical emulsions, lotions, ointments, vaginal suppositories) | | granules. It has a faint, characteristic sweet odor |
| **Cetrimide** | Cetrimide [8044-71-1] |  | Included in nonparenteral medicines licensed in the UK. Included in the Canadian List of Acceptable Non-medicinal Ingredients | Antimicronial preservative; antiseptic; cationic surfactant; disinfectant | Cetrimide is a white to creamy white, free-flowing powder, with a faint but characteristic odor and a bitter, soapy taste |
| **Cetyl alcohol** | Hexadecan-1-ol [36653-82-4] |  | Included in the FDA IIG (ophthalmic preparations, oral capsules and tablets, otic and rectal preparations, topical aerosols, creams, emulsions, ointments and solutions, and vaginal preparations) | Coating agent; emulsifying agent; stiffening agent | Cetyl alcohol occurs as waxy, white flakes, granules, cubes, or castings. It has a faintcharacteristic odor and bland taste |
| **Cetylpyridinium chloride** | 1-Hexadecylpyridinium chloride [123-03-5] 1-Hexadecylpyridinium chloride monohydrate [6004-24-6] |  | Included in nonparenteral formulations licensed in the UK. Included in the FDA IIG, for use in inhalation and oral preparations. Reported in the EPA TSCA Inventory. It is not approved for use in Japan. Included in the Canadian List of Acceptable Non-medicinal Ingredients | Antimicrobial preservative; antiseptic; cationic surfactant; disinfectant; solubilizing agent; wetting agent | Cetylypyridinium chloride is a white powder with a characteristic odor. It is slightly soapy to the touch |

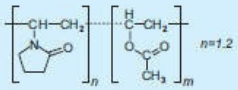
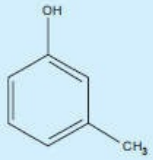
# Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| **Chitosan** | Poly-β-(1,4)-2-Amino-2-deoxy-D-glucose [9012-76-4] |  | Chitosan is registered as a food supplement in some countries | Coating agent; disintegrant; film-forming agent; mucoadhesive; tablet binder; viscosity increasing agent | Chitosan occurs as odorless, white or creamy-white powder or flakes. Fiber formation is quite common during precipitation and the chitosan may look 'cottonlike' |
| **Chlorhexidine** | *N,N″*-Bis(4-chlorophenyl)-3,12-diimino-2,4,11,13 tetraazate-tradecanediim-idamide [55-56-1] |  | Chlorhexidine salts are included in nonparenteral and parenteral medicines licensed in the UK. Included in the Canadian List of Acceptable Non-medicinal Ingredients | Antimicrobial preservative; antiseptic | Chlorhexidine occurs as an odorless, bitter tasting, white crystalline powder. *See* Section 17 for information on chlorhexidine salts |
| **Chlorobutanol** | 1,1,1-Trichloro-2-methyl-2-propanol [57-15-8] |  | Included in the FDA IIG (IM, IV, and SC injections, inhalations, nasal, otic, ophthalmic, and topical preparations) | Antimicrobial preservative; plasticizer | Volatile, colorless or white crystals with a musty, camphoraceous odor |
| **Chlorocresol** | 4-Chloro-3-methylphenol [59-50-7] |  | Included in the FDA IIG (topical creams and emulsions). Included in nonparenteral and parenteral medicines licensed in the UK | Antimicrobial preservative; disinfectant | Colourless or almost colorless, dimorphous crystals or crystalline powder with a characteristic phenolic odor |

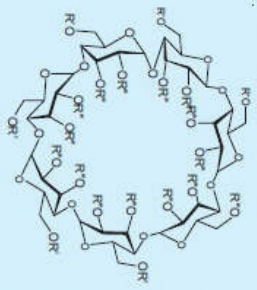
## Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| **Chloroxylenol** | 4-Chloro-3,5-xylenol [88-04-0] |  | Included in the FDA IIG (otic preparations; topical creams and emulsions). Included in nonparenteral medicines licensed in the UK | Antimicrobial preservative; antiseptic; disinfectant | White or cream-colored crystals or crystalline powder with a characteristic phenolic odor. Volatile in steam |
| **Cholesterol** | Cholest-5-en-3β-ol [57-88-5] |  | Included in the FDA IIG (injections, ophthalmic, topical and vaginal preparations) Included in nonparenteral medicines licensed in the UK. Included in the Canadian List of Acceptable Non-medicinal Ingredients | Emollient; emulsifying agent | Cholesterol occurs as white or faintly yellow, almost odorless, pearly leaflets, needles, powder, or granules. On prolonged exposure to light and air, cholesterol acquires a yellow to tan color |
| **Citric acid monohydrate** | |  | GRAS listed. The anhydrous form is accepted for use as a food additive in Europe. Included in the FDA IIG (inhalations; IM, IV, and other injections; ophthalmic preparations; oral capsules, solutions, suspensions and tablets; topical and vaginal preparations) | Acidifying agent; antioxidant; buffering agent; chelating agent; flavor enhancer | Citric acid monohydrate occurs as colorless or translucent crystals, or as a white crystalline, efflorescent powder. It is odorless and has a strong acidic taste. The crystal structure is orthorhombic |
| **Colloidal silicon dioxide** | Silica [7631-86-9] | $SiO_2$ | GRAS listed. Included in the FDA IIG (oral capsules, suspensions, and tablets; | Adsorbent; anticaking agent; emulsion stabi- | Colloidal silicon-dioxide is a submicroscopic |

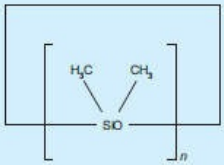
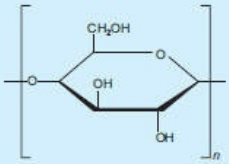
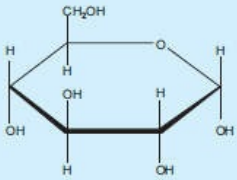
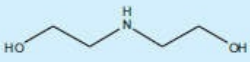
## Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| | | | transdermal and vaginal preparations) | lizer; glidant; suspending agent; tablet disintegrant; thermal stabilizer; viscosity-increasing agent | fumed silica with a particle size of about 15 nm. It is a light, loose, bluish-white-colored, odorless, tasteless, nongritty amorphous powder |
| **Copovidone** | Acetic acid ethenyl ester, polymer with 1-ethenyl-2-pyrrolidinone [25086-89-9] |  | Copovidone is included in the FDA IIG (oral tablets, oral film-coated tablets, sustained action) | Film-former; granulating agent; tablet binder | It is a white to yellowish-white amorphous powder. It is typically spray-dried with a relatively fine particle size. It has a slight odor and a faint taste |
| **Cresol** | Methylphenol [1319-77-3] |  | Included in the FDA IIG (IM, IV, intradermal, and SC injections). Included in parenteral medicines licensed in the UK. Included in the Canadian List of Acceptable Non-medicinal Ingredients | Antimicrobial preservative; disinfectant | It is a colorless, yellowish to pale brownish-yellow, or pink-colored liquid, with a characteristic odor similar to phenol but more tarlike |
| **Croscarmellose sodium** | Cellulose, carboxymethyl ether, sodium salt, crosslinked [74811-65-7] | Croscarmellose sodium is a crosslinked polymer of carboxymethylcellulose sodium | Included in the FDA IIG (oral capsules, granules, sublingual tablets, and tablets). Included in nonparenteral medicines licensed in the UK | Tablet and capsule disintegrant | Croscarmellose sodium occurs as an odorless, white or grayish-white powder |

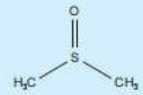
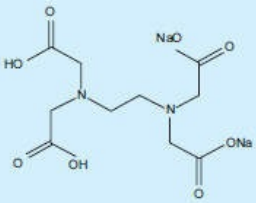
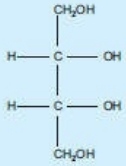
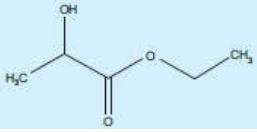
## Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| **Crospovidone** | 1-Ethenyl-2-pyrrolidinone homopolymer [9003-39-8] | $(C_6H_9NO)n$ | Accepted for use as a food additive in Europe. Included in the FDA IIG (IM injections, oral capsules and tablets; topical, transdermal, and vaginal preparations) | Tablet disintegrant | Crospovidone is a white to creamy-white, finely divided, free-flowing, practically tasteless, odorless or nearly odorless, hygroscopic powder |
| **Cyclodextrins** | α-Cyclodextrin [10016-20-3] β-Cyclodextrin [7585-39-9] γ-Cyclodextrin [17465-86-0] |  R′, R″= H for 'natural' α-, β- and γ- cyclodextrins R′, R″= $CH_3$ for methyl cyclodextrins R′, R″= $CHOHCH_3$ for 2-hydroxyethyl cyclodextrins R′, R″= $CH_2CHOHCH_3$ for 2-hydroxypropyl cyclodextrins | β-cyclodextrin is included in the FDA IIG (IM, IV injections, and other injection preparations) | Solubilizing agent; stabilizing agent | Cyclodextrins are cyclic oligosaccharides containing at least six D-(+)-glucopyranose unit sattached by a(1→) glucoside bonds. The three natural cyclodextrins, α, β and γ differ in their ring size and solubility. They contain 6, 7, or 8 glucose units, respectively. Cyclodextrins occur as white, practically odorless, fine crystalline powders, having a slightly sweet taste. Some cyclodextrin derivatives occur as amorphous powders |
| **Cyclomethic-one** | Cyclopolydimeth-ylsiloxane [69430-24-6] | | Included in the FDA IIG (oral powder for reconstitution). Included | Emollient; humectant; viscosity- | Cyclomethicone occurs as a clear, colorless and tasteless volatile liquid |

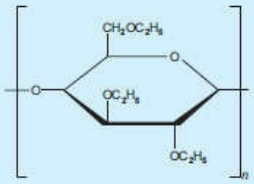
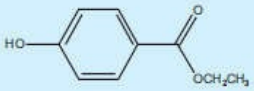
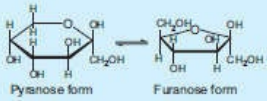
## Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| | |  | in nonparenteral medicines licensed in the UK. Included in the Canadian List of Acceptable Non-medicinal Ingredients | increasing agent | |
| **Dextrin** | Dextrin [9004-53-9] |  | GRAS listed. Included in the FDA IIG (IV injections, oral tablets and topical preparations). Included in nonparenteral medicines licensed in the UK. Included inthe Canadian List of Acceptable Non-medicinal Ingredients | Stiffening agent; suspending agent; tablet binder; tablet and capsule diluent | Dextrin is partially hydrolyzed maize (corn) or potato starch. It is a white, pale yellow or brown-colored powder with a slight characteristic odor |
| **Dextrose** | D-(+)-Glucose monohydrate [5996-10-1] |  | Included in the FDA IIG (capsules; inhalations; IM, IV, and SC injections; tablets, oral solutions, and syrups). Included in nonparenteral and parenteral medicines licensed in the UK. Included in the Canadian List of Acceptable Non-medicinal Ingredients | Tablet and capsule diluent; therapeutic agent; tonicity agent; sweetening agent | Dextrose occurs as odorless, sweet-tasting, colorless crystals or as a white crystalline or granular powder. The JP 2001 describes dextrose as dextrose anhydrous; the PhEur 2005 specifies dextrose anhydro |
| **Diethanola-mine** | 2,2′-Iminobisethanol [111-42-2] |  | Included in the FDA IIG (IV infusions, ophthalmic solutions, and topical preparations). Included in medicines | Alkalizing agent; emulsifying agent | At about room temperature it is a white, deliquescent solid. Above room temperature |

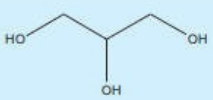
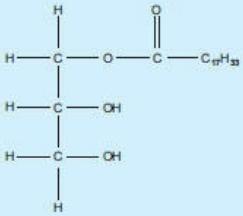
## Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| | | | licensed in the UK. Included in the Canadian List of Acceptable Non-medicinal Ingredients | | diethanolamine is a clear, viscous liquid with a mildly ammoniacal odor |
| **Dimethyl sulfoxide** | Sulfinylbisme-thane [67-68-5] |  | Included in the FDA IIG (IV infusions, SC implants, and topical preparations) | Penetration enhancer; solvent | Dimethyl sulfoxide occurs as a colorless, viscous liquid, or as colorless crystals that are miscible with water, alcohol, and ether |
| **Disodium edetate** | Ethylenediam-ine tetraacetic acid, disodium salt [139-33-3] Disodium ethylenediami-netetraacetate dihydrate [6381-92-6] |  | GRAS listed. Included in the FDA IIG (inhalations; injections; ophthalmic preparations; oral capsules, solutions, suspensions, syrups, and tablets; rectal topical, and vaginal preparations) | Chelating agent | Disodium edetate occurs as a white crystalline, odorless powder with a slightly acidic taste |
| **Erythritol** | Erythritol [149-32-6] |  | GRAS listed. Accepted for use as a food additive in Europe | Sweetening agent; tablet and capsule diluent; taste masking agent | Erythritol is a sugar alcohol (polyol) that occurs as a white or almost white powder orgranular or crystalline substance |
| **Ethyl lactate** | 2-Hydroxy-propanoic acid ethyl ester [97-64-3] |  | GRAS listed. Reported in the EPA TSCA inventory | Film-former; flavoring agent; solvent or co-solvent in liquid formulations | Ethyl lactate occurs as a clear colorless liquid with a sharp characteristic odor |

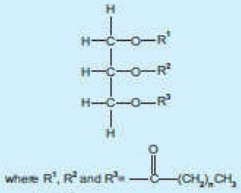
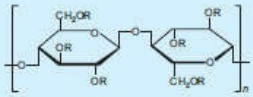
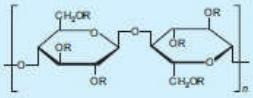
## Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| **Ethylcellulose** | Cellulose ethyl ether [9004-57-3] |  | GRAS listed. Accepted for use as a food additive in Europe. Included in the FDA IIG (oral capsules, suspensions and tablets; topical emulsions and vaginal preparations) | Coating agent; flavoring fixative; tablet binder; tablet filler; viscosity-increasing agent | Ethylcellulose is a tasteless, free-flowing, white to light tan-colored powder |
| **Ethylparaben** | Ethyl-4-hydroxybenzo-ate [120-47-8] |  | Accepted as a food additive in Europe. Included in the FDA IIG (oral, otic, and topical preparations) | Antimicrobial preservative | Ethylparaben occurs as a white, odorless or almost odorless, crystalline powder |
| **Fructose** | D-Fructose [57-48-7] |   | Included in the FDA IIG (oral solutions and suspensions; rectal preparations) | Dissolution enhancer; flavor enhancer; sweetening agent; tablet diluent | Fructose occurs as odorless, colorless crystals or a white crystalline powder with a very sweet taste |
| **Gelatin** | Gelatin [9000-70-8] | Gelatin is a generic term for a mixture of purified protein fractions obtained either by partialacid hydrolysis (type A gelatin) or by partial alkaline hydrolysis (type B gelatin) of animalcollagen. Gelatin may also be a mixture of both types | Included in the FDA IIG (dental preparations; inhalations; injections; oral capsules, pastilles, solutions, syrups and tablets; topical and vaginal preparations) | Coating agent; film-former; gelling agent; suspending agent; tablet binder; viscosi tyincreasing agent | Gelatin occurs as a light-amber to faintly yellow-colored, vitreous, brittle solid. It is practically odorless and tasteless and is available as translucent sheets and granules, or as a powder |

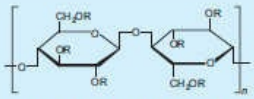
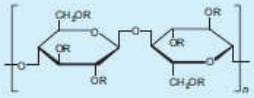
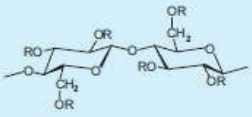
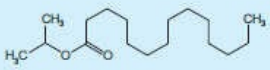
## Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| **Glycerin** | Propane-1,2,3-triol [56-81-5] |  | GRAS listed. Accepted for use as a food additive in Europe. Included in the FDA IIG (dental pastes; buccal preparations; inhalations; injections; nasal and ophthalmic preparations; oral capsules, suspensions and tablets; solutions, otic, rectal, topical, transdermal, and vaginal preparations) | Antimicrobial preservative; emollient; humectant; plasticizer; solvent; sweetening agent; tonicity agent | Glycerin is a clear, colorless, odorless, viscous, hygroscopic liquid; it has a sweet taste, approximately 0.6 times as sweet as sucrose |
| **Glyceryl Monooleate** | 9-Octadecenoic acid (Z), monoester with 1,2,3-propanetriol [25496-72-4] |  | GRAS listed. Included in the FDA IIG (oral capsules, oral powder, oral tablets; creams, controlled-release transdermal films) | Bioadhesive; emollient; emulsifying agent; emulsion stabilizer; gelling agent; mucoadhesive; nonionic surfactant; sustained-release agent | They are defined by the nominal content of monoacylglycerols and obtained by partial glycerolysis of vegetable oils mainly containing triacylglycerols of oleic acid or by esterification of glycerol by oleic acid, this fatty acid being of vegetable or animal origin |
| **Guar Gum** | Galactomannan polysaccharide [9000-30-0] | Guar gum consists of linear chains of (1→4)-β-D-mannopyranosyl units with | GRAS listed. Accepted for use as a food additive in Europe. Included in the | Suspending agent; tablet binder; tablet | Guar gum occurs as an odorless or nearly odorless, white to |

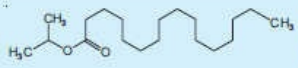
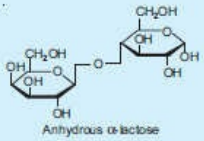
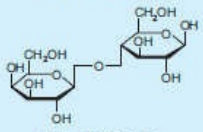
## Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| | | α-D-galactopyranosylunits attached by (1–6) linkages. The ratio of D-galactose to D-mannose is between 1 : 1.4 and 1 : 2 | FDA IIG (oral suspensions, syrups, and tablets; topical preparations; vaginal tablets) | disintegrant; viscosity-increasing agent | yellowish-white powder with a bland taste |
| **Hard fat** | Medium chain triglycerides [73398-61-5] |   | Included in the FDA Inactive Ingredients Guide (rectal and vaginal preparations). Included in non-parenteral medicines licensed in the UK | Suppository bases | A white or almost white, practically odorless, waxy, brittle mass. When heated to 50°C it melts to give a colorless or slightly yellowish liquid |
| **Hydroxyethyl cellulose** | Cellulose, 2-hydroxyethyl ether [9004-62-0] |  where R is H or $[—CH_2CH_2O—]mH$ | Included in the FDA IIG (ophthalmic preparations; oral syrups and tablets; otic and topical preparations). Included in nonparenteral medicines licensed in the UK | Coating agent; suspending agent; tablet binder; thickening agent; viscosity -increasing agent | Hydroxyethyl cellulose occurs as a light tan or cream to white-colored, odorless andtasteless, hygroscopic powder |
| **Hydroxypro-pyl cellulose** | Cellulose, 2-hydroxypropyl ether [9004-64-2] |  | GRAS listed. Accepted for use as a food additive in Europe. Included in the FDA IIG (oral capsules and tablets; topical and transdermal preparations). Included in non-parenteral medicines licensed in the UK | Coating agent; emulsifying agent; stabilizing agent; suspending agent; tablet binder; thickening agent; viscosity-increasing agent | Hydroxypropyl cellulose is a white to slightly yellow-colored, odorless and tasteless powder |

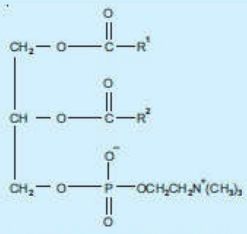
# Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| **Hydroxypropyl Cellulose, Low-substituted** | Cellulose, 2-hydroxypropyl ether (low-substituted) [78214-41-2] |  R is H or $[CH_2CH(CH_3)O]_mH$ | Approved for use in pharmaceuticals in Europe, Japan, USA, and other countries. Included inthe Canadian List of Acceptable Non-medicinal Ingredients | Tablet and capsule disintegrant; tablet binder | It occurs as a white to yellowish white powder or granules. It is odorless or has a slight, characteristic odor, and it is tasteless |
| **Hypromellose** | Cellulose hydroxypropyl methyl ether [9004-65-3] |  where R is H, $CH_3$, or $CH_3CH(OH)CH_2$ | GRAS listed. Accepted for use as a food additive in Europe. Included in the FDA IIG (ophthalmic preparations; oral capsules, suspensions, syrups, and tablets; topical and vaginal preparations) | Coating agent; film-former; rate-controlling polymer for sustained release; stabilizing agent; suspending agent; tablet binder; viscosity-increasing agent | Hypromellose is an odorless and tasteless, white or creamy-white fibrous or granular powder |
| **Hypromellose Acetate Succinate** | Cellulose, 2-hydroxypropyl-methyl ether, acetate hydrogen butanedioate [71138-97-1] |  | Included in the FDA IIG for use in oral preparations (capsules, and delayed-action preparations) | Component of controlled-release or sustained-release dosage forms; enteric coating agent; film-forming agent; solid dispersion vehicle | Hypromellose acetate succinate is a white to off-white powder or granules. It has a faintacetic acid-like odor and a barely detectable taste. |
| **Isopropyl Myristate** | 1-Methylethyl tetradecanoate [110-27-0] |  | Included in the FDA IIG (otic, topical, transdermal, and vaginal preparations) | Emollient; oleaginous vehicle; skin penetrant; solvent | sopropyl myristate is a I clear, colorless, practically odorless liquid of low viscosity that congeals at about 5°C |

# Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| **Isopropyl Palmitate** | 1-Methylethyl hexadecanoate [142-91-6] |  | Included in the FDA IIG (topical and transdermal preparations). Usedin nonparenteral medicines licensed in the UK | Emollient; oleaginous vehicle; skin penetrant; solvent | Isopropyl palmitate is a clear, colorless to pale yellow-colored, practically odorless viscous liquid that solidifies at less than 16°C |
| **Lactose, Anhydrous** | *O*-β-D-galactopyranosyl-(1→4)-β-D-glucopyranose [63-42-3] |  Anhydrous α-lactose  Anhydrous β-lactose | Included in the FDA IIG (IM, IV, and SC injections; oral capsules and tablets; inhalation preparations; rectal, transdermal, and vaginal preparations). Included in non-parenteral and parenteral medicines licensed in the UK. Included in the Canadian List of Acceptable Non-medicinal Ingredients | Binding agent; directly compressible tableting excipient; lyophilization aid; tablet and capsule filler | Lactose occurs as white to off-white crystalline particles or powder. Several different brands of anhydrous lactose are commercially available which contain anhydrous β-lactose and anhydrous α-lactose |
| **Lanolin** | Anhydrous lanolin [8006-54-0] | The USP 28 describes lanolin as the purified wax-like substance obtained from the wool of the sheep, Ovis aries Linné (Fam. Bovidae), that has been cleaned, decolorized, and deodorized | Included in the FDA IIG (ophthalmic, otic, topical, and vaginal preparations). Included in non-parenteral medicines licensed in the UK. Included in the Canadian List of Acceptable Non-medicinal Ingredients | Emulsifying agent; ointment base | Lanolin is a pale yellow-colored, unctuous, waxy substance with a faint, characteristic odor. Melted lanolin is a clear or almost clear, yellow liquid |
| **Lecithin** | Lecithin [8002-43-5] | | GRAS listed. Accepted for use as a food additive in | Emollient; emulsifying | Lecithins vary greatly in their physical form, |

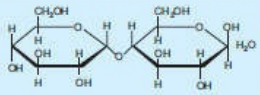
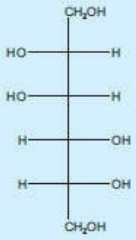
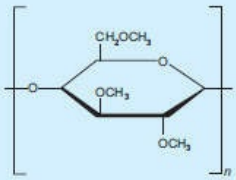
## Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| | |  α-Phosphatidylcholine $R^1$ and $R^2$ are fatty acids, which may be different or identical | Europe. Included in the FDA IIG (inhalations; IM and IV injections; otic preparations; oral capsules, suspensions and tablets; rectal, topical, and vaginal preparations) | agent; solubilizing agent | from viscous semiliquids to powders, depending upon the free fatty acid content. They may also vary in color from brown to light yellow, depending upon whether they are bleached or unbleached or on the degree of purity |
| **Macrogol 15 Hydroxystea-rate** | Polyethylene glycol-15-hydroxystearate [70142-34-6] | Macrogol 15-hydroxystear-ate as a mixture of mainly monoestersand diesters of 12-hydroxystearic acid and macrogols obtained by the ethoxylation of 12-hydroxystearic acid | | Dissolution enhancer; nonionic surfactant; solubilizing agent | Macrogol 15-hydroxystearate is a yellowish-white waxy mass at room temperature, which becomes liquid at approximately 30°C |
| **Magnesium Aluminum Silicate** | Aluminum magnesium silicate [12511-31-8] Magnesium aluminum silicate [1327-43-1] | The complex is composed of a three-lattice layer of octahedral alumina and two tetrahedral silica sheets. The aluminum is substituted to varying degrees by magnesium (with sodium or potassium for balance of electrical charge) | Included in the FDA IIG (oral granules, solutions, suspensions and tablets; rectal; and topical preparations; vaginal preparations) | Adsorbent; stabilizing agent; suspending agent; tablet and capsule disintegrant; tablet binder; viscosity-increasing agent | Magnesium aluminum silicate as a blend of colloidal montmorillonite and saponite that has been processed to remove grit and nonswellable ore components |

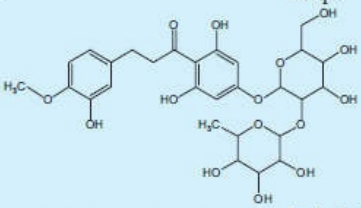
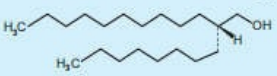
## Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| **Magnesium carbonate** | Magnesium carbonate anhydrous [546-93-0] | Light magnesium carbonate, $(MgCO_3)_3 \cdot Mg(OH)_2 \cdot 3H_2O$, and magnesium carbonate hydroxide, $(MgCO_3)_4 \cdot Mg(OH)_2 \cdot 5H_2O$ | GRAS listed. Accepted as a food additive in Europe. Included in the FDA IIG (oral capsules and tablets). Included in non-parenteral medicines licensed inthe UK | Adsorbent; antacid; tablet and capsule diluent | Magnesium carbonate occurs as light, white-colored friable masses or as a bulky, white colored powder. It has a slightly earthy taste and is odorless but, since it has a high absorptive ability, magnesium carbonate can absorb odors |
| **Magnesium oxide** | Magnesium oxide [1309-48-4] | MgO | GRAS listed. Accepted for use as a food additive in Europe. Included in the FDA IIG (oral capsules and tablets) | Anticaking agent; emulsifying agent; glidant; tablet and capsule diluent | Two forms of magnesium oxide exist: a bulky form termed light magnesium oxide and adense form termed heavy magnesium oxide |
| **Magnesium silicate** | Silicic acid, magnesium salt [1343-88-0] | $MgO \cdot SiO_2 \cdot xH_2O$ | GRAS listed. Accepted for use as a food additive in Europe. Included in the FDA IIG (oral tablets) | Anticaking agent; glidant | Magnesium silicate occurs as an odorless and tasteless, fine, white-colored powder that is free from grittiness |
| **Magnesium stearate** | Octadecanoic acid magnesium salt [557-04-0] | $[CH_3(CH_2)_{16}COO]_2Mg$ | GRAS listed. Accepted as a food additive in the UK. Included in the FDA IIG (oral capsules, powders, and tablets; buccal and vaginal tablets; topical preparations) | Tablet and capsule lubricant | Magnesium stearate is a very fine, light white, precipitated or milled, impalpable powder of low bulk densit, having a faint odor of stearic acid and a characteristic taste |

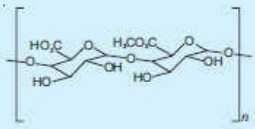
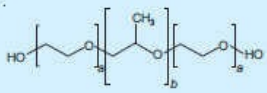
# Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| **Maltose** | 4-*O*-α-D-Glucopyranosyl-β-D-glucopyranose anhydrous [69-79-4]4-*O*-α-D-Glucopyranosyl-β-D-glucopyranose monohydrate [6363-53-7] |  | In the USA, maltose is considered as a food by the FDA and is therefore not subject to food additive and GRAS regulations. Included in the FDA IIG (oral solutions) | Sweetening agent; tablet diluent | Maltose occurs as white crystals or as a crystalline powder. It is odorless and has a sweet taste approximately 30% that of sucrose |
| **Mannitol** | D-Mannitol [69-65-8] |  | GRAS listed. Accepted for use as a food additive in Europe. Included in the FDA IIG (IP, IM, IV, and SC injections; infusions; buccal, oral and sublingual tablets, powders and capsules; ophthalmic preparations; topical solutions) preparations. | Diluent; diluent for lyphilized free-flowing granules. It sweetening agent; tablet and capsule diluent; tonicity agent | Mannitol occurs as a white, odorless, crystalline powder, or has a sweet taste, approximately as sweet as glucose and half as sweet as sucrose, and imparts a cooling sensation in the mouth |
| **Methylcellu-lose** | Cellulose methyl ether [9004-67-5] |  | GRAS listed. Accepted as a food additive in Europe. Included in the FDA IIG (sublingual tablets; IM injections; nasal preparations; ophthalmic preparations; | Coating agent; emulsifying agent; suspending agent; tablet and capsule disintegrant; | Methylcellulose occurs as a white, fibrous powder or granules. It is practically odorless and tasteless. It should be labeled to indicate its viscosity type |

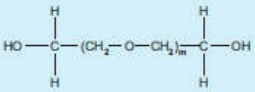
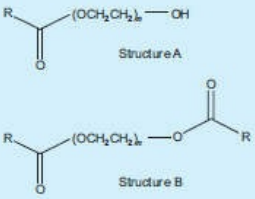
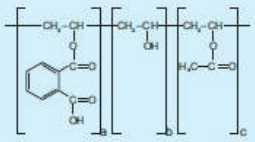
## Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| | | | oral capsules, oral suspensions, and oral tablets; topical and vaginal preparations) | tablet binder; viscosity-increasing agent | |
| **Neohesperidin Dihydrochalcone** | 1-[4-[[2-*O*-(6-Deoxy-α-L-mannopyranosyl)-β-D glucopyranosyl]oxy]-2,6dihydroxyphenyl]-3-(3-hydroxy-4-methoxyphenyl) propan-1-one [20702-77-6] |  | GRAS listed. Accepted for use as a food additive in Eurpe | Flavour enhancer; sweetening agent | Neohesperidin dihydrochalcone occurs as a white or yellowish-white powder with an intensely sweet taste |
| **Octyldodecanol** | Octyldodecanol [5333-42-6] |  | Included in the FDA IIG (topical, transdermal, and vaginal preparations) | Emollient; emulsifying agent; lubricant; solvent; thickening agent | Octyldodecanol occurs as a clear, colorless, or yellowish, oily liquid |
| **Palmitic acid** | Hexadecanoic acid[57-10-3] |  | GRAS listed. Included in the FDA IIG (oral tablets). Included in non-parenteral medicines licensed in the UK | Emulsifying agent; skin penetrant; tablet and capsule lubricant | Palmitic acid occurs as white crystalline scales with a slight characteristic odor and taste |
| **Paraffin** | Paraffin [8002-74-2] | Paraffin is a purified mixture of solid saturated hydrocarbons having the general formula C*n*H2*n*+2 | Accepted in the UK for use in certain food applications. Included in the FDA IIG (oral capsules and tablets, topical emulsions, and ointments) | Ointment base; stiffening agent | Paraffin is an odorless and tasteless, translucent, colorless, or white solid. It feels slightly greasy to the touch and may show a brittle fracture |

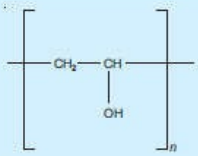
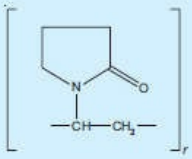
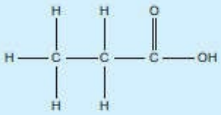
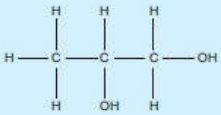
## Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| **Pectin** | Pectin[9000-65-5] |  | GRAS listed. Accepted for use as a food additive in Europe. Included in the FDA IIG (dental paste; oral powders; topical pastes) | Adsorbent; emulsifying agent; gelling agent; thickening agent; stabilizing agent | Pectin occurs as a coarse or fine, yellowish-white, odorless powder that has a mucilaginous taste |
| **Petrolatum and Lanolin Alcohols** | Petrolatum [8009-03-8] and Lanolin alcohols [8027-33-6] | A mixture of petrolatum and lanolin alcohols | Accepted for use in topical pharmaceutical formulations and cosmetics. Included in the Canadian List of Acceptable Non-medicinal Ingredients | Emollient; ointment base; plasticizer | A pale ivory-colored, soft solid with a faint, characteristic sterol odor |
| **Petrolatum** | Petrolatum [8009-03-8] | Petrolatum is a purified mixture of semisolid saturated hydrocarbons having the general formula C*n*H2*n*+2, and is obtained from petroleum | GRAS listed. Accepted for use in certain food applications in many countries worldwide. Included in the FDA IIG otic, topical, and transdermal preparations) oral capsules and tablets, | Emollient; ointment base even when melted | Petrolatum is a pale yellow to yellow-colored, translucent, soft unctuous mass. It is odorless, tasteless, and not more than slightly |
| **Poloxamer** | α-Hydro-ω-hydroxypoly (oxyethylene) poly (oxypropylene) poly (oxyethylene) block copolymer [9003-11-6] |  | Included in the FDA IIG (IV injections; inhalations, ophthalmic preparations; oral powders, solutions, suspensions, and syrups; topical preparations) | Dispersing agent; emulsifying and coemulsifying agent; solubilizing agent; tablet lubricant; wetting agent | Poloxamers generally occur as white, waxy, free-flowing prilled granules, or as cast solids. They are practically odorless and tasteless. At room temperature |
| **Polycarbophil** | Polycarbophil [9003-01-4] | Polycarbophils are polymers of acrylic acid crosslinked with divinyl glycol | GRAS listed. Included in the FDA IIG (vaginal gel; oral, troche). Included in | Adsorbent; bioadhesive; controlled- | Polycarbophil occurs as fluffy, white to off-white, mildly acidic |

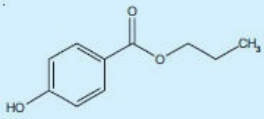
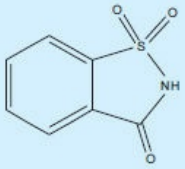
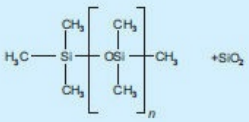
## Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| | | | non-parenteral medicines licensed in the UK | release tablet binder; emulsifying agent; thickening agent; suspending agent | polymer powder with slightly acetic odor |
| **Polyethylene Glycol** | α-Hydro-ω-hydroxypoly (oxy-1,2-ethanediyl) [25322-68-3] |  | Included in the FDA IIG (dental preparations; IM and IV injections; ophthalmic preparations; oral capsules, solutions, syrups, and tablets; rectal, topical, and vaginal preparations) | Ointment base; plasticizer; solvent; suppository base; tablet and capsule lubricant | Liquid grades (PEG 200–600) occur as clear, colorless or slightly yellow-colored, viscous liquids. They have a slight but characteristic odor and a bitter, slightly burning taste |
| **Polyoxyethylene Stearates** | Polyethylene glycol stearate [9004-99-3] pharmaceutical Excipients 2215 Polyethylene glycol distearate [9005-08-7] |  Structure A  Structure B | Included in the FDA IIG (dental solutions; IV injections; ophthalmic preparations; oral capsules and tablets; otic suspensions; topical creams, emulsions, lotions, ointments, and solutions; and vaginal preparations) | Emulsifying agent; solubilizing agent; wetting agent | Varied from pasty sold waxt solid to solid for polyoxyl 12 stereate to polyoxyl 150 distereate |
| **Polyvinyl Acetate Phthalate** | Polyvinyl acetate phthalate [34481-48-6] |  | Included in the FDA IIG (sustained-action oral tablet). Included in non-parenteral medicines licensed in Europe | Coating agent | Polyvinyl acetate phthalate is a free-flowing white to off-white powder and may have a slight odor of acetic acid. The material is essentially amorphous |

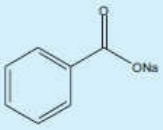
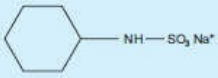
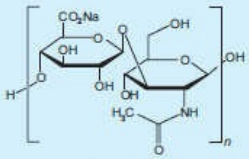
## Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| **Polyvinyl alcohol** | Ethenol, homopolymer [9002-89-5] |  | Included in the FDA IIG (ophthalmic preparations and oral tablets). Included in non-parenteral medicines licensed in the UK | Coating agent; lubricant; stabilizing agent; viscosity-increasing agent | Polyvinyl alcohol occurs as an odorless, white to cream-colored granular powder |
| **Povidone** | 1-Ethenyl-2-pyrrolidinone homopolymer [9003-39-8] |  | Accepted for use in Europe as a food additive. Included in the FDA IIG (IM and IV injections; ophthalmic preparations; oral capsules, drops, granules, suspensions, and tablets; sublingual tablets; topical and vaginal preparations) | Disintegrant; dissolution aid; suspending agent; tablet binder | Povidone occurs as a fine, white to creamy-white-colored, odorless or almost odorless, hygroscopic powder |
| **Propionic acid** | Propionic acid [79-09-4] |  | GRAS listed. Accepted for use in Europe as a food additive. In Japan, propionic acid is restricted to use as a flavoring agent | Acidifying agent; antimicrobial preservative; antioxidant; esterifying agent | Propionic acid occurs as a corrosive, oily liquid having a slightly pungent, disagreeable, rancid odor. It is flammable |
| **Propylene glycol** | 1,2-Propanediol [57-55-6] (–)-1,2-Propanediol [4254-14-2] (+)-1,2-Propanediol [4254-15-3] |  | GRAS listed. Accepted for use as a food additive in Europe. Included in the FDA IIG (dental preparations, IM and IV injections, inhalations, ophthalmic, oralotic, percutaneous, rectal, topical, and vaginal preparations) | Antimicrobial preservative; disinfectant; humectant; plasticizer; solvent; stabilizer for vitamins; water-miscible cosolvent | Propylene glycol is a clear, colorless, viscous, practically odorless liquid with a sweet, slightly acrid taste resembling that of glycerin |

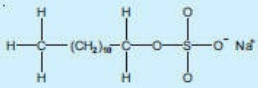
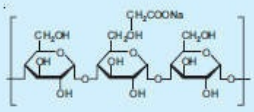
## Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| **Propylparaben** | Propyl 4-hydroxybenzoate [94-13-3] |  | Accepted as a food additive in Europe. Included in the FDA IIG (IM,IV, and SC injections; inhalations; ophthalmic preparations; oral capsules, solutions, suspensions, and tablets; otic, rectal, topical, and vaginal preparations) | Antimicrobial preservative | Propylparaben occurs as a white, crystalline, odorless, and tasteless powder |
| **Saccharin** | 1,2-Benzisothiazol-3(2*H*)-one 1,1-dioxide [81-07-2] |  | Accepted for use as a food additive in Europe. Note that the EU number 'E954' is applied to both saccharin and saccharin salts. Included in the FDA IIG (oral solutions, syrups, tablets, and topical preparations) | Sweetening agent | Saccharin occurs as odorless white crystals or a white crystalline powder. It has an intensely sweet taste, with a metallic aftertaste that at normal levels of use can be detected by approximately 25% of the population |
| **Simethicone** | α-(Trimethylsilyl-ω-methylpoly[oxy(dimethylsilylene)], mixture with silicon dioxide [8050- 81-5] |  where $n$ = 200–350 | GRAS listed. Included in the FDA IIG (oral emulsions, powders, solutions, suspensions, tablets, and rectal and topical preparations) | Antifoaming agent; tablet diluent; water-repelling agent | Simethicone occurs as a translucent, gray-colored, viscous fluid |

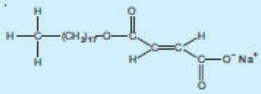
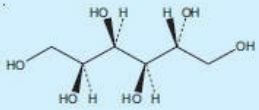
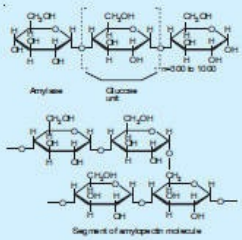
# Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| **Sodium alginate** | Sodium alginate [9005-38-3] | Sodium alginate consists chiefly of the sodium salt of alginic acid, which is a mixture of polyuronic acids composed of residues of D-mannuronic acid and L-guluronic acid | GRAS listed. Accepted in Europe for use as a food additive. Included in the FDA IIG (oral suspensions and tablets) | Stabilizing agent; suspending agent; tablet and capsule disintegrant; tablet binder; viscosity increasing agent | Sodium alginate occurs as an odorless and tasteless, white to pale yellowish-brown colored powder |
| **Sodium benzoate** | Sodium benzoate [532-32-1] |  | GRAS listed. Accepted as a food additive in Europe. Included in the FDA IIG (dental preparations; IM and IV injections; oral capsules, solutions and tablets; rectal; and topical preparations) | Antimicrobial preservative; tablet and capsule lubricant | Sodium benzoate occurs as a white granular or crystalline, slightly hygroscopic powder. It is odorless, or with faint odor of benzoin and has an unpleasant sweet and saline taste |
| **Sodium cyclamate** | Sodium *N*-cyclohexylsul-famate [139-05-9] |  | Included in the FDA IIG (oral powder, solutions and suspensions). Included in nonparenteral medicines licensed in the UK | Sweetening agent | Sodium cyclamate occurs as white, odorless or almost odorless crystals or as a crystalline powder with an intensely sweet taste |
| **Sodium hyaluronate** | Sodium hyaluronate [9067-32-7] |  | Included in the FDA IIG (topical gel preparation) | Humectant; lubricant; matrix for sustained release | Sodium hyaluronate occurs as white to off-white powder or granules. It is very hygroscopic |
| **Sodium lactate** | Sodium lactate[72-17-3] | Sodium lactate solution is a mixture of the enantiomers of sodium 2-hydro- | GRAS listed (not for infant formulas). Included in the FDA IIG (epidural, IM, IV, | Antimicrobial preservative; buffering agent; | Sodium lactate occurs as a clear, colorless, slightly syrupy liquid. |

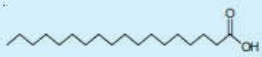
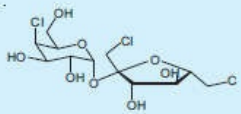
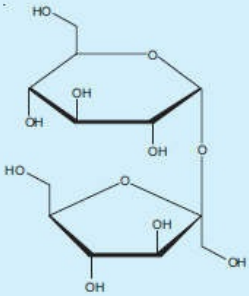
## Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| | | xypropanoate in approximately equal proportions | and SC injections; oral suspensions; topical gels and solutions) | emulsifying agent; flavoring agent; humectant | It is odorless, or has a slight odor with a characteristic saline taste. It is hygroscopic |
| **Sodium lauryl sulfated** | Sulfuric acid monododecyl ester sodium salt [151-21-3] |  | GRAS listed. Included in the FDA IIG (dental preparations; oral capsules, suspensions, and tablets; topical and vaginal preparations) | Anionic surfactant; detergent; emulsifying agent; skin penetrant; tablet and capsule lubricant; wetting agent | Sodium lauryl sulfate consists of white or cream to pale yellow-colored crystals, flakes, or powder having a smooth feel, a soapy, bitter taste, and a faint odor of fatty substances |
| **Sodium metabisulfite** | Sodium pyrosulfite [7681-57-4] | $Na_2S_2O_5$ | GRAS listed. Accepted for use as a food additive in Europe. Included in the FDA IIG (epidural, inhalation; IM, and IV injections; ophthalmic solutions; oral preparations, rectal, topical, and vaginal preparations) | Antioxidant | Sodium metabisulfite occurs as colorless, prismatic crystals or as a white to creamy-white crystalline powder that has the odor of sulfur dioxide and an acidic, saline taste. Sodium metabisulfite crystallizes from water as a hydrate containing seven water molecules |
| **Sodium starch glycolate** | Sodium carboxymethyl starch [9063-38-1] |  | Included in the FDA IIG (oral capsules and tablets). Included in nonparenteral medicines licensed in the UK | Tablet and capsule disintegrant | Sodium starch glycolate is a white to off-white, odorless, tasteless, free-flowing powder |

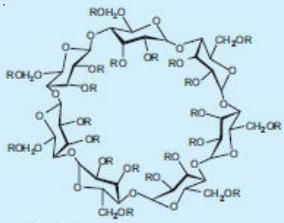
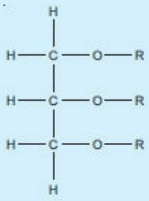
## Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| **Sodium stearyl fumarate** | 2-Butenedioic acid, monooctadecyl ester, sodium salt [4070-80-8] |  | GRAS listed. Permitted by the FDA for direct addition to food for human consumption as a conditioning or stabilizing agent in various bakery products, flour-thickened foods, dehydrated potatoes, and processed cereals up to 0.2–1.0% by weight of the food | Tablet and capsule lubricant | Sodium stearyl fumarate is a fine, white powder with agglomerates of flat, circular-shaped particles |
| **Sorbitol** | D-Glucitol [50-70-4] |  | GRAS listed. Accepted for use as a food additive in Europe. Included in the FDA IIG (intra-articular and IM injections; nasal; oral capsules, solutions, suspensions, syrups and tablets; rectal, topical, and vaginal preparations) | Humectant; plasticizer; sweetening agent; tablet and capsule diluent | Sorbitol occurs as an odorless, white or almost colorless, crystalline, hygroscopic powder |
| **Starch** | Starch [9005-25-8] |  | GRAS listed. Included in the FDA IIG (buccal tablets, oral capsules, powders, suspensions and tablets; topical preparations; and vaginal tablets) | Glidant; tablet and capsule diluent; tablet and capsule disintegrant; tablet binder | Starch occurs as an odorless and tasteless, fine, white-colored powder comprising very small spherical or ovoid granules whose size and shape are characteristic for each botanical variety |

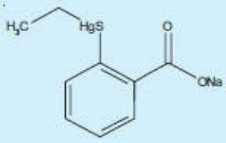
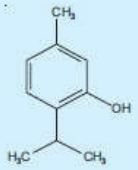
## Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| **Stearic Acid** | Octadecanoic acid [57-11-4] |  | GRAS listed. Accepted as a food additive in Europe (fatty acids). Included in the FDA IIG (sublingual tablets; oral capsules, solutions, suspensions, and tablets; topical and vaginal preparations) | Emulsifying agent; solubilizing agent; tablet and capsule lubricant | Stearic acid is a hard, white or faintly yellow-colored, somewhat glossy, crystalline solid or a white or yellowish white powder. It has a slight odor and taste suggesting tallow |
| **Stearyl Alcohol** | 1-Octadecanol [112-92-5] |  | Included in the FDA IIG (oral tablets, rectal topical, and vaginal preparations) | Stiffening agent | Stearyl alcohol occurs as hard, white, waxy pieces, flakes, or granules with a slight characteristic odor and bland taste |
| **Sucralose** | 1,6-Dichloro-1, 6-dideoxy-β-D-fructofuranosyl-4-chloro-4-deoxy-α-Dgalactopyran-oside [56038-13-2] |  | The FDA, in April 1998, approved sucralose for use as a tabletop sweetener and as an additive in a variety of food products | Sweetening agent | Sucralose is a white to off-white-colored, free-flowing, crystalline powder |
| **Sucrose** | β-D-fructofuranosyl-α-D-glucopyrano-side [57-50-1] |  | GRAS listed. Included in the FDA IIG (injections; oral capsules, solutions, syrups, and tablets; topical preparations). Included in non-parenteral and parenteral medicines licensed in the UK. Included in the Canadian | Base for medicated confectionery; coating agent; granulating agent; sugar coating adjunct; suspending | Sucrose is a sugar obtained from sugar cane (Saccharum officinarum Linné (Fam.Gramineae), sugarbeet (Beta vulgaris Linné (Fam. Chenopodiaceae), and other sources. Sucrose |

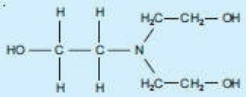
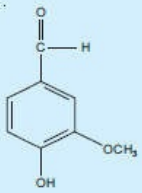
## Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| | | | List of Acceptable Non-medicinal Ingredients | agent; sweetening agent; tablet binder; tablet and capsule diluent; tablet filler; viscosity-increasing agent | occurs as colorless crystals, as crystalline masses or blocks, or as a white crystalline powder; it is odorless a and has a sweet taste |
| **Sulfobuty lether β-cyclodextrin** | β-Cyclodextrin sulfobutylether, sodium salt [1824100-00-0] |  R = $H_{21-n}$ or ($CH_2CH_2CH_2CH_2SO_2ONa$) *n* where *n* = 6.0–7.1 | Sulfobutylether β-cyclodextrin is included in IV and IM injectable products currently approved and marketed in the USA and Europe. It is included in the FDA IIG for IM and IV use | Dissolution-enhancing agent; drug delivery system; osmotic agent; solubilizing agent; stabilizing agent; tablet and capsule diluent; viscosity -increasing agent | Sulfobutylether β-cyclodextrin occurs as a white amorphous powder |
| **Suppository bases, hard fat** | Hard fat triglyceride esters |  where R = H or $OC(CH_2)nCH_3$; *n* = 7–17 | Included in the FDA IIG (rectal and vaginal preparations). Included in nonparenteral medicines licensed in the UK | Suppository base | A white or almost white, practically odorless, waxy, brittle mass. When heated to 50°C it melts to give a colorless or slightly yellowish liquid |

## Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| **Talc** | Talc [14807-96-6] | Talc is a purified, hydrated, magnesium silicate, approximating to the formula $Mg_6(Si_2O_5)_4(OH)_4$ | Accepted for use as a food additive in Europe. Included in the FDA IIG (buccal tablets; oral capsules and tablets; rectal and topical preparations) | Anticaking agent; glidant; tablet and capsule diluent; tablet and capsule lubricant | Talc is a very fine, white to grayish-white, odorless, impalpable, unctuous, crystalline powder. It adheres readily to the skin and is soft to the touch and free from grittiness |
| **Thimerosal** | Ethyl[2-mercaptoben-zoato(2–)-*O*,*S*]-mercurate(1–) sodium [54-64-8] |  | Included in the FDA IIG (IM, IV, and SC injections; ophthalmic, otic, and topical preparations) | Antimicrobial preservative; antiseptic | Thimerosal is a light cream-colored crystalline powder with a slight, characteristic odor |
| **Thymol** | Thymol [89-83-8] |  | GRAS listed. Included in the FDA IIG (inhalation, liquid; oral, powder for solution). Included in non-parenteral medicines licensed in the UK | Antioxidant; antiseptic; cooling agent; disinfectant; flavoring agent; skin penetrant; therapeutic agent | Thymol occurs as colorless or often large translucent crystals, or as a white crystalline powder with a herbal odor (aromatic and thyme-like) and a pungent caustic taste |
| **Titanium dioxide** | Titanium oxide [13463-67-7] | $TiO_2$ | Accepted as a food additive in Europe. Included in the FDA IIG (dental paste; intrauterine suppositories; ophthalmic preparations; oral capsules, suspensions, tablets; topical and transdermal preparations) | Coating agent; opacifier; pigment | White, amorphous, odorless, and tasteless nonhygroscopic powder |

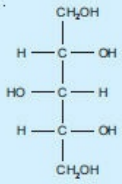
## Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| **Tragacanth** | Tragacanth gum [9000-65-1] | The gum consists of a mixture of water-insoluble and water-soluble polysaccharides. Bassorin, which constitutes 60–70% of the gum, is the main water-insoluble portion, while the remainder of the gum consists of the water-soluble material tragacanthin | GRAS listed. Accepted for use as a food additive in Europe. Included in the FDA IIG (buccal/ sublingual tablets, oral powders, suspensions, syrups, and tablets). Included in nonparenteral medicines licensed in the UK | Suspending agent; viscosity-increasing agent | Tragacanth gum occurs as flattened, lamellated, frequently curved fragments, or as straight or spirally twisted linear pieces from 0.5 to 2.5 mm in thickness; it may also be obtained in apowdered form |
| **Triethanol-amine** | 2,2′,2″-Nitrilotrieth-anol [102-71-6] |  | Included in the FDA IIG (rectal, topical, and vaginal preparations). Included in nonparenteral medicines licensed in the UK | Alkalizing agent; emulsifying agent | It is a clear, colorless to pale yellow-colored viscous liquid having a slight ammoniacal odor. It is a mixture of bases, mainly 2,2′,2″-nitrilotriethanol although it also contains 2,2′-iminobisethanol (diethanolamine) and smaller amounts of 2-aminoethanol (monoethanolamine) |
| **Vanillin** | 4-Hydroxy-3-methoxyben-zaldehyde [121-33-5] |  | GRAS listed. Included in the FDA IIG (oral solutions, suspensions, syrups, and tablets). Included in nonparenteral medicines licensed in the UK | Flavouring agent | White or cream, crystalline needles or powder with characteristic vanilla odor and sweet taste |

## Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| **Wax, anionic emulsifying** | Anionic emulsifying wax [8014-38-8] | anionic emulsifying wax contain cetostearyl alcohol, purified water, and either sodium lauryl sulfate or a sodium salt of a similar sulfated higher primary aliphatic alcohol | Included in nonparenteral medicines licensed in the UK. Included in the Canadian List of Acceptable Non-medicinal Ingredients | Emulsifying agent; stiffening agent | An almost white or pale yellow colored, waxy solid or flakes which when warmed become plastic before melting |
| **Wax, carnauba** | Carnauba wax [8015-86-9] | Carnauba wax consists primarily of a complex mixture of esters of acids and hydroxy acids, mainly aliphatic esters, ω-hydroxy esters, *p*-methoxycinnamic aliphatic esters, and *p*-hydroxycinnamic aliphatic diesters composed of several chain lengths, in which $C_{26}$ and $C_{32}$ alcohols are the most prevalent | GRAS listed. Accepted for use as a food additive in Europe. Included in the FDA IIG (oral capsules and tablets) | Coating agent | Carnauba wax occurs as a light brown- to pale yellow-colored powder, flakes, or irregular lumps of a hard, brittle wax. It has a characteristic bland odor and practically no taste |
| **Wax, white** | White beeswax [8012-89-3] | Beeswax consists of 70–75% of a mixture of various esters of straight-chain monohydric alcohols with even-numbered carbon chains from $C_{24}$ to $C_{36}$ esterified with straight-chain acids. These straight-chain acids also have even numbers of carbon atoms up to $C_{36}$ together with some $C_{18}$ hydroxy acids | GRAS listed. Accepted for use as a food additive in Europe. Included in the FDA IIG (oral capsules and tablets, rectal, topical, and vaginal preparations) | Controlled-release vehicle; stabilizing agent; stiffening agent | White wax consists of tasteless, white or slightly yellow-colored sheets or fine granules with some translucence. Its odor is similar to that of yellow wax but is less intense |

## Some commonly used excipients in pharmaceutical formulations (Appendix I)

| Name of excipient | Chemical name and CAS registry number | Structural formula | Regulatory status | Application in pharmaceutical formulation | Description |
|---|---|---|---|---|---|
| **Xanthan gum** | Xanthan gum [11138-66-2] | $(C_{35}H_{49}O_{29})n$ Approximately $2 \times 106$ | GRAS listed. Accepted for use as a food additive in Europe. Included in the FDA IIG (oral solutions, suspensions, and tablets; rectal and topical preparations) | Stabilizing agent; suspending agent; viscosity-increasing agent | Xanthan gum occurs as a cream- or white-colored, odorless, free-flowing, fine powder |
| **Xylitol** | *xylo*-Pentane-1,2,3,4,5-pentol [87-99-0] |  | GRAS listed. Accepted for use as a food additive in Europe. Included in the FDA IIG (oral solution, chewing gum) | Antimicrobial preservative; base for medicated confectionery; coating agent; emollient; humectant; sweetening agent; tablet and capsule diluent | It is odorless, with a sweet taste that imparts a cooling sensation |
| **Zinc stearate** | Octadecanoic acid zinc salt [557-05-1] | $C_{36}H_{70}O_4Zn$ | GRAS listed. Included in the FDA IIG (oral capsules and tablets). Included in non-parenteral medicines licensed in the UK | Tablet and capsule lubricant | Zinc stearate occurs as a fine, white, bulky, hydrophobic powder, free from grittiness and with a faint characteristic odor |

# 13:
# Tablets

# PART I: TABLET DESIGN

When a new drug is discovered, one of the first questions a pharmaceutical company asks is whether or not the drug can be effectively administered for its intended effect by the oral route. If not, the drug is primarily relegated to administration in a hospital setting or physician's office. If self-administration by patient cannot be achieved, the sales of the drug constitute only a small fraction of what the market would be otherwise. Of drugs that are administered orally, solid oral dosage forms represent the preferred class of product. The reasons for this preference are as follows.

Tablets and capsules represent unit dosage forms in which one usual dose of the drug has been accurately placed. By comparison, liquid oral dosage forms, such as syrups, suspensions, emulsions, solutions, and elixirs, are usually designed to contain one dose of medication in 5 to 30 ml. The patient is then asked to measure his or her own medication using a teaspoon, tablespoon, or other measuring device. Such dosage measurements are typically in error by a factor ranging from 20 to 50% when the drug is self-administered by the patient.

Liquid oral dosage forms have other disadvantages and limitations when compared with tablets. They are much more expensive to ship (one liquid dosage weighs 5 g or more versus 0.25 to 0.40 g for the average tablet), and breakage or leakage during shipment is a more serious problem with liquids than with tablets. Taste masking of the drug is often a problem (if the drug is in solution even partially). In addition, liquids are less portable and require much more space per number of doses on the pharmacist's shelf. Drugs are in general less stable (both chemically and physically) in liquid form than in a dry state and expiration dates tend to be shorter. Careful attention is required to assure that the product will not allow a heavy microbiologic burden to develop on standing or under normal conditions of use once opened (preservation requirements).

## Advantages

Of the two oral solid dosage forms commonly employed in this country, the tablet and the capsule, the tablet has a number of advantages. One of the major advantages of tablets over capsules, which has recently proved significant, is that the tablet is an essentially tamperproof dosage form. The major advantage of capsules—their ability to hide their contents from sight and to mask or hide the taste or odor of their contents—makes them the most vulnerable to tampering of all dosage forms. In contrast, any adulteration of a tablet after its manufacture is almost certain to be observed. Even though improved packaging provides some consumer protection for such dosage forms as capsules, which are susceptible to tampering, *no* packaging is completely tamperproof.

A major disadvantage of capsules over tablets is their higher cost. Capsules, whether hard gelatin or soft elastic capsules, employ a capsule shell to contain the drug contents. The cost of this shell is approximately several tenths of a cent or more, depending on whether the capsule is banded, printed with identification, or otherwise treated. In addition to this is the cost of filling. This filling cost is higher than the typical total cost of tablet production, now that direct compression methods of tablet manufacture exist, since the capsule filling operation is far slower than the tablet compression operation. In consideration of these few comparisons to capsules, the following may be cited as the primary potential advantages of tablets:

1. They are a unit dose form, and they offer the greatest capabilities of all oral dosage forms for the greatest dose precision and the least content variability.
2. Their cost is lowest of all oral dosage forms.
3. They are the lightest and most compact of all oral dosage forms.
4. They are in general the easiest and cheapest to package and ship of all oral dosage forms.
5. Product identification is potentially the simplest and cheapest, requiring no additional processing steps when employing an embossed or monogrammed punch face.
6. They may provide the greatest ease of swallowing with the least tendency for “hang-up” above the stomach, especially when coated, provided that

tablet disintegration is not excessively rapid.

7. They lend themselves to certain special-release profile products, such as enteric or delayed-release products.
8. They are better suited to large-scale production than other unit oral forms.
9. They have the best combined properties of chemical, mechanical and microbiologic stability of all the oral forms.

The development pharmacist should know fully what the potential advantages of tablets are as a dosage form class. If these general advantages together with the specific criteria specifications for the product are not met, an optimum or even near-optimum product may not have been achieved.

## *Disadvantages*

The disadvantages of tablets include the following:

1. Some drugs resist compression into dense compacts, owing to their amorphous nature or flocculent, low-density character.
2. Drugs with poor wetting, slow dissolution properties, intermediate to large dosages, optimum absorption high in the gastrointestinal tract, or any combination of these features may be difficult or impossible to formulate and manufacture as a tablet that will still provide adequate or full drug bioavailability.
3. Bitter-tasting drugs, drugs with an objectionable odor, or drugs that are sensitive to oxygen or atmospheric moisture may require encapsulation or entrapment prior to compression (if feasible or practical), or the tablets may require coating. In such cases, the capsule may offer the best and lowest cost approach.

In summary of the foregoing advantages and disadvantages of tablets in comparison to other oral dosage forms, tablets do provide advantages to the pharmacist, in minimal storage space requirements as well as ease of dispensing and possibly control; to the patient in convenience of use, optimum portability, and lowest cost; and to the physician in flexibility of dosage (with bisected tablets), and in accuracy and precision of dosage in general.

## TYPES AND CLASSES OF TABLETS

Tablets are classified by their route of administration or function, by the type of drug delivery system they represent within that route, and by their form and method of manufacture. Table 13.1 lists the various classes of tablets, with the primary classification being the route of administration or function.

**Table 13.1:** Types and classes of tablets

| |
|---|
| *Oral tablets for ingestion* |
| Compressed tablets or standard compressed tablets (CT) |
| Multiple compressed tablets (MCT) |
| Layered tablets |
| Compression-coated tablets |
| Chewable tablets |
| Sugar- and chocolate-coated tablets |
| Film-coated tablets |
| Repeat-action tablets |
| Delayed-action and enteric coated tablets Controlled release tablets |
| *Tablets used in the oral cavity* |
| Buccal and sublingual tablets |
| Troches and lozenges |
| Dental cones |
| *Tablets administered by other routes* |
| Implantation tablets |
| Vaginal tablets |
| *Tablets used to prepare solutions* |
| Effervescent tablets |
| Dispensing tablets (DT) |
| Hypodermic tablets (HT) |
| Tablet triturates (TT) |

## Tablets Ingested Orally

Well over 90% of the tablets manufactured today are ingested orally. Orally ingested tablets are designed to be swallowed intact, with the exception of chewable tablets.

### *Compressed Tablets or Standard Compressed Tablets*

This category refers to standard uncoated tablets made by compression and employing any of the three basic methods of manufacture: wet granulation, double compaction, or direct compression. Tablets in this category are usually intended to provide rapid disintegration and drug release. Most tablets containing drugs intended to exert a local effect in the gastrointestinal tract are of this type. These drugs are typically water-insoluble and include such therapeutic categories as the antacids and adsorbents. Other drugs in this group are intended to produce a systemic effect. These drugs have some aqueous solubility, dissolve from the tablet and disintegrated tablet fragments in GI contents, and are then absorbed and distributed in the body.

### *Multiple Compressed Tablets*

There are two classes of multiple compressed tablets: layered tablets and compression-coated tablets. Both types may be either two-component or three-component systems: two- or three-layer tablets, a tablet within a tablet, or a tablet within a tablet within a tablet. Both types of tablets usually undergo a light compression as each component is laid down, with the main compression being the final one. Tablet machine production speeds for multiple compressed tablets are appreciably slower than for standard compressed tablets, especially in the case of compression-coated tablets.

Tablets in this category are usually prepared for one of two reasons: to separate physically or chemically incompatible ingredients, or to produce repeat-action or prolonged-action products. In some cases, a two-layer tablet may provide adequate surface separation of reactive ingredients; if complete physical separation is required for stability purposes, the three-layer tablet may be employed. The layered tablet is preferred to the compression-coated tablet; surface contact between layers is lessened, and production is simpler and more rapid.

Multiple compressed tablets readily lend themselves to repeat-action

products, wherein one layer of the layered tablet or the outer tablet of the compression-coated tablet provides the initial dose, rapidly disintegrating in the stomach. The other layer or the inner tablet is formulated with components that are insoluble in gastric media but are released in the intestinal environment. The shortcoming of this category of dosage form for repeat-action products is that its performance is highly dependent on gastric emptying. If the second layer or core tablet quickly leaves the stomach following release of the initial fast-release dose, an entirely different blood level profile results than if there is a several-hour or longer delay before the second fraction is emptied. It is probably for this reason that relatively few repeat-action or controlled-release products using this approach are marketed.

## *Chewable Tablets*

Chewable tablets are intended to be chewed in the mouth prior to swallowing and are not intended to be swallowed intact. The purpose of the chewable tablet is to provide a unit dosage form of medication which can be easily administered to infants and children or to the elderly, who may have difficulty swallowing a tablet intact. The types of sugars and other components employed in chewable tablets have been designated in this chapter under the heading “Tablet Design and Formulation.” The most common chewable tablet on the market is the chewable aspirin tablet intended for use in children. Bitter or foul-tasting drugs are not good candidates for this type of tablet, and this fact restricts the use of the chewable tablet dosage form. Many antacid tablet products are of the chewable type. The chewable tablet offers two major advantages to the delivery of a solid antacid dosage form. First, the dose of most antacids is large, so that the typical antacid tablet would be too large to swallow. Second, as noted previously, the activity of an antacid is related to its particle size. If the tablet is chewed prior to swallowing, better acid neutralization may be possible from a given antacid dose.

## *Sugar- and Chocolate-coated Tablets*

Chocolate-coated tablets are nearly a thing of the past. They are too easily mistaken for candy by children. Sugar-coated tablets suffer the same disadvantage. Their primary historical role was to produce an elegant, glossy, easy-to-swallow tablet dosage form. Also, they permit separation of incompatible ingredients between coating and core, and this fact has been

widely utilized in preparing many multivitamin and multivitamin mineral combinations. The process as originally developed was time-consuming and required skilled coating artisans to be conducted properly. Earlier sugar coatings typically doubled tablet weight. Today, water-soluble polymers are often incorporated in the sugar solution, automated-spray coating equipment is employed, and high-drying-efficiency side-vented coating pans are used. The result is that the coatings are more elastic and mechanically stable, coat weight may be 50% or less of the core weight, and the process may be completed in a day or less.

## *Film-coated Tablets*

Film-coated tablets were developed as an alternative procedure to the preparation of coated tablets in which drug was not required in the coating. The initial film-coating compositions employed one or more polymers, which usually included a plasticizer for the polymer and possibly a surfactant to facilitate spreading, all delivered to the tablets in solution from an organic solvent. The filmcoating process was an attractive tablet coating method since it permitted the completion of the tablet coating operation in a period of one or two hours. An airless spray coating procedure was typically employed for such film-coating compositions, using either conventional coating pans or side-vented equipment. During the decade of the 1970s, several factors began to make solvent-based film coating less attractive. These factors were the increase in cost of the organic solvents, OSHA restrictions on worker exposure to solvent vapors, and EPA limitations on solvent vapor discharge to the atmosphere. As a result of these influences, many companies have now converted their earlier film-coating process to a totally aqueous-based procedure. Polymers such as hydroxypropyl cellulose and hydroxypropyl methylcellulose, which are dissolved in water with an appropriate plasticizer, are now widely used to produce immediate-release film coatings. The recent development of a colloidal dispersion of ethylcellulose in water also makes it possible to produce slow- or controlled-release film coatings without the use of organic solvents. A 30% ethylcellulose dispersion is marketed under the trade name Aquacoat by the FMC Corporation.

Film-coated tablets offer a number of advantages over sugar-coated tablets. These advantages include better mechanical strength of the coating based on the elasticity and flexibility of the polymer coating, little increase in

tablet weight, the ability to retain debossed markings on a tablet through the thin film coating, the avoidance of sugar, which is contraindicated in the diets of a significant segment of the population, and the employment of a process that may be continuous, or that readily lends itself to automation. The primary disadvantage of film coating compared with sugar coating is that it is difficult to produce film-coated tablets that match the physical appearance and elegance of the sugar-coated product. Film coating in the future will assume increasing importance as a means of controlling drug delivery release rates from both tablets and bead particles as well as from drug crystals. Film-coated tablets, which are basically tasteless, also offer the advantage over sugar-coated tablets of being less likely to be mistaken for candy.

## *Repeat-action Tablets*

The mode of operation of repeat-action tablets, and their limitations based on uncontrolled and unpredictable gastric emptying, have just been mentioned. In addition to multiple compressed tablets being used for this effect, sugar-coated tablets may also be employed. The core tablet is usually coated with shellac or an enteric polymer so that it will not release its drug load in the stomach. The second dose of drug is then added in the sugar coating, either in solution in the sugar syrup or as a part of the dusting powder added for rapid coat buildup. More uniform drug addition occurs if the drug is in solution or fine suspension in the sugar solution, especially if an automated-spray sugar-coating operation is employed. Even so, the coating operation will probably require interruption one or more times while the partially coated tablets are assayed to establish that the correct amount of drug has been applied in the coating.

## *Delayed-action and Enteric Coated Tablets*

The delayed-action tablet dosage form is intended to release a drug after some time delay, or after the tablet has passed through one part of the GI tract into another. The enteric coated tablet is the most common example of a delayed-action tablet product. All enteric coated tablets (which remain intact in the stomach but quickly release in the upper intestine) are a type of delayed-action tablet. Not all delayed-action tablets are enteric or are intended to produce the enteric effect. In veterinary product development, tablets may be designed to pass through the stomach (or several stomachs) of an animal or through all or most of the small intestine before releasing—or

even into the cecum or large bowel, as in the case of treating worm parasites located in this lower region. In a human drug application, a product may be designed to pass through the stomach intact and then release gradually for several hours or longer in the intestines.

The coatings that are used today to produce enteric effects are primarily mixed acid functionality and acid ester functionality synthetic or modified natural polymers. Cellulose acetate phthalate has the longest history of use as an enteric coating. More recently, polyvinyl acetate phthalate and hydroxypropyl methylcellulose phthalate have come into use. All three polymers have the common feature of containing the dicarboxylic acid, phthalic acid, in partially esterfied form. These polymers, being acid esters, are insoluble in gastric media that have a pH of up to about 4; they are intended to hydrate and begin dissolving as the tablets leave the stomach, enter the duodenum (pH of 4 to 6), and move further along the small intestine, where the pH increases to a range of 7 to 8. The primary mechanism by which these polymers lose their film integrity, thereby admitting intestinal fluid and releasing drug, is ionization of the residual carboxyl groups on the chain and subsequent hydration. The presence of esterases in the intestinal fluid that break down ester linkages of the polymer chains may also play some role, as may surface activity effects of bile salts and other components in bile that enter the upper small intestine via the bile duct.

Enteric coatings are employed for a number of therapeutic, safety, and medical reasons. Some drugs are irritating when directly exposed to the gastric mucosa, including aspirin and strong electrolytes such as $NH_4Cl$. While for most people the occasional aspirin tablet may not cause irritation, those on daily doses of aspirin, such as arthritics, may find gastric upset a major problem. Enteric coating is one method of reducing or eliminating irritation from such drugs. There are other drugs that if released in the stomach may produce nausea and vomiting. The low pH of the stomach destroys other drugs (for example, erythromycin), and hence enteric coating may be necessary to bring the drug through that environment to the more neutral intestinal contents. Yet another reason for enteric coating may be the desire to release the drug undiluted and in the highest concentration possible within the intestine. (Examples are intestinal antibacterial or antiseptic agents and intestinal vermifuges.) As in the case of repeat-action and other

controlled-release dosage forms, the influence of altering the release profile of the drug on total drug bioavailability, distribution, and pharmacokinetics must be investigated.

## *Controlled Release Tablets*

Theoretically, tablets offer the lowest-cost approach to sustained- and controlled-release solid dosage forms. Currently, the vast majority of such products are coated pellets placed in capsules. This particulate approach to sustained release has offered several advantages: metered particle emptying from the stomach, utilization of a plurality of coatings, and several release profiles for the various populations of coatings, thereby permitting an immediate release fraction followed by a sustaining fraction. Matrix slow-releasing tablets cannot match these characteristics. For this reason, and based on public expectations (or those of marketing specialists) that controlled-release products are expected to be beads in a capsule, relatively few sustained-release oral products have been in tablet form.

A major recent break in that trend is the highly successful sustained-release theophylline product of Key Pharmaceuticals, Inc., Theo-Dur. This is a unique type of sustained-release tablet that overcomes some of the limitations of earlier matrix tablets. Under gastric pH conditions, the Theo-Dur tablet slowly erodes; however, at a pH corresponding to the upper small intestine, the tablet disintegrates rapidly to release coated particles, which in turn slowly release drug. Two different release mechanisms are operative, neither of which is zero-order—erosion and decreasing surface area, and dissolution of coated particles—but the overall tablet release profile comprising the two mechanisms in sequence is nearly linear for most of the dose in the tablet. The result is the ability to control theophylline blood levels in a narrow range, above the minimum effective level and below the toxic level. This type of sustained-release tablet has clearly shown the potential of the tablet as a reliable sustained-release dosage form with good release profile. More sustained-release tablet forms of this type are sure to follow.

The Oros product of the Alza Corporation is another new zero-order sustained-release tablet product; it is based on osmotic pressure as the rate-controlling process. This concept will also expand the use of tablet dosage forms in controlled release. The first such products are already being marketed in Europe, and the drug indomethacin is about to be marketed in the

United States in an Oros system.

Film coatings that may be applied to tablets to provide diffusion-controlled “membranes” for constant drug release rate profiles as the tablet dosage form moves along the GI tract are also developed. Such a system offers the potential ultimate dosage form as a simple, low-cost, and reproducible physicochemical approach to oral controlled and sustained drug release.

## Tablets Used in the Oral Cavity

### *Buccal and Sublingual Tablets*

These two classes of tablets are intended to be held in the mouth, where they release their drug contents for absorption directly through the oral mucosa. These tablets are usually small and somewhat flat, and are intended to be held between the cheek and teeth or in the cheek pouch (buccal tablets), or beneath the tongue (sublingual tablets). Drugs administered by this route are intended to produce systemic drug effects, and consequently, they must have good absorption properties through the oral mucosa. Drug absorption from the oral mucosa into the bloodstream leads directly to the general circulation. Drug absorption from the gastrointestinal tract leads to the mesenteric circulation, which connects directly to the liver via the portal vein. Thus, drug absorption from the oral cavity avoids first-pass metabolism. The oral route of administration from these two classes of tablet dosage form thus offers several possible advantages: The gastric environment, where decomposition may be extensive (for certain steroids and hormones), may be avoided for drugs that are well absorbed in the mouth. A more rapid onset of drug action occurs than for tablets that are swallowed (an advantage with vasodilators given by this route). The first-pass effect may be avoided as noted previously, and for certain drugs (e.g. methyltestosterone), the nausea produced when the product is swallowed is avoided.

Buccal and sublingual tablets should be formulated with bland excipients, which do not stimulate salivation. This reduces the fraction of the drug that is swallowed rather than being absorbed through the oral mucosa. In addition, these tablets should be designed not to disintegrate but to slowly dissolve, typically over a 15 to 30 min period, to provide for effective absorption.

### *Troches and Lozenges*

These are two other types of tablets used in the oral cavity, where they are intended to exert a local effect in the mouth or throat. These tablet forms are commonly used to treat sore throat or to control coughing in the common cold. They may contain local anesthetics, various antiseptic and antibacterial agents, demulcents, astringents, and anti-tussives. Lozenges were originally termed pastilles, but are more commonly called cough drops. They are usually made with the drug incorporated in a flavored hard-candy sugar base.

Lozenges may be made by compression but are usually formed by fusion or by a candy-molding process. Troches, on the other hand, are manufactured by compression as are other tablets. These two classes of tablets are designed not to disintegrate in the mouth but to dissolve or slowly erode over a period of perhaps 30 min or less.

## *Dental Cones*

Dental cones are a relatively minor tablet form that are designed to be placed in the empty socket remaining following a tooth extraction. Their usual purpose is to prevent the multiplication of bacteria in the socket following such extraction by employing a slow-releasing antibacterial compound, or to reduce bleeding by containing an astringent or coagulant. The usual vehicle of these tablets is sodium bicarbonate, sodium chloride, or an amino acid. The tablet should not be formulated with a component that might provide media for bacterial proliferation. The tablet should be formulated to dissolve or erode slowly in the presence of a small volume of serum or fluid, over a 20-to 40-min period, when loosely packed in the extraction site.

## Tablets Administered by Other Routes

### *Implantation Tablets*

Implantation or depot tablets are designed for subcutaneous implantation in animals or man. Their purpose is to provide prolonged drug effects, ranging from one month to a year. They are usually designed to provide constant drug release rate. These tablets are usually small, cylindric, or rosette-shaped forms, and are typically not more than 8 mm in length. Since there are two major safety problems with this form of drug administration, this class of dosage form has achieved little use in humans. The safety problems include the need for a surgical technique to discontinue therapy, and tissue toxicity problems in the area of the implantation site. A special injector utilizing a hollow needle and plunger (the Kern injector) may be used to administer rod-shaped tablets. Surgical techniques may be required for administering tablets of other shapes. Implantation tablets have been largely replaced by other dosage forms, such as diffusion-controlled silicone tubes filled with drug or biodegradable polymers that contain entrapped drug in a variety of forms. The primary application of current implantation tablets and depot forms is to the administration of growth hormones to food-producing animals. In this case, the implant or depot should be made in an animal structure that is not consumed. The ear of the animal is typically used, and appropriate drug release to the animal from the ear site must be achieved.

### *Vaginal Tablets*

Vaginal tablets or inserts are designed to undergo slow dissolution and drug release in the vaginal cavity. The tablets are typically ovoid or pear-shaped to facilitate retention in the vagina. This tablet form is used to release antibacterial agents, antiseptics, or astringents to treat vaginal infections, or possibly to release steroids for systemic absorption. The tablets are often buffered to promote a pH favorable to the action of a given antiseptic agent. The buffer pH, however, should not be greatly differ from physiologic pH. The vehicle of these tablets is typically a slowly soluble material similar to agents described for the preparation of buccal and sublingual tablets. The tablets should be designed to be compatible with some type of plastic tube inserter, which is usually employed to place the tablet in the upper region of the vaginal tract.

## Tablets Used to Prepare Solutions

### *Effervescent Tablets*

Effervescent tablets are designed to produce a solution rapidly with the simultaneous release of carbon dioxide. The tablets are typically prepared by compressing the active ingredients with mixtures of organic acids—such as citric acid or tartaric acid—and sodium bicarbonate. When such a tablet is dropped into a glass of water, a chemical reaction is initiated between the acid and the sodium bicarbonate to form the sodium salt of the acid, and to produce carbon dioxide and water. The reaction is quite rapid and is usually completed within one minute or less. In addition to having the capability of producing clear solutions, such tablets also produce a pleasantly flavored carbonated drink, which assists in masking the taste of certain drugs. For many years, various saline cathartics were prepared as effervescent mixtures and powders. The most widely produced effervescent tablet today is one that contains aspirin. If a clear solution is to be produced, the drug that is incorporated in the tablet must be soluble at a neutral or slightly alkaline pH, and any lubricant or other additive employed to facilitate tablet compression must be water-soluble.

The advantage of the effervescent tablet as a dosage form is that it provides a means of extemporaneously preparing a solution containing an accurate drug dose. As in the case of aspirin, this dosage form may provide other advantages as well. The solution produced by the most widely marketed effervescent aspirin tablet has a pH of about 8. If the volume of the solution and the pH of the solution are adequate to raise the gastric contents to neutral or near-netural pH, the aspirin remains in solution and is rapidly available upon emptying from the stomach. Some literature has been published to indicate that this form of aspirin is less irritating to the stomach mucosa. In addition, neutralization of gastric contents may be rapidly obtained from solutions of this type of tablet. The product does, however, represent a “systemic” antacid effect, with an appreciable dose of sodium or potassium, and thus does not represent a recommended method of producing routine gastric neutralization.

The disadvantage of the effervescent tablet, and one reason for its somewhat limited utilization, is related to the difficulty of producing a chemically stable product. Even the moisture in the air during product 

preparation may be adequate to initiate effervescent reactivity. During the course of the reaction, water is liberated from the bicarbonate, which autocatalyzes the reaction. Providing adequate protection of effervescent tablets in the hands of the consumer is another problem. The moisture to which tablets are exposed after opening the container can also result in a rapid loss of product quality in the hands of the consumer. It is for this reason that effervescent tablets are specially packaged in hermetic-type foil pouches or are stack-packed in cylindric tubes with minimal air space. Another reason for such packing is the fact that the tablets are usually compressed to be soft enough to produce an effervescent reaction that is adequately rapid.

A number of investigators have looked at alternative effervescent components in recent years in an attempt to produce a more chemically stable system. Such studies have included investigation of malic acid, fumaric acid, and various acid anyhdrides, in combination with newer carbonate sources such as sodium glycine carbonate and various sesqui-carbonates. If, in the future, more chemically stable effervescent mixtures are identified that continue to provide rapid reactivity in water, the effervescent tablet system may expand as a method of producing extemporaneous drug-containing solutions.

## *Dispensing Tablets*

Dispensing tablets are intended to be added to a given volume of water by the pharmacist or the consumer, to produce a solution of a given drug concentration. Materials that have been commonly incorporated in dispensing tablets include mild silver proteinate, bichloride of mercury, merbromin, and quaternary ammonium compounds. The dispensing tablet must typically comprise totally soluble components, and the excipient ingredients of the tablet must not produce deleterious effects in the intended application of the solution or undesirable physical or chemical interactions with the active agent. In some cases, as in applications where the solution is to be used in contact with mucous membranes or on wounds, the tablet may also contain components to provide buffering or isotonicity. Dispensing tablets are less commonly used than formerly, since they cannot be employed on a routine basis with water of known quality to produce sterile solutions. Another difficulty with dispensing tablets is that some of the components previously used in this dosage form are highly toxic and are extremely hazardous, and even lethal, if mistakenly swallowed. Great care must be taken in the

packaging and labeling of such tablets to attempt to prevent their oral consumption. In the past, bichloride of mercury was usually prepared in coffin-shaped tablets, with an embossed skull and crossbones to emphasize its toxicity.

## *Hypodermic Tablets*

Hypodermic tablets are composed of one or more drugs with other readily water-soluble ingredients and are intended to be added to sterile water or water for injection. Such extemporaneous preparation of an injectable solution was once widely used in medicine, because the physician, especially the rural physician, could carry many vials of such tablets in his bag with only one bottle of sterile water for injection, to prepare a great many types of injectable medications as the need arose. Hypodermic tablets are little used today in this country because their use increases the likelihood, of administering a nonsterile solution, even though portable sterile filtration equipment exists to help assure the sterility and freedom from particulate matter in such a product. Furthermore, since physicians today practice most of their medicine from an office or a hospital, the advantage of portability of tablets for injection is far outweighed by the hazards and disadvantages of this dosage form in most medical situations.

## *Tablet Triturates*

Tablet triturates are small, usually cylindric, molded, or compressed tablets. Though rarely used today, they provided an extemporaneous method of preparation by the pharmacist. The drugs employed in such products were usually quite potent and were mixed with lactose and possibly a binder, such as powdered acacia, after which the mixture was moistened to produce a moldable, compactable mass. This mass was forced into the holes of a mold board fabricated from wood or plastic, after which the tablets were ejected using a pegboard, whose pegs matched the holes in the mold. The tablets were then allowed to dry and were available for dispensing. Since virtually every conceivable drug that would be useful in a tablet dosage form is available in that form, or in capsule form, there is virtually no need today for pharmacists to prepare tablets extemporaneously. Since in preparing this form of molded tablet, alcohol was commonly used to wet the powder mass to expedite drying of the tablets, tablet triturates were usually soft and quite friable. Many of the drugs employed in these tablets were highly potent, and

drug migration could occur as the alcohol evaporated, so content uniformity of such tablets was often questionable. Because of these problems and the question of producing bioavailable dosage forms from such extemporaneous preparations, the tablet triturate is rarely seen today.

# TABLET DESIGN AND FORMULATION

## Tablet Design

The objective of the design and manufacture of the compressed tablet is to deliver orally the correct amount of drug in the proper form, at or over the proper time and in the desired location, and to have its chemical integrity protected to that point. Aside from the physical and chemical properties of the medicinal agent(s) to be formulated into a tablet, the actual physical design, manufacturing process, and complete chemical makeup of the tablet can have a profound effect on the efficacy of the drug(s) being administered.

A tablet (1) should be an elegant product having its own identity while being free of defects such as chips, cracks, discoloration, contamination, and the like, (2) should have the strength to withstand the rigors of mechanical shocks encountered in its production, packaging, shipping, and dispensing and (3) should have the chemical and physical stability to maintain its physical attributes over time. Pharmaceutical scientists now understand that various physical properties of tablets can undergo change under environmental or stress conditions, and that physical stability, through its effect on bioavailability in particular, can be of more significance and concern in some tablet systems than chemical stability.

On the other hand, the tablet (4) must be able to release the medicinal agent(s) in the body in a predictable and reproducible manner and (5) must have a suitable chemical stability over time so as not to allow alteration of the medicinal agent(s). In many instances, these sets of objectives are competing. The design of a tablet that emphasizes only the desired medicinal effects may produce a physically inadequate product. The design of a tablet emphasizing only the physical aspects may produce tablets of limited and varying therapeutic effects. As one example of this point, Meyer and associates present information on 14 nitrofurantoin products, all of which passed the compendial physical requirements, but showed statistically significant bioavailability differences.

### *Formulation Components*

Regardless of how tablets are manufactured, conventional oral tablets for ingestion usually contain the same classes of components in addition to the

active ingredients, that may be selected to provide essential manufacturing technology functions (binders, lubricants, glidants), modify drug release (disintegrants, polymers), enhance stability (antioxidants), enhance patient acceptance (flavours, sweeteners), or aid in product identification (colorants). All nondrug components of a formula are termed excipients. Excipients are critical to the design of the tablet and play a major role in determining its quality and performance. A long list of possible excipients is available, but certain external factors such as cost, availability, functional reliability, and international acceptance govern their selection.

Regardless of why an excipient is selected, tablet excipients must meet certain criteria in the formulation.

These include the following:

1. They must be nontoxic and acceptable to the regulatory agencies in all countries where the product is to be marketed.
2. They must be commercially available in an acceptable grade in all countries where the product is to be manufactured.
3. Their cost must be acceptably low.
4. They must not be contraindicated by themselves (e.g. sucrose) or because of a component (e.g. sodium) in any segment of the population.
5. They must be physiologically inert.
6. They must be physically and chemically stable by themselves and in combination with the drug(s) and other tablet components.
7. They must be free of any unacceptable microbiologic “load.”
8. They must be color-compatible (not produce any off-color appearance).
9. If the drug product is also classified as a food (certain vitamin products), the diluent and other excipients must be approved direct food additives.
10. They must have no deleterious effect on the bioavailability of the drug(s) in the product.

## *Diluents*

Diluents are fillers designed to make up the required bulk of the tablet when the drug dosage itself is inadequate to produce this bulk. The dose of some drugs is sufficiently high that no filler is required (e.g. aspirin and certain

antibiotics). Round tablets for ingestion are usually in a size range of 3/16 to ½ inch. Tablets below 3/16 inch may be difficult for the elderly to handle, and those larger than ½ inch become difficult to swallow. This provides a tablet weight range of perhaps 120 to 700 mg for standard density organic materials. By using oval tablets, which may be easier to swallow, tablets weighing up to 800 mg or more may be produced. Tablet formulations may contain a diluent for secondary reasons: to provide better tablet properties such as improved cohesion, to permit use of direct compression manufacturing, or to promote flow.

There are cited cases of pharmaceutical manufacturers actually producing products in which an excipient reduced the bioavailability of a drug, or in which chemical incompatibilities existed. The former situation occurred with the marketing of an antibiotic that utilized a calcium salt as the diluent. The tetracycline product made with calcium phosphate filler had less than half the bioavailability of the standard product. Divalent and trivalent cations form insoluble complexes and salts with a number of amphoteric or acid functionality antibiotics, which greatly reduces their absorption (which is also why milk should not be coadministered with these drugs). A classic case of a chemical incompatibility that went unrecognized for several years was the interaction of certain amine drugs with the commonly used diluent lactose, in the presence of a metal stearate lubricant (such as magnesium stearate); the resulting tablets were gradually discolored with time (Maillard reaction). Tablet formulators should remember that physical and chemical interactions between formulation components may be promoted by the intimate contact between potential reactants that are tightly compressed together in a tablet compact. Thus, materials that are capable of forming a eutectic mixture, for example, may pose no problem when loosely packed as a powder in a capsule, while the same formulation when compressed in a tablet forms a compact that quickly softens and becomes unacceptable.

Table 13.2 lists some of the commonly used tablet diluents. Note that several of the diluents listed exist as hydrates (dibasic calcium phosphate and calcium sulfate). Diluents that exist in their common salt form as hydrates, containing appreciable bound water as water of crystallization, may nevertheless be excellent for very water-sensitive drugs, provided that the bound water is not released under any elevated storage condition to which the product might be exposed. Dibasic calcium phosphate and calcium sulfate

have the advantages of possessing low concentrations of unbound moisture and having a low affinity for atmospheric moisture. These are required features for any excipient material to be combined with a water-sensitive drug. The bound water of calcium sulfate is not released until a temperature of approximately 80°C is reached. Such bound water is usually unavailable for chemical reaction. Such excipients containing tightly bound water but having a low remaining moisture demand may be vastly superior to an anhydrous diluent, which has a moderate to high moisture demand.

**Table 13.2:** Diluents used in tablet formulation

| **Diluent** | **Proprietary name** | **Directly compressible** |
|---|---|---|
| Calcium carbonate* | Cal-Carb, Millicarb, Pharma-Carb, Sturcal | X |
| Calcium phosphate dibasic* | Cyfos, Calstar, Calipharm, Emcompress | √ |
| Calcium phosphate tribasic* | Tricafos, Tri-Cal, Tri-Tab | √ |
| Calcium sulfate* | Cal-Tab, Compactrol | √ |
| Cellulose, microcrystalline (MCC) | Avicel, Emcocel, Vivacel | √ |
| MCC silicified | Prosolv | X |
| Dextrates | Emdex | √ |
| Dextrose | Tabfine | X |
| Fructose | Fructofin | X |
| Lactitol | Finlac | √ |
| Lactose monohydrate | Fast-Flo, Lactochem, Zeparox Pharmatose, Tablettose | √ |
| Maltitol | Maltisorb, Maltit | X |

| | | |
|---|---|---|
| Maltodextrin | Glycidex, Lycatab, Maltrin | √ |
| Maltose | Advantose | X |
| Mannitol | Pearlitol | √ |
| Sorbitol | | √ |
| Starch | | X |
| Starch, pregelatinized | Pharma-Gel, Pre-Jel, Sepistab, Starch 1500, Starx 1500 | √ |
| Sucrose | | X |
| Sugar, compressible | Dipac, Nutab | X |
| Sugar spheres | Nu-Core, Nu-Pareil | X |
| Talc | | X |
| Xylitol | Xylifin, Xylitab | √ |

* Diluents insoluble in water

Lactose is the most widely used diluent in tablet formulation. Lactose is an excipient that has no reaction with most drugs, whether it is used in the hydrous or anhydrous form. Anhydrous lactose has the advantage over lactose in that it does not undergo the Maillard reaction, which can lead to browning and discoloration with certain drugs. The anhydrous form, however, picks up moisture when exposed to elevated humidity. Such tablets may have to be carefully packaged to prevent moisture exposure. When a wet granulation process is employed, the hydrous form of lactose should generally be used. Two grades of lactose are commonly available commercially: a 60 to 80 mesh (coarse) and an 80-to 100-mesh (regular) grade. In general, lactose formulations show good drug release rates, their granulations are readily dried, and the tablet disintegration times of lactose tablets are not strongly sensitive to variations in tablet hardness. Lactose is a low-cost diluent, but it may discolor in the presence of amine drug bases or salts of alkaline compounds.

Spray-dried lactose is one of several diluents now available for direct compression following mixing with the active ingredient, and possibly, a

disintegrant and a lubricant. If this form of lactose is allowed to dry out and the moisture content falls below the usual 3% level, the material loses some of its direct compressional characteristics. In addition to its direct compression properties, spray-dried lactose also has good flow characteristics. It can usually be combined with as much as 20 to 25% of active ingredient without losing these advantageous features. Spray-dried lactose is especially prone to darkening in the presence of excess moisture, amines, and other compounds, owing to the presence of a furaldehyde. A neutral or acid lubricant should be used when spray-dried lactose is employed.

Starch, which may come from corn, wheat or potatoes, is occasionally used as a tablet diluent. The USP grade of starch, however, has poor flow and compression characteristics and possesses a high typical moisture content of between 11 and 14%. Specially dried types of starch that have a standard moisture level of 2 to 4% are available, but at a premium price. Use of such starches in wet granulation is wasteful since their moisture levels increase to 6 to 8% following moisture exposure.

Various directly compressible starches are now available commercially. Sta-Rx 1500 is one such free-flowing, directly compressible starch; it may be used as a diluent, binder, and/or disintegrating agent. Since it is self-lubricating, it may be compressed alone, but when combined with as little as 5 to 10% of drug, it typically requires addition of a lubricant, and possibly a flow promoter such as 0.25% of a colloidal silicone dioxide. Sta-Rx 1500 contains about 10% moisture and is reportedly prone to softening when combined with excessive amounts (more than 0.5%) of magnesium stearate.

Two hydrolyzed starches are Emdex and Celutab, which are basically 90 to 92% dextrose and about 3 to 5% maltose. They are free-flowing and directly compressible. These materials may be used in place of mannitol in chewable tablets because of their sweetness and smooth feeling in the mouth. These materials contain about 8 to 10% moisture and may increase in hardness after compression.

Dextrose is also used as a tablet diluent. It is available from one supplier under the name Cerelose and comes in two forms: as a hydrate, and in anhydrous form for when low moisture contents are required. Dextrose is sometimes combined in formulation to replace some of the spray-dried lactose, which may reduce the tendency of the resulting tablets to darken.

Mannitol is perhaps the most expensive sugar used as a tablet diluent, but because of its negative heat of solution, its slow solubility, and its pleasant feeling in the mouth, it is widely used in chewable tablets. It is relatively nonhygroscopic and can be used in vitamin formulation, in which moisture sensitivity may be a problem. Mannitol formulations typically have poor flow characteristics and usually require fairly high lubricant levels.

Sorbitol is an optical isomer of mannitol and is sometimes combined in mannitol formulations to reduce diluent cost; however, sorbitol is hygroscopic at humidities above 65%. Both of these sugars have a low caloric content and are noncariogenic.

Sucrose, or sugar, and various sucrose-based diluents are employed in tablet making. Some manufacturers avoid their use in products that would subject a diabetic to multiple gram quantities of sugar. Some of the sucrose-based diluents have such tradenames as Sugartab (90 to 93% sucrose plus 7 to 10% invert sugar), DiPac (97% sucrose plus 3% modified dextrins), and Nu Tab (95% sucrose and 4% invert sugar with a small amount of corn starch and magnesium stearate). All of these diluents are available for direct compression, and some are also employed, with or without mannitol, in chewable tablets. All have a tendency to pick up moisture when exposed to elevated humidity.

Microcrystalline cellulose, often referred to by the tradename Avicel, is a direct compression material. Two tablet grades exist: pH 101 (powder) and pH 102 (granules). The flow properties of the material are generally good, and the direct compression characteristics are excellent. This is a somewhat unique diluent in that while producing cohesive compacts, the material also acts as a disintegrating agent. It is, however, a relatively expensive material when used as a diluent in high concentration and is thus typically combined with other materials. As in the case of starch, microcrystalline cellulose is often added to tablet formulation for several possible functions. It is a commonly employed excipient.

While a careful search of the literature reveals over 50 chemicals that have been evaluated and advocated as tablet diluents, those listed in Table 13.2 probably represent 80 to 90% of currently used diluents.

## *Binders and Adhesives*

These materials are added either dry or in liquid form during wet granulation

to form granules or to promote cohesive compacts for directly compressed tablets. Acacia and tragacanth are natural gums (listed in Table 13.3), and are employed in solutions ranging from 10 to 25% concentration, alone or in combination. These materials are much more effective when they are added as solutions in the preparation of granulations than when they are added dry to a direct compression formula. These natural gums have the disadvantage of being variable in their composition and performance based on their natural origin, and they are usually fairly heavily contaminated with bacteria. When these materials are used, their wet granulation masses should be quickly dried at a temperature above 37° to reduce microbial proliferation.

Gelatin is a natural protein and is sometimes used in combination with acacia. It is a more consistent material than the two natural gums, is easier to prepare in solution form, and forms tablets equally as hard as acacia or tragacanth. Starch paste has historically been one of the most common granulating agents. It is prepared by dispersing starch into water, which is then heated for some prescribed time. During the heating, the starch undergoes hydrolysis to dextrin and to glucose. A properly made paste is translucent rather than clear (which would indicate virtually complete conversion to glucose) and produces cohesive tablets that readily disintegrate when properly formulated. Liquid glucose, which is a 50% solution in water, is a fairly common wet granulating agent. Its properties are similar to those of sucrose solutions, which are commonly employed in concentrations between 50 and 74%. These sugar solutions are capable of producing wet granulations, which when tabletted, produce hard but somewhat brittle compacts. These materials have the advantage of being low-cost adhesives. Unless the sugar solutions are highly concentrated, bacterial proliferation may be a problem.

Modified natural polymers, such as the alginates and cellulose derivatives (methylcellulose, hydroxypropyl methylcellulose, and hydroxypropyl cellulose), are common binders and adhesives. Used dry for direct compression, they have some binder capabilities, while their aqueous solutions have adhesive properties. Hydroxypropyl cellulose may also be used as an alcohol solution to provide an anhydrous adhesive. Ethyl-cellulose may be used only as an alcoholic solution, and it may be expected to retard disintegration and dissolution time of drugs in the resulting tablets when wet granulation is employed. Polyvinylpyrrolidone is a synthetic polymer that

may be used as an adhesive in either an aqueous solution or alcohol. It also has some capabilities as a dry binder.

**Table 13.3:** Binders used in tablet formulation

| Binder | Proprietary name | Concentration used (%) |
|---|---|---|
| Acacia mucilage | | Up to 20 1–5 |
| Alginic acid | | |
| Carbomer | Carbopo | l5–10 |
| Carboxymethylcellulose sodium | Nymcel | 5–15 |
| Cellulose, microcrystalline | Avicel, Emcocel, Vivacel | |
| Ethyl cellulose | Aquacoat | 1–3 |
| Gelatin | | 5–20 |
| Glucose, liquid | | Up to 50 |
| Guar gum | | 1–10 |
| Hydroxyethyl | Cellosize | 2–6 |
| Hydroxypropyl cellulose | Klucel | 2–6 |
| Hydroxypropylmethyl cellulose | Methocel, Pharmacoat | 2–5 |
| Magnesium aluminum silicate | Pharmasorb, Veegum | 2–10 |
| Maltodextrin | Glucidex, Lycatab, Maltrin | 2–10 |
| Methylcellulose | Celacol, Methocel | 1–5 |
| Polydextrose | Litesse | |
| Polyethylene oxide | Polyox | 5 |
| Povidone | Kollidon, Plasdone | 0.5–05 |
| Sodium alginate | Manucol | 1–3 |
| Starch paste | | 5–25 |

| | |
|---|---|
| Sucrose (syrup) | Up to 70 |
| Water | |

## *Disintegrants*

A disintegrant is added to most tablet formulations to facilitate a breakup or disintegration of the tablet when it contacts water in the gastrointestinal tract. Disintegrants may function by drawing water into the tablet, swelling, and causing the tablet to burst apart. Such tablet fragmentation may be critical to the subsequent dissolution of the drug and to the attainment of satisfactory drug bioavailability. Starch USP and various starch derivatives are the most common disintegrating agents. They also have the lowest cost. Starch is typically used in a concentration range of 5 to 20% of tablet weight. Such modified starches as Primogel and Explotab, which are low substituted carboxylmethyl starches, are used in lower concentrations (1 to 8%, with 4% usually reported as optimum). Various pre-gelatinized starches are also employed as disintegrants, usually in a 5% concentration (Table 13.4).

Clays such as Veegum HV and bentonite have been used as disintegrants at about a 10% level. Such use of these materials is limited unless the tablets are colored, since the clays produce an off-white appearance. The clays are typically less effective as disintegrants than some of the newer modified polymers and starches, which can increase in volume in the presence of water by 200 to 500%.

The disintegrating characteristics of microcrystalline cellulose have been reported previously in this chapter.

The disintegrating agent may be mixed at two stages (i) during the formation of granules, prior to wetting with the granulating fluid—intragranular and (ii) at the second mixing stage during compaction of granules into tablets—extragranular. Extragranular disintegrating agents cause the tablet to disintegrate quickly into the granules, while intragranular disintegrants break down the granules, giving a finer product. For many years, starch was the disintegrating agent of choice. Recently “super disintegrants” have been introduced, which markedly reduce tablet disintegration time. Such substances include sodium starch glycolate, crospovidone, croscarmellose, and polacrilin potassium (Table 13.4).

**Table 13.4:** Disintegrants used in tablet formulation

| Disintegrant | Proprietary name | Concentration used (%) |
|---|---|---|
| Alginic acid | | 2–10 |
| Carboxymethylcellulose sodium | | 1–5 |
| Cellulose, microcrystalline | Avicel, Emcocel, Vivacel | Up to 10 |
| Croscarmellose sodium* | Ac-di-Sol, Solutab | 0.5–5 |
| Crospovidone* | Kollidon CL, Polyplasdone XL | 2–5 |
| Docusate sodium | | 0.5–1 |
| Guar gum | | 2–8 |
| Magnesium aluminum silicate | Veegum | 2–10 |
| Methylcellulose | Celacol, Methocel | 2–10 |
| Polacrilin potassium* | Amberlite | 2–10 |
| Poloxamer | | 5–10 |
| Povidone | Kollidon, Plasdone | 0.5–5 |
| Sodium alginate | Manucol | 2.5–10 |
| Sodium glycine carbonate | | |
| Sodium lauryl sulfate | Empicol | 0.5–2 |
| Sodium starch glycolate* | Explotab, Primojel | 2–8 |
| Starch | | 2–10 |
| Starch, pregelatinized | Lycatab, Pharma-Gel, Pre-Jel, Sepistab, | 5–10 |

* Used as superdisintergrant

## *Lubricants, Antiadherents and Glidants*

These three classes of materials are typically described together because they have overlapping functions. A material that is primarily described as an antiadherent is typically also a lubricant, with some glidant properties as well. The differentiation between these terms is as follows: Lubricants are intended to reduce the friction during tablet ejection between the walls of the tablet and the walls of the die cavity in which the tablet was formed. Antiadherents have the purpose of reducing sticking or adhesion of any of the tablet granulation or powder to the faces of the punches or to the die wall. Glidants are intended to promote flow of the tablet granulation or powder materials by reducing friction between the particles.

Lubricants act by interposing an intermediate layer between the tablet and the die wall. Since they are required to act at the material/tooling interface, lubricants should be incorporated at the final mixing step just before tablet compression.

In addition to the lubricants listed in Table 13.5, hydrocarbon oils such as mineral oil have been employed by application to granulation as a fine spray, either directly or in a solvent solution. The problem with using this type of lubricant is the production of oil spots. The most widely used lubricants have been stearic acid and various stearic acid salts and derivatives. Calcium and magnesium stearate are the most common salts employed. Stearic acid is a less effective lubricant than these salts and also has a lower melting point. Activity of metallic salts of stearic acid is believed to derive from adhesion of the polar metallic portion of the lubricant to the powder particle surface. As a consequence, the hydrocarbon portion of the lubricant becomes oriented away from the surface. Thus, a nonpolar layer is presented to adjacent powder particles and the press tooling responsible for lubrication.

Talc is probably the second most commonly used tablet lubricant, historically. Most talc samples are found to contain trace quantities of iron, and talc should be added carefully in any formulation containing a drug whose breakdown is catalyzed by the presence of iron. The higher-molecular-weight polyethylene glycols and certain polymeric surfactants have been used as water-soluble lubricants. These materials are much less effective as lubricants, however, than the materials previously cited. Since lubrication is basically a coating process, the finer the particle size of the lubricant, the more effective the lubricant action is likely to be. However, for water insoluble lubricants, this has two deleterious consequences. The first is that

each powder particle presents a hydrophobic exterior, slows dissolution, and has been shown to cause bioavailability problems. The second consequence is that direct contact between powder particles is, at least in part, replaced by contact between adjacent hydrocarbon layers and thus the tablet structure is weakened.

**Table 13.5:** Lubricants and glidants used in tablet formulation

| **Lubricant** | **Proprietary name** | **Concentration used (%)** |
|---|---|---|
| Calcium stearate | | 0.5–2 |
| Fumaric acid | | 5 |
| Glyceryl behenate | | 0.5–4 |
| Glyceryl palmitostearate | Precirol | 0.5–5.0 |
| Hydrogenated vegetable oil | Lubritab, Sterotex | 1–6 |
| Magnesium lauryl sulfate | | 1–2 |
| Magnesium stearate | | 0.25–5 |
| PEG 4000 or 6000 | Macrogols, Carbowax | 2–5 |
| Sodium lauryl sulfate | Empicol, Stearowet | 1–2 |
| Sodium stearyl fumarate | Pruv | 0.5–2.0 |
| Starch | | 2–10 |
| Stearic acid | | 1–3 |
| Talc | | 1–10 |
| Zinc stearate | | 0.5–2 |
| **Glidant** | | |
| Calcium silicate | | 0.5–2 |
| Cellulose, powdered | Elcema, Solka, Floc | 1–2 |
| Magnesium carbonate | | 1–3 |
| Magnesium oxide | | 1–3 |
| Magnesium silicate | | 0.5–2 |

| | | |
|---|---|---|
| Silicon dioxide, colloidal | Aerosil, Cab-o-Sil | 0.05–0.5 |
| Starch | | 2–10 |
| Talc | | 1–10 |

The hydrophobic properties of hydrocarbon based lubricants can be countered to a certain extent by the inclusion of a wetting agent such as sodium lauryl sulfate into the formulation. Alternatively, more soluble hydrophilic materials such as polyethylene glycols have been investigated. However, water soluble lubricants do not appear to be as efficient in lubricating tablet formulations as their water insoluble counterparts.

As previously noted, most of the materials listed as lubricants, with the possible exception of those that are water-soluble, also function as antiadherents. Talc, magnesium stearate, and starch as well as starch derivatives possess antiadherent properties. In addition, various colloidal silicas have been used as antiadherents.

Materials used as glidants, or flow promoters, are typically talc at a 5% concentration, corn starch at a 5 to 10% concentration, or colloidal silicas such as Cab-O-Sil, Syloid, or Aerosil in 0.25 to 3% concentrations. Glidant is thought to act by lodging in the surface irregularities of the granule, forming a more rounded structure and hence reducing interparticulate friction. Colloidal silica has the added advantage of acting as a moisture scavenger, thereby providing a drier environment.

## *Colors, Flavours and Sweeteners*

The use of colors and dyes in tablet making has served three purposes over the years: disguising of off-color drugs, product identification, and production of a more elegant product. With the continual decertification of many synthetic dyes, pharmaceutical manufacturers are becoming quite concerned as to how future tablet formulations will be colored. The availability of natural vegetable colors is limited, and these colors are often unstable. Two forms of color have typically been used in tablet preparation. These are the FD&C and D&C dyes—which are applied as solutions, typically in the granulating agent—and the lake forms of these dyes. Lakes are dyes that have been absorbed on a hydrous oxide and usually are employed as dry powders for coloring. In addition to concerns regarding

possible delisting in the future, several other precautions should be considered when colors are employed. When using water-soluble dyes, pastel shades usually show the least mottling from uneven distribution in the final tablet. When wet granulation is employed, care should be taken to prevent color migration during drying. In any colored tablet, the formulation should be checked for resistance to color changes on exposure to light. Various artificial light sources are available that simulate the ultraviolet spectrum of sunlight. Methods of quantifying color are given later in the chapter under the heading, “Organoleptic Properties.”

Flavors are usually limited to chewable tablets or other tablets intended to dissolve in the mouth. In general, flavors that are water-soluble have found little acceptance in tablet making because of their poor stability. Flavor oils are added to tablet granulations in solvents, are dispersed on clays and other absorbents, or are emulsified in aqueous granulating agents. Various dry flavors for use in pharmaceutical products are also available from flavor suppliers. Usually, the maximum amount of oil that can be added to a granulation without influencing its tabletting characteristics is 0.5 to 0.75%.

The use of sweeteners is primarily limited to chewable tablets to exclude or limit the use of sugar in the tablets. Various sugars used as tablet excipients have been described earlier. Mannitol is reportedly about 72% as sweet as sucrose. Until recently, saccharin was the only artificial sweetener available. This material is about 500 times sweeter than sucrose. Its major disadvantages are that it has a bitter aftertaste and has been reported to be carcinogenic. A new artificial sweetener that is expected to largely replace saccharin is aspartame. The primary disadvantage of aspartame is its lack of stability in the presence of moisture. When aspartame is used in a formulation, e.g. a chewable tablet with hygroscopic components, it will be necessary to determine its stability under conditions in which the product can adsorb atmospheric moisture.

Examples of tablet formulations are shown in the Appendices to this chapter. Not only do the formulations illustrate the use of common ingredients, but they also illustrate the use of the ingredients in tablets to be made by wet granulation, dry granulation, and direct compression processes.

It has previously been noted that while the excipients are the inactive part of a tablet formulation, they have a direct influence on the quality and effectiveness of the final product. Figure 13.1 describes, for example, the

influence of compression force on disintegration time for various direct compression materials. Some materials have a maximum disintegration time of no higher than 200 to 250 sec, regardless of the compression force applied over the range studied. One material rapidly increased in disintegration time to over 500 sec at a compression force of less than 1000 kg. Similar relationships between important tablet properties and processing characteristics can be shown for many other tablet excipients.

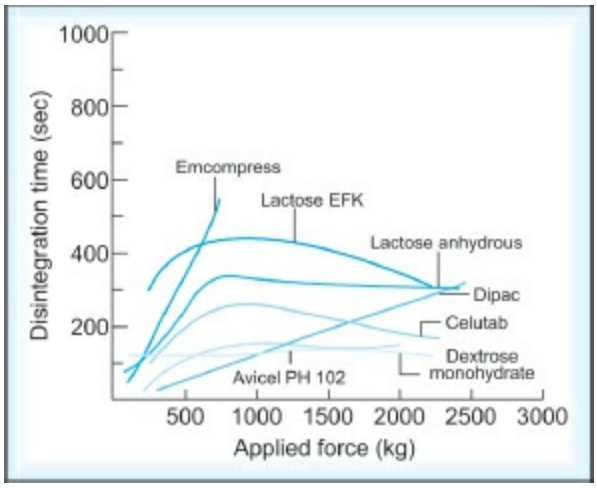



**Fig. 13.1:** Disintegration time *vs* applied force for tablets prepared from various direct compression diluents

Another important consideration that many pharmaceutical formulators must consider in tablet formulation is the worldwide acceptability of their formulation components. An excipient used in the United States, for example, may not be permitted in a major market area such as Japan or Europe, or vice versa. Companies with major international markets strive to have tablet formulations that are equally acceptable around the world and that contain components that are not likely to be delisted in any country.

## TABLET MANUFACTURING

The manufacture of granulations for tablet compression may follow one or a combination of three established methods: the direct compression, compression granulation, and wet granulation. Table 13.6 outlines the combination of properties that necessitate or allow the various processing options whereas Fig. 13.2 compares the type and number of processing steps commonly required with each technique. A consideration of the important aspects of these processes illustrates the advantages and disadvantages of each.

**Table 13.6:** Schematic for process selection

| Drug flow | Drug compressibility | Drug dose | Excipient concentration | Tablet reparation methodology |
|---|---|---|---|---|
| Good | Good | High | Low | Direct compression |
| Good | Bad | – | – | Wet granulation |
| Good/Bad | Good/Bad | Low | High | Direct compression |
| Bad | Good | – | – | Dry granulation |
| Bad | Bad | – | – | Wet granulation |

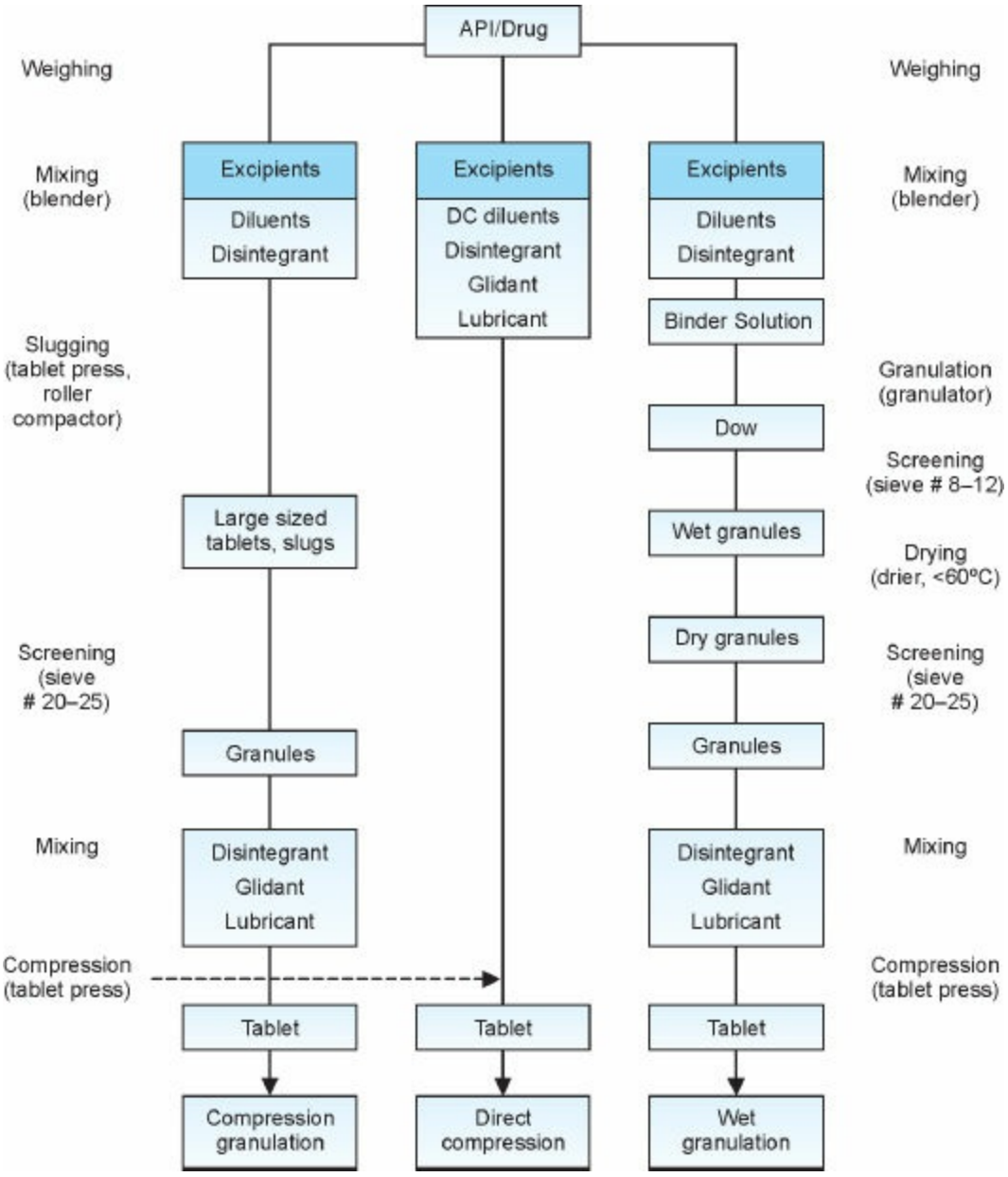



**Fig. 13.2:** Processing steps commonly required in the various tablet preparation techniques

## Direct Compression Method

There are a few crystalline substances, such as sodium chloride, sodium bromide, and potassium chloride that may be compressed directly. The vast majority of medicinal agents are rarely so easy to tablet, however. In addition, the compression of a single substance may produce tablets that do not disintegrate. If disintegration is a problem, other components are needed, which in turn may interfere with the compressibility of the active ingredient and thus minimize the usefulness of the method. Most materials possess relatively weak intermolecular attraction or are covered with films of adsorbed gases that tend to hinder compaction. Thus, most large-dose drugs do not lend themselves to this process. With many other drugs having small doses, uniform blends of the drug and coarser direct compression diluents cannot be achieved, which makes this process impractical. However, the use of compressible diluents with many moderate-dose drugs makes this process the most streamlined method of tablet manufacture.

A directly compressible diluent is an inert substance that may be compacted with little difficulty and may compress even when quantities of drugs are mixed with it. Compression capacity is still maintained when other tablet materials necessary for flow, disintegration, and so forth are blended in. Directly compressible excipients are examined in detail previously in this chapter (Table 13.2). Direct compression materials, in addition to possessing good flow and compressibility, must be inert, tasteless, reworkable, able to disintegrate, and inexpensive.

The most important advantage of the direct compression process is its simplicity, low labor input and hence economy. Being a dry process, risk of deterioration of the active ingredient, is decreased. A further advantage is that tablets disintegrate into their primary particles rather than granular aggregates. The resultant increased surface area available for dissolution should result in faster drug release.

Even though direct compression has some important advantages (low labor input, a dry process, fewest processing steps) there are some limitations to the technique:

1. Differences in particle size and bulk density between the drug and diluent may lead to stratification within the granulation. The stratification may then result in poor content uniformity of the drug in the compressed

tablet. The stratification and resultant content uniformity problems are of special concern with low-dose drugs.

2. A large-dose drug may present problems with direct compression if it is not easily compressible by itself. To facilitate compression, noncompressible large-dose drugs, which are usually restricted to about 30% of a direct compression formula, could require an amount of diluent so large that the resultant tablet is costly and difficult to swallow.
3. In some instances, the direct compression diluent may interact with the drug. A good example of such a reaction is that which occurs between amine compounds and spray-dried lactose, as evidenced by a yellow discoloration.
4. Because of the dry nature of direct compression, static charge buildup can occur on the drug during routine screening and mixing, which may prevent a uniform distribution of the drug in the granulation.

The equipment and procedures used in direct compression are basically screening or milling and mixing. These topics are covered in Chapter 1, “Mixing,” and Chapter 2, “Milling.”

## Dry or Compression Granulation Method

Compression granulation has been used for many years, and is a valuable technique in situations where the effective dose of a drug is too high for direct compaction, and the drug is sensitive to heat, moisture, or both, which precludes wet granulation. Many tablet formulations such as those of aspirin and vitamin are prepared by compression granulation.

Compression granulation involves the compaction of the components of a tablet formulation by means of a tablet press or specially designed machinery, followed by milling and screening, prior to final compression into a tablet. When the initial blend of powders is forced into the dies of a large-capacity tablet press and is compacted by means of flat-faced punches, the compacted masses are called *slugs*, and the process is referred to as "slugging."

In slugging, large tablets (e.g. 2-in. diameter) are made using very heavy-duty machines. These tablets may not have good fill weight uniformity, but this does not matter, because they are broken up in the next step by coarse milling.

The slugs are then screened or milled to produce a granular form of tabletting material, which now flows more uniformly than the original powder mixture. When a single slugging process is insufficient to confer the desired granular properties to the material, the slugs are sometimes screened, slugged again, and screened once more.

Slugging is just an elaborate method of subjecting a material to increased compression time. The act of slugging followed by screening and subsequent compression of the particles is roughly equivalent to an extended dwell time during compression in a tablet machine. The two or more times that the material is subjected to compaction pressures causes a strengthening of the bonds that hold the tablet together. The resultant granules also increase the fluidity of these powder mixtures, which by themselves do not flow well enough to fill the dies satisfactorily.

As shown in Fig. 13.2, the compression granulation method requires less equipment and space than wet granulation, and eliminates the addition of moisture and the application of heat, as found in the wet massing and drying steps of the wet granulation method.

## *Equipment*

On a large scale, compression granulation can also be performed on a specially designed machine called a *roller compactor*. Roller compactors are capable of producing as much as 500 kg per hour or more of compacted ribbon-like material, which can then be screened or milled into a granulation suitable for compression into tablets.

Roller compactors, utilize two rollers that revolve toward each other (Fig. 13.3). By means of a hydraulic ram forcing one of the rollers against the other, the machine is capable of exerting known fixed pressures on any powdered material that flows between the rollers. Powdered material is fed between the rollers by a screw conveyor system. After passing through the rollers, the compacted mass resembles a thin wide ribbon that has fallen apart into large segments. These are equivalent to the slugs produced by the slugging process. The segments are then screened or milled for the production of granules.

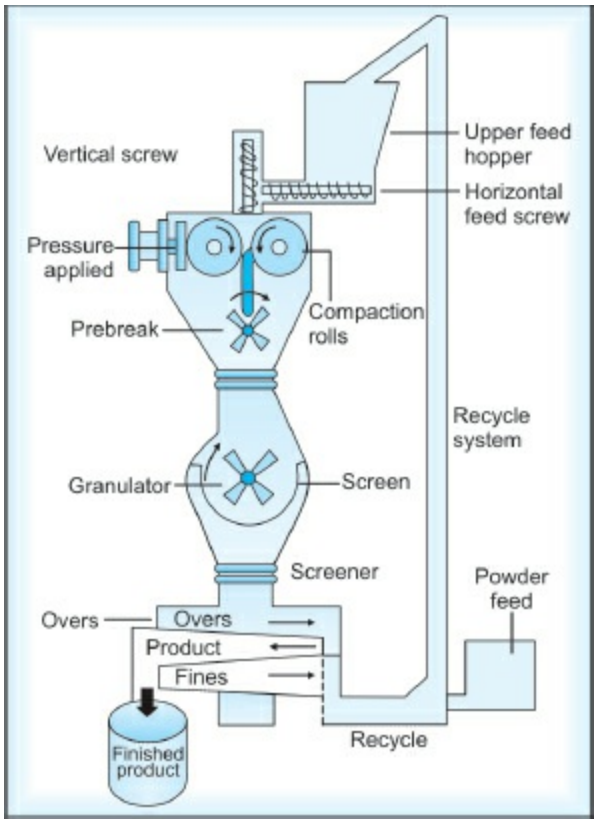



**Fig. 13.3:** Schematic diagram of a Chilsonator roller compactor in a granulation production system

The compaction force of the roller compactor is controlled by three variables (1) the hydraulic pressure exerted on the compaction rolls, (2) the rotational speed of the compaction rolls and (3) the rotational speed of the feed screws. The roll speed and the feedscrew speed have the greatest effect on the compaction process. The feed screws on most modern compactors consist of a variable-speed horizontal and vertical screw. The horizontal screw picks up the powder from the hopper and maintains a continuous flow to the vertical screw. The vertical screw delivers the powder to the compaction rolls. The vertical screw speed is critical for uniform compaction. It serves to deaerate the powder and maintains a constant flow onto the compaction rolls. Any variation in deaeration or load causes extreme changes in the compact. The vertical feed screw is usually set so that it delivers more material than the compaction rolls accept, assuring constant loading during the compaction process. The speed of the compaction rolls controls the pressure dwell time, which has a great effect on the density and hardness of the compact.

A standard procedure for testing compaction uniformity and machine capacity is to select a hydraulic pressure in the midranges of the equipment. Set the compaction roll at the slowest speed, and set the feed screw at the highest speed. If the powders are compactible in the first pass, the machine will overload. When this happens, the compaction roll speed should be increased until the loading is constant. Maximum throughput is achieved at this setting for the material being tested. If no overloading occurs, the powder should be passed through a second time, using the same procedure. The roller compactor offers the advantages over the slugging process of increased production capacity, greater control of compaction pressure and dwell time, and no need for excessive lubrication of the powder.

There are many modifications available on roll compactors. Roll designs cover a complete range from smooth to curve and serrated surfaces. The shapes and sizes of the screw feed assembly are available in a wide range of designs. Most compactors can be fitted with liquid-cooled rolls and chambers. All manufacturers of roller compactors have pilot plant facilities and offer complete testing programs. Trial runs are advisable, so that the compactor is suitable for the materials to be compacted.

## Wet Granulation Method

The wet granulation technique uses the same preparatory and finishing steps (screening or milling, and mixing) as the two previously discussed granulation techniques. The unique portions of wet granulation process involve the wet massing of the powders, wet sizing or milling, and drying. The theory, equipment, and methods associated with drying are discussed in Chapter 4.

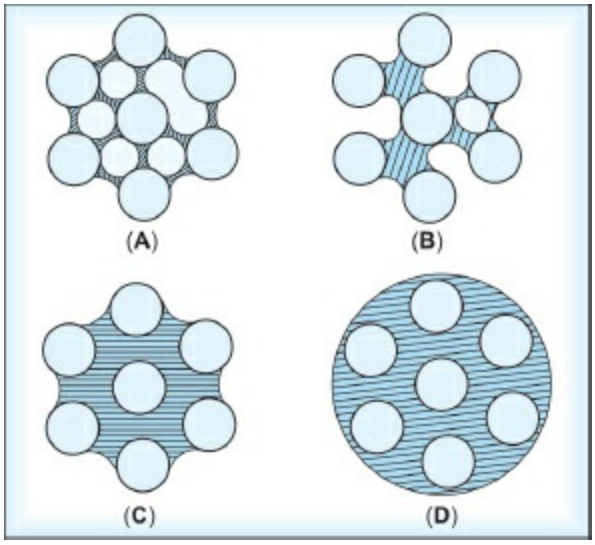
Wet granulation forms the granules by binding the powders together with an adhesive, instead of by compaction. The wet granulation technique employs a solution, suspension, or slurry containing a binder, which is usually added to the powder mixture; however, the binder may be incorporated dry into the powder mix, and the liquid may be added by itself.

Addition of a granulating liquid to a mass of powder may be characterized in a series of stages as illustrated in Fig. 13.4. If the powder particles are wetted during the initial stage (Fig. 13.4A), liquid films will be formed on their surface and may combine to produce discrete liquid bridges at points of contact. The surface tension and negative capillary pressure in such bridges provide the cohesive force and result in a condition called the *pendular* state, which has comparatively low mechanical strength.

As the liquid content increases, several bridges may coalesce, giving rise to the *funicular* state (Fig. 13.4B), and a further modest increase in the strength of the moist granule. Eventually, as more liquid is added and the mass is kneaded to bring particles into closer proximity, the void spaces within the granule are entirely eliminated. At this point, bonding is affected by interfacial forces at the granule surface and by a negative capillary pressure throughout the interior liquid-filled space, a condition referred to as the *capillary* state (Fig. 13.4C). Further addition of liquid results in *droplet* formation (Fig. 13.4D), in which the particles are still held together by surface tension, but without intragranular forces; such structures are weaker. The capillary state coincides with the maximum strength of the wet granules, and optimization of many granulation processes involves ensuring that this state has been achieved. For example, granulation equipment can be instrumented with torque measuring devices, which sense the change in agitator power requirements at the capillary stage.

The liquid plays a key role in the granulation process. Liquid bridges are

developed between particles, and the tensile strength of these bonds increases as the amount of liquid added is increased. These surface tension forces and capillary pressure are primarily responsible for initial granule formation and strength. Once the granulating liquid has been added, mixing continues until a uniform dispersion is attained and all the binder has been activated.




**Figs 13.4A to D:** Stages in the development of moist granules as the proportion of liquid is increased: (A) Pendular; (B) Funicular; (C) Capillary; (D) Droplet.

The method of introducing the binder depends on its solubility and on the components of the mixture. Since, in general, the mass should merely be moist rather than wet or pasty, there is a limit to the amount of solvent that may be employed. Therefore, when only a small quantity is permissible, the binder is blended in with the dry powders initially; when a large quantity is required, the binder is usually dissolved in the liquid. The solubility of the binder also has an influence on the choice of methods, since the solution should be fluid enough to disperse readily in the mass.

During granulation, particles and agglomerates are subjected to consolidating forces by action of machine parts and of interparticulate forces. Granulation in large blenders requires 15 min to an hour. The length of time depends on the wetting properties of the powder mixture and the granulating fluid, and upon the efficiency of the mixer. A rough way of determining the end point is to press a portion of the mass in the palm of the hand; if the ball crumbles under moderate pressure, the mixture is ready for the next stage in

processing, which is wet screening.

The wet screening process involves converting the moist mass into coarse, granular aggregates by passage through a hammer mill or oscillating granulator, equipped with screens having large perforations. The purpose is to further consolidate granules, increase particle contact points, and increase surface area to facilitate drying. Overly wet material dries slowly and forms hard aggregates, which tend to turn to powder during subsequent dry milling. There are many instances in which wet milling may be omitted, with a considerable saving of time. The formulator should be alert to these opportunities and not follow the old method blindly.

A drying process is required in all wet granulation procedures to remove the solvent that was used in forming the aggregates and to reduce the moisture content to an optimum level of concentration within the granules. During drying, interparticulate bonds result from fusion or re-crystallization and curing of the binding agent, with Van der Waals forces playing a significant role.

After drying, the granulation is screened again. The size of the screen depends upon the grinding equipment used and the size of the tablet to be made.

The use of volatile or inflammable solvents for wet granulation creates other problems. Safety considerations demand that at a minimum, the work areas be well-ventilated to reduce direct toxic effects or to keep the solvent vapor concentration below explosion limits. Also, all equipment should be electrically grounded to prevent sparks that could initiate explosions. Explosion-proof or explosion-resistant motors may also be required. If solvent granulating systems are to be used, the entire process should be thoroughly discussed, and the facilities should be inspected by the company’s safety engineer.

Exhausting solvent vapors or drying granulations made with solvents also requires special precautions. Environmental Protection Agency (EPA) regulations limit the amount of solvent vapors that can be exhausted into the atmosphere. Such EPA limits could require recovery or burning of the solvent vapors, which are expensive operations. Ovens employed for drying granulations wetted with explosive solvents should employ high airflow rates, to stay well below vapor explosive limit concentrations in air. Such ovens should also contain appropriate controls to prevent explosions due to

accumulation of vapors following a power outage or during later resumption of power.

## *Equipment*

When traditional equipment is used in the conventional wet granulation scheme, the entire process is labor-intensive and timeconsuming. The equipment used for granulation is not highly effective for dry mixing. Therefore, in many instances, a different mixer is used for dry mixing prior to granulating. Examples are sigma blade and planetary mixers. Granulating mixers are slow, are generally poor powder mixers, and require care for even addition of granulating liquids. Also, considerable time is needed to distribute the binder properly throughout the mass.

While some tablets are still made in the traditional manner, newer equipment has been developed that can accomplish both dry mixing and wet granulation efficiently and in much less time. New mixer/granulators such as Lödige, Diosna and Gral, allow several processes of wet granulation to be conducted in rapid succession or to be combined in one piece of equipment. These high speed mixer/granulators are provided with secondary chopper blades which function as a lump and agglomerate breaker so that sieving is no longer an essential prerequisite of powder blending. When liquid granulating agents are added to dry powders, the liquid enters the mixer under pressure through the liquid nozzle immediately above the chopper assembly, or assemblies, and is immediately dispersed. Power usage is often reflected in the readings of an ammeter or wattmeter mounted on the equipment and may be useful in helping to identify the proper end point for the wet granulation process. The description of these equipments is discussed in Chapter 1.

# TABLET COMPRESSION OPERATION

## Tablet Compression Machines

Tablets are made by compressing a formulation containing a drug or drugs with excipients on stamping machines called *presses*. Tablet compression machines or tablet presses are designed with the following basic components:

1. Hopper(s) for holding and feeding granulation to be compressed.
2. Dies that define the size and shape of the tablet.
3. Punches for compressing the granulation within the dies.
4. Cam tracks for guiding the movement of the punches.
5. A feeding mechanism for moving granulation from the hopper into the dies.

Tablet presses are classified as either singlepunch or multi-station rotary presses. Figure 13.5 illustrates in cross-section the compression process on a single punch machine. Note that all of the compression is applied by the upper punch, making the single punch machine a “stamping press.”

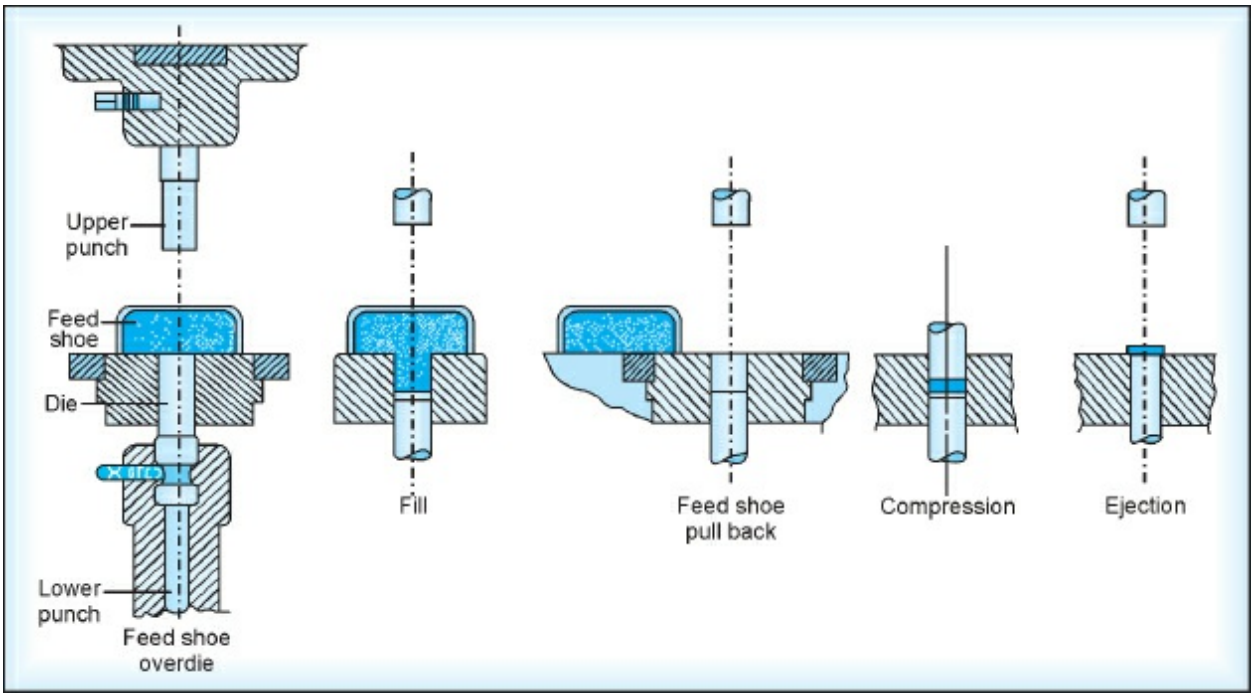



**Fig. 13.5:** The compression cycle of a single-punch tablet press

Multi-station presses are termed *rotary* because the head of the tablet machine that holds the upper punches, dies, and lower punches in place rotates. As the head rotates, the punches are guided up and down by fixed cam tracks, which control the sequence of filling, compression, and ejection. The portions of the head that hold the upper and lower punches are called the upper and lower turrets respectively, and the portion holding the dies is called the *die table*. At the start of a compression cycle (Fig. 13.6), granulation stored in a hopper (not shown), empties into the feed-frame (A), which has several interconnected compartments (Fig. 13.7). These compartments spread the granulation over a wide area to provide time for the dies (B) to fill (Fig. 13.6). The pull-down cam (C) of Fig. 13.6 guides the lower punches to the bottom of their vertical travel, allowing the dies to overfill. The punches then pass over a weight-control cam (E), which reduces the fill in the dies to the desired amount. A wipe-off blade (D) at the end of the feed-frame removes the excess granulation and directs it around the turret and back into the front of the feed-frame. Next, the lower punches travel over the lower compression roll (F) while simultaneously the upper punches ride beneath the upper compression roll (G). The upper punches enter a fixed distance into the dies, while the lower punches are raised to squeeze and compact the granulation within the dies. To regulate the upward movement of *the* lower punches, the height of the lower pressure roll is changed. After the moment of compression, the upper punches are withdrawn as they follow the upper punch raising cam (H); the lower punches ride up the cam (I), which brings the tablets flush with or slightly above the surface of the dies. The exact position is determined by a threaded bolt called the *ejector knob*. The tablets strike a sweep-off blade affixed to the front of the feed-frame (A) and slide down a chute into a receptacle. At the same time, the lower punches re-enter the pull down cam (C), and the cycle is repeated.

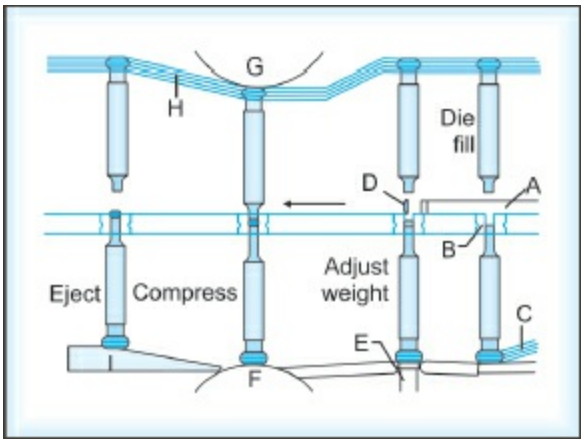



**Fig. 13.6:** The compression cycle of a rotary tablet press. (*See* text far explanation of lettered labels)

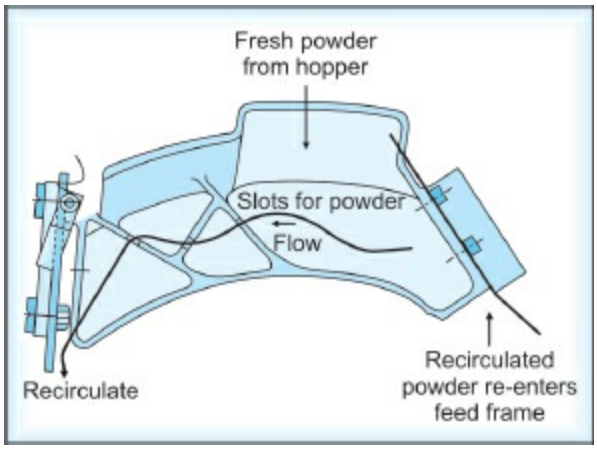



**Fig. 13.7:** Granulation flow in an open feed frame of a rotary tablet press

Many production tablet machines are designed so that the compression cycle is accomplished more than once (requiring additional granulation hoppers, feed frames, cam tracks, and compression rolls) while the machine head makes a single revolution.

All other parts of a tablet press are designed to control the functioning of the components just listed. Such features as capacity, speed, maximum weight, and pressure vary with the design of the equipment, but the basic elements remain essentially the same. Although tablet compressing machinery has undergone numerous mechanical modifications over the years, the compaction of materials between a pair of moving punches within a stationary die has remained unchanged. The principal modification from earlier equipment has been an increase in production rate rather than any

fundamental change in the process. Better control and simplification have been corollary benefits.

In recent years, there has been a change by manufacturers from activities concerned with production rate to problems of process improvement and control. Growth of governmental and pharmacopoeial tests related to inter-tablet weight and potency variation, as noted earlier in the chapter, have created some of the new requirements for tablet compressing machinery. As tablet production rates have increased with modern equipment, for example, the need for automatic tablet weight control independent of operator vigilance has become a matter of increasing concern. This topic is discussed later, in the section "Auxiliary Equipment."

Tabletting presses vary principally in the number of tooling stations available for compression and in special application features. Table 13.7 tabulates the maximum and minimum tablet manufacturing output capable within the various press models of several manufacturers.

A tablet machine's output is regulated by three basic characteristics of its design:

**Table 13.7:** Characteristics of some selected rotary tablet presses

| Manufacturer | Number of stations available | | Output (tablets per minute) | | US Representative |
|---|---|---|---|---|---|
| | Min | Max | Min | Max | |
| Colton | 12 | 90 | 480 | 16,000 | Vector Corp.<br>Marion, IA 52302 |
| Wilhelm Fette,<br>Gmbh Hamburg<br>W. Germany | 20 | 55 | 300/900 | 3300/8250 | Raymond Automation<br>Company, Inc.<br>Norwalk, CT 06856 |
| Kilian & Co.,<br>Gmbh Koln<br>W. Germany | 14 | 67 | 140/383 | 1083/10,000 | INPPEC<br>Fairfield, CT 06430 |
| Manesty Machines Ltd.<br>Liverpool, England | 16 | 69 | 600/1500 | 3330/10,000 | Thomas Engineering<br>Hoffman Estates, IL 60172 |
| Stokes-Merrill | 33 | 65 | 1200/3300 | 3500/10,000 | Stokes-Merrill Division<br>Penwalt Corp.<br>Oak Brook, IL 60521 |
| Korsch Maschinenfabrik<br>Berlin, W. Germany | 20 | 55 | 540/1100 | 1640/5500 | Aeromatic<br>East Towaco, NJ 07082 |
| Hata Ironworks<br>Hon Engineering Co.<br>Osaka, Japan | 28 | 71 | 420/1420 | 1960/7100 | Elizabeth-Hata<br>International Inc.<br>MeKeesport, PA 15132 |

- Number of tooling sets
- Number of compression stations
- Rotational speed of the press

In general, all rotary presses are engineered for fast and economical production of all kinds of tablets. Larger machines can readily produce several million tablets each in a working day, and their performance can be geared to continuous low-maintenance operation. Figure 13.8 is an example of a modern high-speed tabletting machine.

There are many modifications and options that can be obtained from various manufacturers. One modification, which is found on most modern high-speed tablet presses, is the use of hydraulic or pneumatic pressure to control the pressure rolls in place of the older spring type pressure. Either of these alternatives gives a smoother pressure or compressive load force over a longer period of time. Hydraulic or pneumatic pressure is much more accurate and can be set with closer tolerances, which do not change with time or fatigue.


**Fig. 13.8:** The 16-Station rotary tablet press (*Courtesy* of ERWEKA)

Special adaptations of tablet machines allow for the compression of "layered" tablets and coated tablets. Precompression stations are also available to help in compressing difficult granulations. Available with certain Fette machines is a device that chills the compression components to allow for the compression of low-melting-point substances such as waxes, thereby

making it possible to compress products with low melting points, such as suppositories.

There are many basic and optional features available in tablet machines, including some not mentioned in this text. Manufacturers' brochures should be closely checked for available features. One should attend equipment shows, if possible, to obtain up-to-date information on equipment developments. In some instances, test runs on machinery may be made before a final decision to purchase new high-speed tablet equipment or specialized granulation or drying equipment.

## Compression Machine Tooling

As mentioned earlier, the size and shape of a tablet as well as certain identification markings are determined by the compression machine tooling. Each tooling set consists of a die and upper and lower punches. Since each tablet is formed by a tooling set, the tooling must meet many requirements to satisfy the needs of dosage uniformity, production efficiency, and aesthetic appearance.

The terminology used with tooling is illustrated in Fig. 13.9. The most common tools employed are referred to as BB tooling and are 5.25 inches in length, and have a nominal barrel diameter of 0.75 inches and 1-inch head diameter. B tooling is identical to the BB type except that the lower punch is only 3⅜ inches long. D tooling is popular for large tablets, utilizing a 1-inch barrel diameter, 3¼ Winch head diameter, and 5.25-inch length. The dies that are used with the above punches are either a 0.945-inch outside diameter (OD) die capable of making a ⅞-inch round tablet or 1⅜-inch capsule-shaped tablet; or a ⅝-inch OD die capable of handling a ⅝-inch round or ¾-inch capsule shaped tablet.

Several types of steel are normally used in the manufacture of compression tooling. These steels differ in toughness, to withstand the cyclic compacting forces (ductility), and in wear resistance. Unfortunately, no single steel type has a high resistance to abrasive wear and a high ductility. Therefore, the selection of the best steel for a specific application must be based on experience and an accumulated history of the product being tabletted. In the selection of the proper steel for a specific use, one should also consider the shape of the punch tip, whether or not debossing is to be employed on the tooling, the expected compression forces involved, and whether the materials to be processed are abrasive or corrosive.

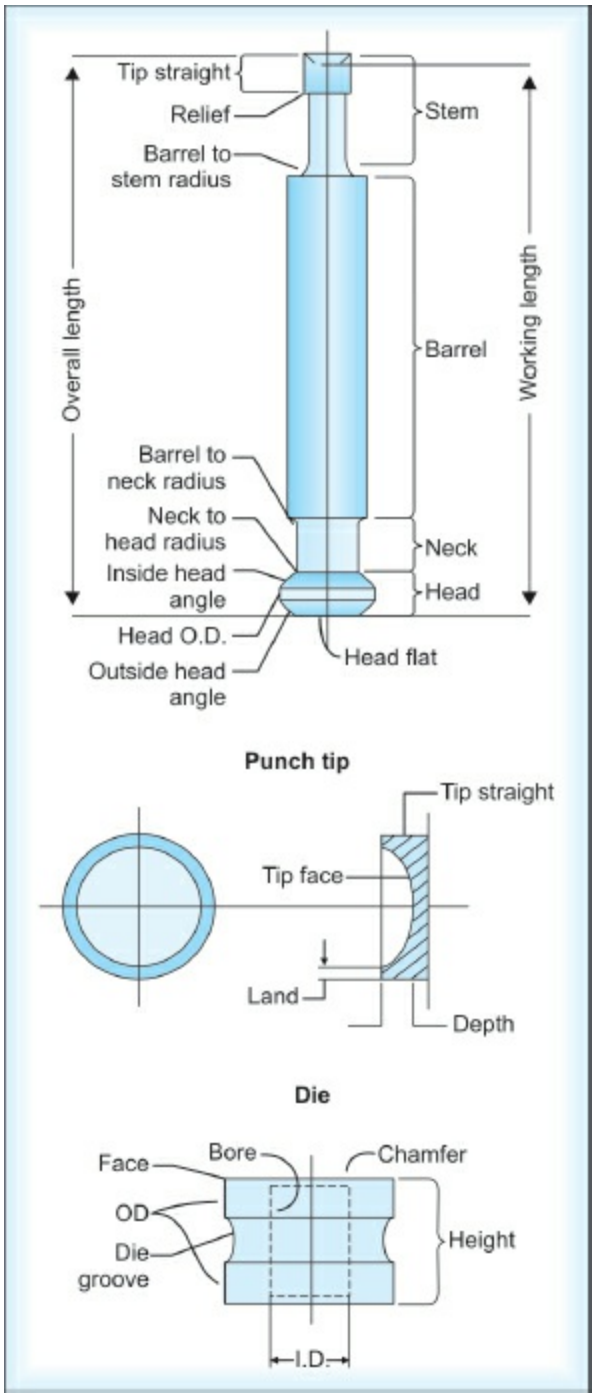



**Fig. 13.9:** Tablet press tooling nomenclature

The size, shape, and contour of a tablet is almost unlimited within the given limits of the specified die size. A survey of the PDR Product Identification section reveals numerous variations on tablet size and shape. In addition, tooling can be made with certain other information to aid in producing a visibly unique tablet product. Company names or symbols, trade

names, dosage strength, or National Drug Code (NDC) numbers can be cut or engraved into a punch face, or the punches may be scored, to produce uniquely embossed or engraved tablets. Even though tooling design would appear to be limitless, certain practical aspects do limit design implementation. Because of the movement of tooling during a compression operation, certain tablet shapes or contour configurations perform better than others. Round tablets perform better than irregularly shaped tooling since they do not require “keying” to maintain the proper upper punch orientation with the die. When the tip on an upper punch is not round, it must not rotate, or it will strike the edge of the die hole as it descends for compression. To prevent this, a slot is cut longitudinally into the barrel of the punch, and a key is inserted. This key protrudes a short distance so that it engages a similar slot cut into the upper punch guides on the tablet press. Lower punches do not need keys because their tips remain within the die bore, which controls the axial movement of the punch. Because keyed punches cannot rotate, wear is distributed unevenly, and punch life is shortened.

When a press is set up with keyed punches, the upper punches are inserted first to determine the placement of the dies. Once the dies are properly aligned and seated, they are locked in place, and the lower punches are inserted. The more curvature that is built into a tablet contour, the more difficult it is to compress, especially if the tablet tends to laminate or cap. The engraving or embossing on a tablet must be designed to be legible, must not add to compression problems, and must fit on the tablet surface. Many considerations, at close tolerances, must be incorporated in tooling design to produce tablets that are uniform and aesthatic. Manufacturing specifications for tooling have been standardized by the Industrial Pharmaceutical Technology Section of the Academy of Pharmaceutical Sciences in its Standard Specifications of Tabletting Tools.

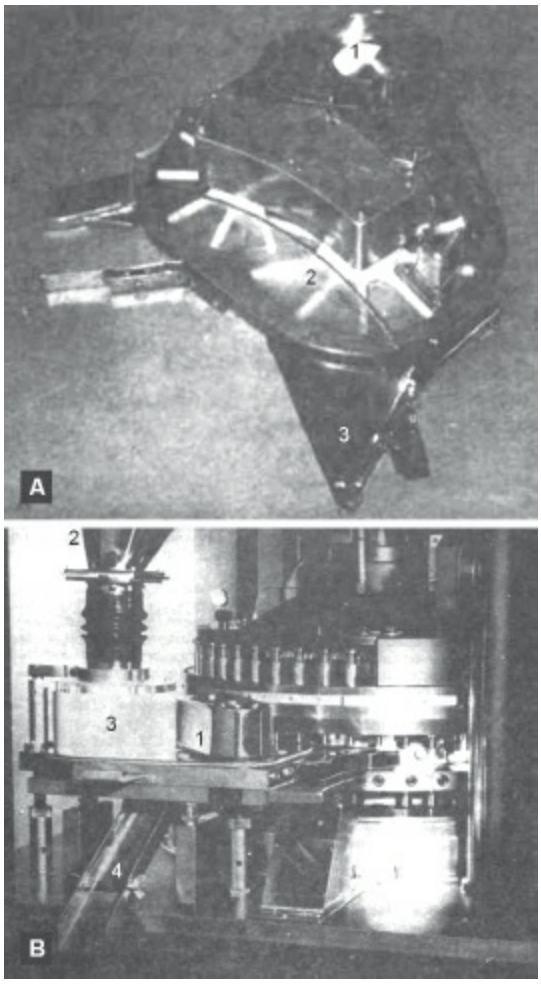
Because of its hard steel structure, tablet tooling may appear to be indestructible. During normal use, however, the punches and dies become worn, and the cyclic application of stress can cause the steel to fatigue and break. Improper storage and handling can readily result in damage that necessitates discarding of an entire tooling set. The punch tips are especially delicate and susceptible to damage if the tips make contact with each other, the dies, or the press turret upon insertion or removal of the tools from the tablet machine. A good tool control system must be employed to maintain the history of each tool set, not only to maintain a constant surveillance of critical

tolerances altered by wear, but also to eliminate product mixups by preventing the wrong tooling from being used for a product.

To avoid tooling damage, compressive loads or pressures at the pressure rolls must be translated into a calculation of pressure at the punch tips. As tablet punch diameter decreases, less force is required to produce the same pressure at the punch face, since the face represents a smaller fraction of a unit area (square inch). The formula for the area of a circle is $\pi r^2$ where r is the radius of the circle. Given a flat punch face, the area of a ¼-inch-diameter punch would thus be $3.14 \times (⅛)^2$ or $3.14 \times$ 1/16, or approximately 1/20 square-inch. If a 1-ton load is being applied by the pressure roll, this area is translated as 2000 pounds on 1/20 square inch, or 40,000 pounds on 1 square inch, a gross overload.

## Auxiliary Equipment

There are some common auxiliary pieces of equipment that increase the efficiency of the tablet compression operation. In many cases, the speed of the die table is such that the dwell time of a die under the feed frame is too short to allow for adequate or consistent gravity filling of the die with granulation. Improper filling of the dies with granulation results in unsatisfactory weight variation and content uniformity of the resulting tablets. A similar result can occur with a poorly flowing granulation. To help alleviate these problems, mechanized feeders can be employed to force granulation into the dies (Fig. 13.10).




**Figs 13.10A and B:** (A) Granulation feeding device: (1) Granulation input port from tablet machine feed hopper, (2) Rotating feed fingers, (3) Compressed tablet scrape-off blade, (B) Granulation feeding device mounted

on a rotary tablet machine: (1) Tablet machine turret, (2) Tablet machine hopper, (3) Feeding device, (4) Compressed tablet output chute

The high tablet output rates of modern presses demand that the granulation hoppers be refilled at frequent intervals; the larger the tablet is, the more frequently the hopper needs to be replenished. Allowing a tablet machine to run “dry” results in a series of rapidly degenerating and unacceptable events. First, low-weight tablets and tablets with poor weight variation are produced. Then, the soft granulation is unable to be formed into tablets. Finally, the tooling is usually ruined, particularly with thin tablets, by the punches being forced together without any granulation between them. Because of the relatively low volume of press hoppers, the filling of hoppers by hand on high-speed presses is inefficient, increases the risk of punch damage, and can contribute to weight variation problems. Therefore, mechanized equipment has been developed to load granulation into the press hoppers.

A popular method of handling large quantities of material is to place bulk granulation containers directly above tabletting machines to gravity-feed the granulation into hoppers. This can be accomplished by several means. Bulk granulation containers can be placed on floors above a tablet machine, and granulation can then be directed through openings in the floor into the hoppers. In a similar fashion, granulation containers can be held on mezzanines above tablet machines. If such overhead room is unavailable, hoists and mechanical lifts can be used to elevate granulation containers or material transfer devices directly in position above the press. Granulation level sensors can be used to stop the press automatically when the granulation level drops to a critical level in the hopper.

The high rate of tablet output with modern presses calls for a higher frequency or even continuous monitoring of tablet weight. Electronic monitoring devices, such as the Thomas Tablet Sentinel (Fig. 13.11), Pharma-kontroll, and the Kilian Control System-MC, monitor the force at each compression station, which correlates with tablet weight. These monitors are also capable of initiating corrective actions, altering the amount of die fill to maintain a fixed force, ejecting tablets that are out of specification, counting and documenting the machine operation throughout the run.

**Fig. 13.11:** Thomas Thorns Table Sentinel II (*Courtesy* of Thomas Entering, Hoffman Estates, IL)

In almost all cases, tablets coming off a tablet machine bear excess powder and are run through a tablet deduster to remove that excess.

Numerous unit processes are involved in making tablets, including particle milling, mixing, granulating, drying, compaction, and (frequently) coating. Various factors associated with these processes can seriously affect content uniformity, bioavailability, or stability. Some of these are given in the following list:

**Particle milling**

- Non-uniform milling can lead to segregation or demixing problems.
- Change in the crystalline state during milling can affect solubility.
- Development of electrostatic forces inhibits blending.

**Mixing**

- Non-uniform mixing can lead to segregation or demixing problems.
- Non-homogeneous distribution of drug substance.
- Overblending of water insoluble lubricant lowers dissolution rate.

**Granulation**

- Non-homogeneous distribution of binder and drug gives drug-rich or drug-poor fines.
- Hydrolysis of drug substance due to residual moisture.

- Compaction or uniformity problems due to uneven size of granules.

**Tableting**

- Compaction pressures affect dissolution.
- Poor flow of mix gives poor content uniformity.
- Shearing of lubricant in feed frame lowers dissolution rates.

## EVALUATION/QUALITY CONTROL TEST

During the compression of tablets, in-process tests are routinely run to monitor the process including tests for tablet weight, weight variation, hardness, thickness, disintegration and various evaulations of elegance. The inprocess tests are performed by production and/or quality control (QC) personnel. In addition, many in-process tests are performed during product development by the formulator. Such testing during development has become increasingly important in recent years for process validation purposes. The data supplied by the formulator is usually employed by QC personnel to establish the test limits. At the start-up of a tablet compression operation, the identity of the granulation is verified, along with the set-up of the proper tabletting machine and proper tooling.

To design tablets and later monitor tablet production quality, quantitative evaluations and assessments of a tablet's chemical, physical, and bioavailability properties must be made. Not only could all three property classes have a significant stability profile, but the stability profiles may be interrelated, i.e. *chemical* breakdown or interactions between tablet components may alter *physical* tablet properties, greatly changing the *bioavailability* of a tablet system.

## General Appearance

The general appearance of a tablet, its visual identity and overall “elegance,” is essential for consumer acceptance, for control of lot-to-lot uniformity and general tablet-to-tablet uniformity, and for monitoring trouble-free manufacturing. The control of the general appearance of a tablet involves the measurement of a number of attributes such as a tablet’s size, shape, color, presence or absence of an odor, taste, surface texture, physical flaws and consistency, and legibility of any identifying markings.

## Size and Shape

The size and shape of the tablet can be dimensionally described, monitored, and controlled. A compressed tablet's shape and dimensions are determined by the tooling during the compression process. The thickness of a tablet is the only dimensional variable related to the process. At a constant compressive load, tablet thickness varies with changes in die fill, with particle size distribution and packing of the particle mix being compressed, and with tablet weight, while with a constant die fill, thickness varies with variations in compressive load. Tablet thickness is consistent batch to batch or within a batch only if the tablet granulation or powder blend is adequately consistent in particle size and size distribution, if the punch tooling is of consistent length, and if the tablet press is clean and in good working order.

The crown thickness of individual tablets may be measured with a micrometer, which permits accurate measurements and provides information on the variation between tablets. Other techniques employed in production control involve placing 5 or 10 tablets in a holding tray, where their total crown thickness may be measured with a sliding caliper scale. This method is much more rapid than measurement with a micrometer in providing an overall estimate of tablet thickness in production operations, but it does not as readily provide information on variability between tablets; however, if the punch and die tooling has been satisfactorily standardized and the tablet machine is functioning properly, this method is satisfactory for production work.

Tablet thickness should be controlled within a ±5% variation of a standard value. Any variation in tablet thickness within a particular lot of tablets or between manufacturer's lots should not be apparent to the unaided eye for consumer acceptance of the product. In addition, thickness must be controlled to facilitate packaging. Difficulties may be encountered in the use of unit dose and other types of packaging equipment if the volume of the material being packaged is not consistent. A secondary packaging problem with tablets of variable thickness relates to consistent fill levels of the same product container with a given number of dosage units.

The physical dimensions of the tablet, along with the density of the materials in the tablet formulation and their proportions, determine the weight of the tablet. The size and shape of the tablet can also influence the choice of

tablet machine to use, the best particle size for the granulation, production lot sizes that can be made, the best type of tablet processing that can be used, packaging operations, and the cost to produce the tablet. The shape of the tablet alone can influence the choice of tablet machine used. Shaped tablets requiring “slotted punches” must be run at slower speeds than are possible with round tablets, using conventional punches. Because of the nonuniform forces involved within a tablet during compression, the more convex the tablet surface, the more likely it is to cause capping problems, forcing the use of a slower tablet machine or one with precompression capabilities.

## Unique Identification Markings

Pharmaceutical companies manufacturing tablets often use some type of unique markings on the tablet in addition to color, to aid in the rapid identification of their products. These markings utilize some form of embossing, engraving, or printing. A look into the product identification section of the current Physician's Desk Reference (PDR), provides a quick reference to the multitude of marking variations, both artistic and informational, that can be produced.

The type of informational markings placed on a tablet usually includes the company name or symbol, a product code such as that from the National Drug Code (NDC) number, the product name, or the product potency. In the future, these identifying marks, in conjunction with a greater diversity of tablet sizes and shapes, may provide the sole means of identification of tablets, if the pharmaceutical industry continues to lose the use of approved Food, Drug, and Cosmetic (FD&C) colors.

## Organoleptic Properties

Many pharmaceutical tablets use color as a vital means of rapid identification and consumer acceptance. The color of a product must be uniform within a single tablet (nonuniformity is generally referred to as “mottling”), from tablet to tablet, and from lot to lot. Nonuniformity of coloring not only lacks aesthatic appeal but could be associated by the consumer with nonuniformity of content and general poor quality of the product.

The eye cannot discriminate small differences in color nor can it precisely define color. The eye has limited memory storage capability for color, and the storage of visually acquired data is difficult, which results in people perceiving the same color differently and a single person describing the same color differently at different times. In addition, visual color comparisons require that a sample be compared against some color standard. Color standards themselves are subject to change with time, thus forcing their frequent redefinition, which can lead to a gradual and significant change in acceptable color. Efforts to quantitate color evaluations have used reflectance spectrophotometry, tristimulus colorimetric measurements, and the use of a microreflectance photometer to measure the color uniformity and gloss on a tablet surface.

The presence of an odor in a batch of tablets could indicate a stability problem, such as the characteristic odor of acetic acid in degrading aspirin tablets; however, the presence of an odor could be characteristic of the drug, (vitamins have a characteristic odor), added ingredients (flavoring agents have pleasant odors), or the dosage form (film-coated tablets usually have a characteristic odor).

Taste is important in consumer acceptance of chewable tablets. Many companies utilize taste panels to judge the preference of different flavors and flavor levels in the development of a product. Owing to the subjectiveness of “taste” preference, however, the control of taste in the production of chewable tablets is often simply the presence or absence of a specified taste.

A tablet’s level of flaws such as chips, cracks, contamination from foreign solid substances (e.g. hair, drops of oil, and “dirt”), surface texture (“smooth” versus “rough”), and appearance (“shiny” versus “dull”) may have a zero-defect specification, but the visual inspection techniques used for detecting or evaluating these characteristics are subjective in nature. 

Electronic devices that are currently being developed hold promise for making inspection a more quantitative and reproducible operation.

## Hardness

Tablets require a certain amount of strength, or hardness and resistance to friability, to withstand mechanical shocks of handling in manufacture, packaging, and shipping. In addition, tablets should be able to withstand reasonable abuse when in the hands of the consumer, such as bouncing about in a woman's purse in a partially filled prescription bottle. Adequate tablet hardness and resistance to powdering and friability are necessary requisites for consumer acceptance. More recently, the relationship of hardness to tablet disintegration, and perhaps more significantly, to the drug dissolution release rate, has become apparent. The monitoring of tablet hardness is especially important for drug products that possess real or potential bioavailability problems or that are sensitive to altered dissolution release profiles as a function of the compressive force employed.

Historically, the strength of a tablet was determined by breaking it between the second and third fingers with the thumb acting as a fulcrum. If there was a "sharp" snap, the tablet was deemed to have acceptable strength. More recently, however, tablet hardness has been defined as the force required to break a tablet in a diametric compression test. To perform this test, a tablet is placed between two anvils, force is applied to the anvils, and the crushing strength that just causes the tablet to break is recorded. Hardness is thus sometimes termed the *tablet crushing strength*. Several devices operating in this manner have been and continue to be used to test tablet hardness: the *Monsanto tester*, the *Strong-Cobb tester*, the *Pfizer tester*, the *Erweka tester*, and the *Schleuniger tester*.

One of the earliest testers to evaluate tablet hardness was the Monsanto hardness tester, which was developed approximately fifty years ago. The tester consists of a barrel containing a compressible spring held between two plungers. The lower plunger is placed in contact with the tablet, and a zero reading is taken. The upper plunger is then forced against a spring by turning a threaded bolt until the tablet fractures. As the spring is compressed, a pointer rides along a gauge in the barrel to indicate the force. The force of fracture is recorded, and the zero force reading is deducted from it. To overcome the manual nature of the Monsanto tester and the minute or longer time required to make an individual test, the Strong-Cobb tester was developed about twenty years later. The original design employed a plunger

activated by pumping a lever arm, which forces an anvil against a stationary platform by hydraulic pressure. The force required to fracture the tablet is read from a hydraulic gauge. Later modifications of the Strong-Cobb tester were built with the force applied by air-pressure rather than by a manual pump.

Approximately one decade later, the Pfizer tester was developed and made available to the industry. This tester operates on the same mechanical principle as a pair of pliers. As the plier's handles are squeezed, the tablet is compressed between a holding anvil and a piston connected to a direct force reading gauge. The dial indicator remains at the reading where the tablet breaks and is returned to zero by depressing a reset button. The Pfizer tester became extensively used in comparison to the earlier testers, based on its simplicity, low cost, and the rapidity with which it could be used.

Two testers have been developed to eliminate operator variation. In the Erweka tester, the tablet is placed on the lower anvil, and the anvil is then adjusted so that the tablet just touches the upper test anvil. A suspended weight, motor driven, moves along a rail, which slowly and uniformly transmits pressure to the tablet. A pointer moving along a scale provides the breaking strength value in kilograms. A modified version of Erweka hardness tester operating in a horizontal position is shown in Fig. 13.12. The Schleuniger tester also operates in a horizontal position. An anvil driven by an electric motor presses the tablet at a constant load rate against a stationary anvil until the tablet breaks. A pointer moving along a scale indicator provides the breaking strength value. The instrument reads in both kilograms and Strong-Cobb units. This instrument is currently the most widely employed and has the advantage of being both fast and reproducible.

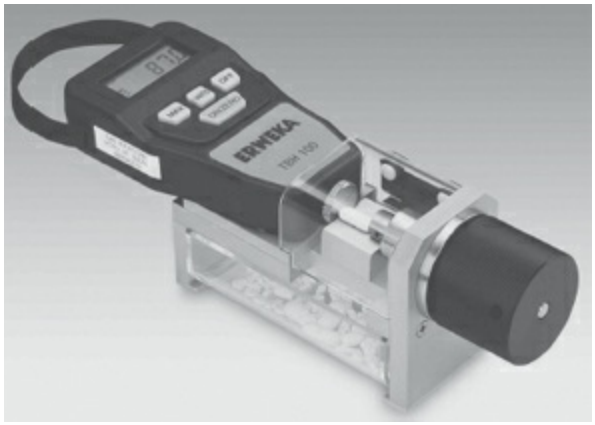


**Fig. 13.12:** The Erweka tablet hardness tester (*Courtesy* of ERWEKA)

Unfortunately, these testers do not produce the same results for the same tablet. Studies have shown that operator variation, lack of calibration, spring fatigue, and manufacturer variation contribute greatly to the lack of uniformity. Even those testers designed to eliminate operator variability have been found to vary.

Operators must be aware of these variations, especially when the tablets are to be evaluated by other persons or in other labs. For accurate comparison, each instrument should be carefully calibrated against a known standard.

The hardness of a tablet, like its thickness, is a function of the die fill and compression force. At a constant die fill, the hardness values increase and thickness decreases as additional compression force is applied. This relationship holds up to a maximum value for hardness and a minimum value for thickness, beyond which increases in pressure cause the tablet to laminate or cap, thus destroying the integrity of the tablet. At a constant compression force (fixed distance between upper and lower punches), hardness increases with increasing die fills and decreases with lower die fills.

When uniform tooling is used, the die-fill/force relationship makes control of tablet hardness a useful method of physically controlling tablet properties during a production operation, particularly when this measurement is combined with measurements of tablet thickness. The fill/force relationship is also the basis for instrumenting tablet machines.

In general, tablets are harder several hours after compression than they are immediately after compression. Lubricants can affect tablet hardness when they are used in too high a concentration or mixed for too long a period. Large tablets require a greater force to cause fracture and are therefore “harder” than small tablets. For a given granulation, a flat beveled tool produces a tablet harder than a deep cup tool.

## Friability

Tablet hardness is not an absolute indicator of strength since some formulations, when compressed into very hard tablets, tend to "cap" on attrition, losing their crown portions. Therefore, another measure of a tablet's strength, its friability, is often measured. Tablets that tend to powder, chip, and fragment when handled lack elegance and consumer acceptance, and can create excessively dirty processes in such areas of manufacturing as coating and packaging. They can also add to a tablet's weight variation or content uniformity problems.

The friability test is official in USP but not in BP and IP The laboratory friability tester is known as the Roche friabilator. This device, shown in Fig. 13.13, subjects a number of tablets to the combined effects of abrasion and shock by utilizing a transparent synthetic polymer chamber with an internal diameter between 283 and 291 mm and a depth between 36 and 40 mm that revolves at 25 ±1 rpm. The tablets are tumbled from a distance of six inches at each turn of the drum by a curved projection. Normally, a preweighed tablet sample is placed in the friabilator ($w$), which is then operated for 100 revolutions. For tablets with a unit weight equal to or less than 650 mg, take a sample of whole tablets corresponding as near as possible to 6.5 g. For tablets with a unit weight of more than 650 mg, take a sample of 10 whole tablets. After testing, the tablets are dusted and reweighed ($w_0$).


**Fig. 13.13:** The friability apparatus (*Courtesy* of ERWEKA)

The friability, *f*, is given by:

$$f = 100 \times \left(1 - \frac{w_o}{w}\right) \qquad \ldots (1)$$

Generally, the test is run once. If obviously cracked, cleaved, or broken tablets are present in the tablet sample after tumbling, the sample fails the test. If the results are difficult to interpret or if the weight loss is greater than the targeted value, the test should be repeated twice and the mean of the three tests determined. A maximum mean weight loss from the three samples of not more than 1.0% is generally considered acceptable for conventional compressed tablets.

Some chewable tablets and most effervescent tablets undergo high friability weight losses, which accounts for the special stack packaging that may be required for these types of tablets. When capping is observed on friability testing, the tablet should not be considered for commercial use, regardless of the percentage of loss seen.

When concave and especially deep concave punches are used in tabletting, and especially when the punches are in poor condition or worn at their surface edges, the tablets produced result in “whiskering” at the tablet edge. Such tablets have higher than normal friability values because the “whiskers” are removed in testing. Tablet friability may also be influenced by the moisture content of the tablet granulation and finished tablets. A low but acceptable moisture level frequently acts as a binder. Very dry granulations that contain only fractional percentages of moisture often produce more friable tablets than do granulations containing 2 to 4% moisture. For this reason, the manufacture of chemically stable tablets that contain some hydrolyzable drugs that are mechanically sound is difficult.

The traditional hardness and friability evaluations performed on tablets involve only a small sample of tablets. How the tablets withstand the mechanical shocks of a production environment is related to the large number of tablets involved, the production equipment used, and the skill of the production personnel? Rough handling tests can be performed to give an indication of how well a tablet will hold up in its specified package and shipping container during shipment. Rough handling tests usually include a vibration test, a drop test, and an incline plane impact test. Some investigators have actually shipped bottled products across the country and back again to estimate the strength of the new tablet product in shipment.

## Drug Content and Release

As mentioned earlier, a physically sound tablet may not produce the desired effects. To evaluate a tablet's potential for efficacy, the amount of drug per tablet needs to be monitored from tablet to tablet and batch to batch, and a measure of the tablet's ability to release the drug needs to be ascertained. The uniformity of dosage units can be demonstrated by either of two methods, Weight Variation or Content Uniformity (Table 13.8).

**Table 13.8:** Application of weight variation (WV) and content uniformity (CU) tests for dosage forms

| Dosage form | Type | Dose and ratio of drug | |
|---|---|---|---|
| | | >25 mg and >25 % | <25 mg and <25 % |
| Tablets | Uncoated/film coated | WV | CU |
| | Coated - others | CU | CU |
| Capsules | Hard | WV | CU |
| | Soft | CU | CU |
| Solids in single-unit containers | Single component | WV | WV |
| | Freeze dried product | WV | WV |
| Suspension, emulsion, or gel for systemic use, in single-unit containers | | CU | CU |
| Solutions for inhalation in glass or plastic ampoules and oral solutions in unit-dose containers | | WV | WV |
| Inhalations in premetered dosage units | | CU | CU |
| Transdermal systems | | CU | CU |
| Suppositories | | CU | CU |
| Others | | CU | CU |

## Weight Variation

With a tablet designed to contain a specific amount of drug in a specific amount of tablet formula, the weight of the tablet being made is routinely measured to help ensure that a tablet contains the proper amount of drug. The weight variation test is run by weighing 10 tablets individually, calculating the average weight, and comparing the individual tablet weights to the average. The requirements for dosage uniformity are met if the acceptance value of the first 10 dosage units is less than or equal to 15%. If the acceptance value is greater than 15%, test the next 20 units and calculate the acceptance value. The requirements are met if the final acceptance value of the 30 dosage units is less than or equal to 15%, and no individual content of any dosage unit is less than nor more than 25%.

The weight variation test would be a satisfactory method of determining the drug content uniformity of tablets if the tablets contain 25 mg or more of drug substance and that comprises 25% or more (by weight) of one tablet. For tablets such as aspirin, which are usually 90% or more active ingredient, the ±5% weight variation should come close to defining true potency and content uniformity (95 to 105% of the label strength) if the average tablet weight is close to the theoretic average weight. The weight variation test is clearly not sufficient to assure uniform potency of tablets of moderate or low-dose drugs, in which excipients make up the bulk of the tablet weight.

## Content Uniformity

The potency of tablets is expressed in terms of grams, milligrams, or micrograms (for some potent drugs) of drug per tablet and is given as the label strength of the product. Official compendia or other standards provide an acceptable potency range around the label potency. For highly potent, low-dose drugs such as digitoxin, this range is usually not less than 90% and not more than 110% of the labeled amount. For most other larger-dose drugs in tablet form, the official potency range that is permitted is not less than 95% and not more than 105% of the labeled amount.

In general, official potency analytic methods require that a composite sample of the tablets be taken, ground up, mixed, and analyzed to produce an average potency value. In composite assays, individual discrepancies can be masked by use of the blended sample. Even though the average assay result looks acceptable, it could mask a wide variation in potency, with the result that a patient could be variably underdosed or overdosed. With such a drug as digitoxin, in which the safe and effective level and the toxic level are close (or even overlapping), exceeding the official or accepted potency range is not only undesirable, but possibly dangerous.

Three factors can directly contribute to content uniformity problems in tablets (1) nonuniform distribution of the drug substance throughout the powder mixture or granulation, (2) segregation of the powder mixture or granulation during the various manufacturing processes and (3) tablet weight variation. As noted in the previous section, the use of weight cannot be used as a potency indicator, except perhaps when the tablets contain 25 mg or more of drug substance and that comprises 25% or more (by weight) of one tablet. In tablets with smaller dosages, a good weight variation does not ensure good content uniformity, but a large weight variation precludes good content uniformity.

To assure uniform potency for tablets of low-dose drugs, a content uniformity test is applied. In this test, 30 tablets are randomly selected for the sample, and at least 10 of them are assayed individually. Nine of the 10 tablets must contain not less than 85% or more than 115% of the labeled drug content. The tenth tablet may not contain less than 75% or more than 125% of the labeled content. If these conditions are not met, the tablets remaining from the 30 must be assayed individually, and none may fall outside of the 85

to 115% range. In evaluating a particular lot of tablets, several samples of tablets should be taken from various parts of the production run to satisfy statistical procedures.

The *purity* of official tablets is usually assured by utilizing raw materials, both active drug and all excipients, that meet official or other rigid specifications. Extraneous substances present in a raw material or a drug that is not specifically allowed by compendial specifications or well-defined manufacturer's specifications may render the product unacceptable for pharmaceutical use. These extraneous substances may be toxic on acute or long-term use or may have an unpredictable or deleterious effect on product stability or efficacy. Certain well-defined impurities often appear in the specification of raw materials or drug substances, or if they are the product of unavoidable decomposition of the drug, they may be listed with an upper tolerance limit. For example, aspirin tablets as specified by the USP may contain no more than 0.15% of free salicylic acid relative to the amount of aspirin present.

## Disintegration

A generally accepted maxim is that for a drug to be readily available to the body, it must be in solution. For most tablets, the first important step toward solution is breakdown of the tablet into smaller particles or granules, a process known as *disintegration*. The time that it takes a tablet to disintegrate is measured in a device described in the USP/NF.

Research has established that one should not automatically expect a correlation between disintegration and dissolution. However, since the dissolution of a drug from the fragmented tablet appears to control partially or completely the appearance of the drug in the blood, disintegration is still used as a guide to the formulator in the preparation of an optimum tablet formula and as an in-process control test to ensure lot-to-lot uniformity.

The USP device to test disintegration uses 6 glass tubes that are 3 inches long, open at the top, and held against a 10-mesh screen at the bottom end of the basket rack assembly (Fig. 13.14). To test for disintegration time, one tablet is placed in each tube, and the basket rack is positioned in a 1-L beaker of water, simulated gastric fluid, or simulated intestinal fluid, at 37°C ± 2°C, such that the tablets remain 2.5 cm below the surface of the liquid on their upward movement and descend not closer than 2.5 cm from the bottom of the beaker. A standard motor-driven device is used to move the basket assembly containing the tablets up and down through a distance of 5 to 6 cm at a frequency of 28 to 32 cycles per minute. Perforated plastic discs may also be used in the test. These are placed on top of the tablets and impart an abrasive action to the tablets. The discs may or may not be meaningful or impart more sensitivity to the test, but they are useful for tablets that float.

**Fig. 13.14:** Tablet disintegration tester

To be in compliance with the USP standards, the tablets must disintegrate, and all particles must pass through the 10-mesh screen in the time specified. If any residue remains, it must have a soft mass with no palpably firm core. Procedures are stated for running disintegration times for uncoated tablets, plain-coated tablets, enteric coated tablets, buccal tablets, and sublingual tablets. Uncoated USP tablets have disintegration time standards as low as 5 min (aspirin tablets), but the majority of the tablets have a maximum disintegration time of 30 min. Enteric coated tablets are to show no evidence of disintegration after 1 hour in simulated gastric fluid. The same tablets are then tested in simulated intestinal fluid and are to disintegrate in 2 hours plus the time specified in the monograph.

## Dissolution

The original rationale for using tablet disintegration tests was the fact that as the tablet breaks down into small particles, it offers a greater surface area to the dissolving media and therefore must be related to the availability of the drug to the body. The disintegration test, however, simply identifies the times required for the tablet to break up under the conditions of the test and for all particles to pass through a 10-mesh screen. The test offers no assurance that the resultant particles will release the drug in solution at an appropriate rate. For this reason, dissolution tests and test specifications have now been developed for nearly all tablet products. The rate of drug absorption for acidic drug moieties that are absorbed high in the GI tract is often determined by the rate of drug dissolution from the tablet. If the attainment of high peak blood levels for the drug is a product objective, obtaining rapid drug dissolution from the tablet is usually critically important. The rate of dissolution may thus be directly related to the efficacy of the tablet product, as well as to bioavailability differences between formulations. Therefore, an evaluation as to whether or not a tablet releases its drug contents when placed in the environment of the gastrointestinal tract is often of fundamental concern to the tablet formulator.

The most direct assessment of a drug's release from various tablet formulations or products is accomplished through in vivo bioavailability measurements. The use of in vivo studies is restricted, however, for several reasons: the length of time needed to plan, conduct, and interpret the study; the highly skilled personnel required for human studies; the low precision and high variability typical of the measurements; the high cost of the studies; the use of human subjects for "nonessential" research; and the necessary assumption that a perfect correlation exists between diseased patients and the healthy human subjects used in the test. Consequently, in vitro dissolution tests have been extensively studied, developed, and used as an indirect measurement of drug availability, especially in preliminary assessments of formulation factors and manufacturing methods that are likely to influence bioavailability. As with any in vitro test, it is critically important that the dissolution test be correlated with in vivo bioavailability tests.

Two objectives in the development of in vitro dissolution tests are to show (1) that the release of the drug from the tablet is as close as possible to

100% and (2) that the rate of drug release is uniform batch to batch and is the same as the release rate from those batches proven to be bio-available and clinically effective. Since 1970, the United States Pharmacopeia and the National Formulary have provided procedures for dissolution testing. They determine compliance with the limits on dissolution as specified in the individual monograph for a tablet (or a capsule). The USPXX/NFXV, Supplement 3, specifies that either of two apparatuses be used for determining dissolution rates.

## *Apparatus 1*

In general, a single tablet is placed in a small wire mesh basket fastened to the bottom of the shaft connected to a variable speed motor. The basket is immersed in the dissolution medium (as specified in the monograph) contained in a 100 ml flask. The flask is cylindric with a hemispherical bottom. The flask is maintained at 37°C ± 0.5°C by a constant temperature bath. The motor is adjusted to turn at the specified speed, and samples of the fluid are withdrawn at intervals to determine the amount of drug in solution.

## *Apparatus 2*

The same equipment as in apparatus 1 is used, except that the basket is replaced by a paddle, formed from a blade and a shaft, as the stirring element. The dosage form is allowed to sink to the bottom of the flask before stirring. Dosage forms may have a “small, loose piece of nonreactive material such as not more than a few turns of wire helix” attached to prevent them from floating. Description of a dissolution test in a USP/NF monograph specifies the dissolution test medium and volume, which apparatus is to be used, the speed (rpm) at which the test is to be performed, the time limit of the test, and the assay procedure. The test tolerance is expressed as a percentage of the labeled amount of drug dissolved in the time limit. For example, for methyldopa tablets, the dissolution test calls for a medium of 900 ml of 0.1N HCl, apparatus 2 turning at 50 rpm, and a time limit of 20 min. The accepted amount dissolved in 20 min is not less than 80% of the labeled amount of methyldopa (based on the cited assay procedure).

Dissolution testing and interpretation can be continued through three stages if necessary. In stage 1($S_1$), six tablets are tested and are acceptable if all of the tablets are not less than the monograph tolerance limit (Q) plus 5%.

If the tablets fail $S_1$, an additional six tablets are tested ($S_2$). The tablets are acceptable if the average of the twelve tablets is greater than or equal to Q and no unit is less than Q minus 15%. If the tablets still fail the test, an additional 12 tablets are tested. The tablets are acceptable if the average of all 24 tablets is greater than or equal to Q and if not more than 2 tablets are less than Q minus 15%.

Industrial pharmacists routinely test their formulations for dissolution. Their results are plotted as concentration versus time. Values for $t_{50\%}$, $t_{90\%}$ and the percentage dissolved in 30 min are used as guides. The value for $t_{50\%}$ is the length of time required for 50% of the drug to go into solution. A value for $t_{90\%}$ of 30 min is often considered satisfactory and is an excellent goal since a common dissolution tolerance in the USP/NF is not less than 75% dissolved in 45 min.

## PROCESSING PROBLEMS

In the normal process of developing formulations, and in the routine manufacture of tablets various problems occur. Sometimes, the source of the problem is the formulation, the compression equipment, or a combination of the two.

## Capping and Lamination

*Capping* is a term used to describe the partial or complete separation of the top or bottom *crowns* of a tablet from the main body of the tablet (Fig. 13.15). *Lamination* is the separation of a tablet into two or more distinct layers. Usually, these processing problems are readily apparent immediately after compression; however, capping and lamination may occur hours or even days later. Subjecting tablets to the friability test described earlier is the quickest way of revealing such problems. Capping and lamination have in the past been attributed to air entrapment. During the compression process, air is entrapped among the particles or granules and does not escape until the compression pressure is released. Research has shown, however, that capping and lamination are due to the deformational properties of the formulation during and immediately following compression.

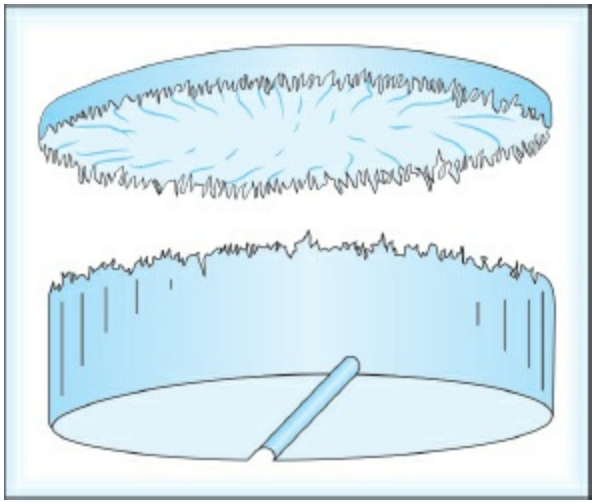


**Fig. 13.15:** Diagram of a “capped” tablet

During compaction, particles undergo sufficient plastic deformation to produce die-wall pressures greater than can be relieved by elastic recovery when the punch pressure is removed. In some materials, this die-wall pressure causes enough internal stress to cause a crack to propagate and initiate fracture of the compact in the die. If the excess stresses do not initiate fracture upon decompression in the die, the compact may laminate or cap upon ejection from the die. As the compact is ejected, the die-wall pressure falls to zero. The emerging portion of the compact expands while the confined portion cannot, thus concentrating shear stresses at the edge of the

die and causing a break to develop. Tablets that do not fracture have the ability to relieve the shear stresses developed during decompression and/or ejection by plastic deformation. This stress relaxation is time-dependent; therefore, the occurrence of tablet fracture is also time-dependent. Intact tablets of acetaminophen, methenamine, and erythromycin can be made if the decompression is extended for several hours. Rapid decompression results in tablets that fracture. Stress relaxation could be the explanation for some practical tabletting problems. Tablet lamination or capping problems are often eliminated by precompression, slowing the tabletting rate, and reducing the final compression pressure. As the stress relaxation time is increased, the amount of stress needing to be relieved is reduced, allowing an intact compact to be formed.

Often, deep concave punches produce tablets that cap. The curved part of such tablets expands radially while the body of the tablet cannot, which establishes a shear stress that produces the fracture. Flat punches may eliminate this additional shear stress.

A certain percentage of moisture is often essential for good compaction. A granulation that is too dry tends to cap or laminate for lack of cohesion. For moisture-critical granulations, the addition of a hygroscopic substance, e.g. sorbitol, methylcellulose, or polyethylene glycol 4000, can help to maintain a proper moisture level. Capping and lamination may also be encountered in direct compression product development. Some powder or fine particulate materials may not be compressible or may have poor compression properties. Relative compressibility of various materials may be reflected by their degree of consolidation (crown thickness) when compressed in standard tooling under identical compression conditions.

Tablet tooling can also be a cause of capping. The concave or beveled edge faces of punches gradually curve inward with use and form a “claw” that can pull off the crowns of a tablet. Wear in the upper punch guides accelerates this claw formation by permitting the punch tips to strike the edges of the die holes. Also, the greater the radius of curvature of the punch face, the greater is the force exerted on the edges and the less on the center of the tablet at the moment of compression.

Dies develop a wear “ring” in the area of compression. As the ring develops, and enlarges, the tablets that are compressed in the rings have a diameter that is too large to pass easily through the narrower portion of the

die above the ring. Upon ejection, this constriction causes the tablet to cap or laminate. A simple solution of this particular problem is to turn the die over so that compression occurs in an unworn area above the ring. On some presses, the depth of penetration of the upper punch can be regulated so that compression may be performed over some range of locations within the die. There are also dies available with tungsten carbide inserts. The carbide is so durable that the casing wears out before the insert does. Wear on tablet tooling increases as the hardness of the material being compressed increases. Most organic materials are soft; certain inorganic materials such as magnesium trisilicate are relatively hard and abrasive.

Another cause of capping is an incorrect setup at the press. When a compressed tablet is ejected from a die, the lower punch must rise flush with or protrude slightly above the face of the die at the point where the tablet strikes the sweep-off blade. If the punch remains below the face of the die, the sweep-off blade cuts off the tablet, leaving the bottom in the die. A less severe result of this incorrect adjustment is that the edge of the tablet catches on the die and chip. An incorrect adjustment of the sweep-off blade can also result in tablet fracture. If the blade is adjusted too high, tablets can start to travel under it, become stuck, and break off. The resulting broken pieces of tablets then enter the feed frame; if they are large enough, they can cause a disruption of the granulation feed, as well as affect the weight and hardness of subsequent tablets.

## Picking and Sticking

“Picking” is a term used to describe the surface material from a tablet that is sticking to and being removed from the tablet’s surface by a punch. Picking is of particular concern when punch tips have engraving or embossing. Small enclosed areas such as those found in the letters “B,” “A,” and “O” are difficult to manufacture cleanly. Tablet materials that stick to the punches can accumulate to the point of obliterating the tip design. “Sticking” refers to tablet material adhering to the die wall. When sticking occurs, additional force is required to overcome the friction between the tablet and the die wall during ejection. Serious sticking at ejection can cause chipping of a tablet’s edges and can produce a rough edge. Also, a sticking problem does not allow the lower punches free movement and therefore can place unusual stresses on the cam tracks and punch heads, resulting in their damage. Sticking can also apply to the buildup of material on punch faces.

These flaws have many remedies. Lettering should be designed as large as possible, particularly on punches with small diameters. The tablet can perhaps be reformulated to a larger size. Plating of the punch faces with chromium is a method for producing a smooth, nonadherent face.

In some cases, colloidal silica added to the formula acts as a polishing agent and makes the punch faces smooth so that material does not cling to them. On the other hand, the frictional nature of this material may require additional lubrication to facilitate release of the tablet from the die. Sometimes, additional binder or a change in binder may make the granules more cohesive, and therefore less adherent, than before.

Low-melting-point substances, either active ingredients or additives such as stearic acid and polyethylene glycol, may soften sufficiently from the heat of compression to cause sticking. Dilution of the active ingredient with additional higher-melting-point materials and a consequent increase in the size of the tablet may help. The level of low-melting-point lubricants may be reduced, or higher-melting-point replacements may be substituted. When a low-melting-point medicament is present in high concentration, refrigeration of the granulation and the press may be in order. Excessive moisture may be responsible for sticking, and further drying of the granulation is then required.

## Mottling

Mottling is an unequal distribution of color on a tablet, with light or dark areas standing out in an otherwise uniform surface. One cause of mottling is a drug whose color differs from the tablet excipients or a drug whose degradation products are colored. The use of colorants may solve the above problem but can create others. A dye can cause mottling by migrating to the surface of a granulation during drying. To overcome this difficulty, the formulator may change the solvent system, change the binder system, reduce the drying temperature, or grind to a smaller particle size. The use of colorants in direct compression formulations can lead to mottling if the dye is not well dispersed or if its particle size is too large.

Certain colored adhesive gel solutions may not be distributed well because they must be hot when added to much cooler powder mixtures. The adhesive then precipitates from solution and carries most of the color with it. Further wetting, even overwetting, is needed to disperse the binder and the color. The additional mixing and increased activation of the binder, however, may result in tablets with increased disintegration times. Therefore, a better practice may be to incorporate fine powder adhesives such as acacia and tragacanth into the product before adding the granulating fluid, or to disperse a dry color additive during the powder blending step.

## Weight Variation

In previous sections, weight variation of tablets has been mentioned as an important in-process control measurement, and weight variation specifications have been given. The weight of a tablet being compressed is determined by the amount of granulation in the die prior to compression. Therefore, anything that can alter the die-filling process can alter tablet weight and weight variation.

### *Granule Size and Size Distribution Before Compression*

Variations in the ratio of small to large granules and in the magnitude of difference between granule sizes influence how the void spaces between particles are filled. Thus, although the apparent volume in the die is essentially the same, different proportions of large and small particles may change the weight of fill in each die. Furthermore, if large granules are being used to fill a small die cavity, relatively few granules are required, and the difference of only a few granules around the average may represent a high percentage weight variation. If hundreds of granules are required on the average for die fill, a variation of a few granules around the average would produce a minor weight variation, giving a narrow particle size range.

### *Poor Flow*

The die-fill process is based on a continuous and uniform flow of granulation from the hopper through the feed frame. When the granulation does not flow readily, it tends to move spasmodically through the feed frame so that some dies are incompletely filled. Similarly, dies are not filled properly when machine speed is in excess of the granulation’s flow capabilities. With poor flow, the addition of a glidant such as talcum or colloidal silica, or an increase in the amount already present, may be helpful.

Also available are induced die feeders, which mechanically “force” the granulation down into the die cavities as they pass beneath the feed frame.

Poor flow through the feed frame is usually a sign that the granulation is not flowing properly out of the hopper. As particulate solids move under the force of gravity through progressively smaller openings, they are subjected to uneven pressures from the mass above and alongside. Depending on the geometry of the hopper, this situation may give rise to one or another of two

causes for poor flow: “arching” or “bridging,” and “rat-holing.” Figure 13.16 illustrates these phenomena. When poor hopper flow occurs, it may be controllable with vibrators attached to the hopper sides to induce the granulation flow.

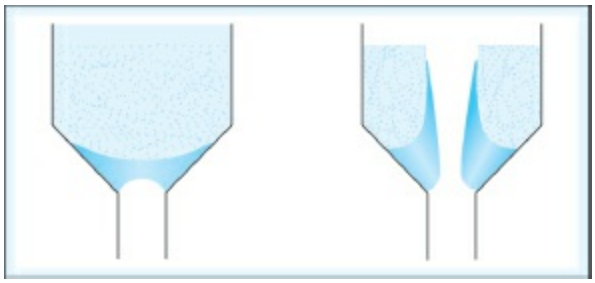


**Fig. 13.16:** Bridging (left); rat-holing (right)

Devices designed to improve poor flow characteristics of materials often introduce another problem, however. Since most tablet granulations consist of materials with a range of particle sizes, the vibration or mixing action of the flow-promoting devices may induce segregation and stratification of the particles. The larger particles tend to drift upward while the smaller particles sift downward. Not only can the resulting “classification” of particle sizes cause appreciable changes in tablet weight and weight variation as described earlier, but it can also lead to poor content uniformity, since drug is often not uniformly distributed between the larger and smaller particles. Poor particulate flow may be caused not by the granulation, but by poor design of the granulation hopper, which can be exaggerated by dents that effectively cut off the flow. Poor weight variation can also be caused by surges of excessive flow. Direct compression granulations fed through typical wet-granulation hoppers and feed frames are prone to this type of flow. Often, restricting the flow out of the hopper corrects the problem. Recently, a patent was issued for a new feed frame design that accommodated excessive flow from the hopper without compromising uniform weight variation.

## *Poor Mixing*

Sometimes, the lubricants and glidants are not thoroughly distributed. The flow of particles is then impaired, and the granules do not move efficiently into the dies. There is a tendency to minimize the mixing time during lubricant addition to prevent or reduce granule friability; however, inadequate mixing at this stage can result in unsatisfactory granulation flow.

## *Punch Variation*

When lower punches are of unequal lengths—the difference may be only a few thousandths of an inch—the fill in each die varies because the fill is volumetric. Only a good punch and die control program can provide tooling of uniform dimensions.

## Hardness Variation

Hardness variation is a problem that has the same causes as weight variation. Hardness depends on the weight of material and the space between the upper and lower punches at the moment of compression. If the volume of material or the distance between punches varies, hardness is likewise inconsistent.

## Double Impression

A last problem for discussion is that of double impression. This involves only punches that have a monogram or other engraving on them. At the moment of compression, the tablet receives the imprint of the punch. On some machines, the lower punch is free to drop and then travel uncontrolled for a short distance before it rides up the ejection cam to push the tablet out of the die. During its free travel, it rotates. At this point, the punch may make a new, although lighter, impression on the bottom of the tablet, resulting in a double imprint. Similar problems can be encountered with engraved upper punches and tablet machines that utilize two compression stages to compress a tablet. The first stage, *precompression*, uses a lower compaction force, than the final compression stage, but the tablet does receive the imprint of the punch. If the upper punch is uncontrolled, it can rotate during the short travel to the final compression stage and thus create a double imprint. The newer presses have antiturning devices as an integral part of their design and construction.

## MANUFACTURING IMPROVEMENTS

### Basic Improvement Areas

Wet granulation has traditionally been a highly labor-intensive and time-consuming process (Table 13.9). In the last 20 years, however, significant improvements in tablet manufacturing efficiency have taken place. These can be attributed to four basic areas: the elimination or combination of processing steps, the improvement of specific unit operations, the design of new equipment specifically oriented to granulation objectives, and the improvement of materials handling techniques and systems. Illustrating the use of these improvement areas, Table 13.9 compares the processing steps of Lederle's old tablet-manufacturing process to its new tabletmanufacturing process.

**Table 13.9:** Unit processing of solid dosage forms (tablet manufacturing)—Lederle laboratories

| Old process (wet granulation) | New process (direct compression) |
|---|---|
| Raw materials | Raw materials |
| Weighing and measuring | Weighing and measuring (automatic weigher and recording system) |
| Screening | Gravity feeding |
| Manual feeding | Blending |
| Blending (slow-speed planetary mixer) | Gravity feeding from the storage tank |
| Wetting (hand addition) | Compression (high-speed rotary press) |
| Subdivision (comminutor) | Aqueous coating (Hi-Coater) |
| Drying (fluid bed dryer) | |
| Subdivision (comminutor) | |
| Premixing (barrel roller) Batching and lubrication (ribbon blender) | |
| Manual feeding | |

Solvent film coating (Wurster column)

Tablet inspection (manual)

## *Elimination or Combination of Steps*

Tables 13.9 indicate the processing steps that may be omitted on conversion from wet granulation to direct compression. As noted, new mixer/granulators allow several processes of wet granulation to be conducted in rapid succession or to be combined in one piece of equipment.

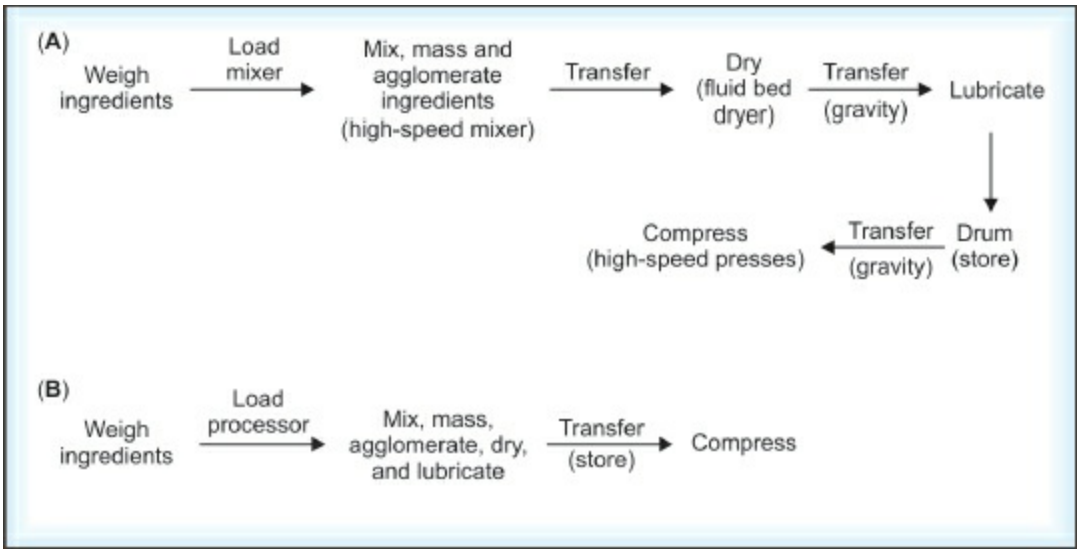
## *Unit Operation Improvement*

The efficiency of new tablet manufacturing methods, as exemplified by the new process, was achieved by enhancing the efficiency of three specific unit operations. First, material blending was improved by replacing slow-speed planetary type mixers with high-speed mixer/granulators. Second, the tablet compression operation was improved by replacing old single-fill-station, gravity-fed compression machines with newer high-speed, multistation presses, with induced die feed and automated weight control. Third, the coating operation was brought into better compliance with EPA standards by switching from organic-solvent based systems to aqueous systems, which were further aided by side-vented coating equipment having greatly improved drying efficiencies.

## *Materials Handling*

A major labor-saving change made in equipment designs allowed material to be moved by gravity. A granulation gravity feed system was designed that eliminated the manual feeding of granulation into the presses.

With use of similar techniques, even wet granulation operations have been made more efficient, and a hypothetical state-of-the-art processing scheme is presented in the flow chart in Fig. 13.17. All processing steps including drying might be combined in a single processor in the wet granulation method of the future. Such equipment will minimize materials handling, labor requirement, and human variables.




**Figs 13.17A and B:** Flow charts depict state-of-the-art wet granulation processing of the 1980s (A), projected wet granulation processing methods of the future and beyond (B)

## *Equipment*

Various equipment improvements that would combine several of the wet granulation processing steps are being investigated. One such method is the use of the sprayer dryer. The components of the formula-diluent, binder, disintegrant, and lubricant may be suspended and/or dissolved in a suitable vehicle according to their nature. The solids should represent at least 50 to 60% of the suspension. Under constant stirring to maintain good distribution, the slurry is pumped to an atomizing wheel, which whirls the material into a stream of hot air. The heat removes the liquid carrier, and the solids fall to the bottom of the dryer as fine, spherical granules ranging from as low as 10 to as high as 250 microns in diameter, the size depending on the speed of the wheel and the flow rate of the feed. The drug may be mixed with this “base” in proportions as high as 1:1. If the drug remains stable with the temperatures and solvents used, it may also be included in the slurry.

## *Computer Process Control*

As the tablet manufacturing processes continue to be improved in the four areas indicated, human worker involvement will continue to decline. As human involvement requirements are reduced, computer control of the process is inevitable. There are many good reasons for implementing

computer process control.

Rigid control enforcement

Operational information

Documentation of the process

Security of the process and its control

Increased consistency

Increased flexibility

Increased reliability

Increased productivity

Merck, Sharp and Dohme’s computer-controlled Aldomet plant has shown that tablet production under computer control with limited human intervention is possible in a continuous mode.

In the common batch mode of tablet production, individual unit operations such as coating processes, fluid bed dryers and tablet press monitors are becoming automated by microcomputers. There are obstacles to computer control, including the need for smaller and more powerful computer devices, better process interfacing sensors, particularly of the “composition” type, and better man/machine interfacing. Therefore, as the price and size of computers continue to decrease, as the availability of sensors increases, and as our knowledge of tablet processes increases, the computer control of tablet, batch operations will rapidly grow in the future.

# APPENDICES

## APPENDIX A

## Common Tablet Ingredients in Wet Granulation Formulas

### *Phenobarbital Tablets*

| Ingredient | Quantity per tablet | Quantity per 10,000 tablets |
|---|---|---|
| Phenobarbital | 65 mg | 650 g |
| Lactose (fine powder) | 40 mg | 400 g |
| Starch (paste) | 4 mg | 40 g |
| Starch (dry) | 10 mg | 100 g |
| Talc | 10 mg | 100 g |
| Mineral oil, 50 cps | 4 mg | 40 g |

### *Procedure*

Mix the phenobarbital and lactose, and moisten with 10% starch paste to proper wetness. Granulate by passing through a 14-mesh screen, and dry at 140°F. When dry, pass through a 20-mesh screen; add the dry starch and talc; mix well. Finally, add the mineral oil, mix again, and compress using ⅝₂-in. standard cup punches.

### *Aminophylline Tablets*

| Ingredient | Quantity per tablet | Quantity per 10,000 tablets |
|---|---|---|
| Aminophylline | 100 mg | 1.0 kg |
| Tricalcium phosphate | 50 mg | 0.5 kg |
| Pregelatinized starch | 15 mg | 0.15 kg |
| Water | q.s. | q.s. |
| Talc | 30 mg | 0.3 kg |

| | | |
|---|---|---|
| Mineral oil, light | 2 mg | 0.02 kg |

## *Procedure*

Mix the aminophylline, tricalcium phosphate, and starch; moisten with water. Pass through a 12-mesh screen, and dry at 100°F. Size the dry granules through a 20-mesh screen; add the talc and mix. Add the mineral oil, mix for 10 min, and compress using 5/16-in. deep cup punches for enteric coating.

## *Chewable Antacid Tablets*

| Ingredient | Quantity per tablet | Quantity per 10,000 tablets |
|---|---|---|
| Aluminum hydroxide (dried gel) | 400 mg | 4.0 kg |
| Magnesium hydroxide (fine powder) | 80 mg | 0.8 kg |
| Sucrose | 20 mg | 0.2 kg |
| Mannitol (fine powder) | 180 mg | 1.8 kg |
| PVP (10% solution) | 30 mg | 0.3 kg |

## *Procedure*

Mix the first four ingredients, and moisten with a 10% PVP solution in 50% ethanol. Granulate by passing through a 14-mesh screen. Dry at 140 to 150°F. Size through a 20-mesh screen, add the oil of peppermint mixed with the Cab-O-Sil and finally the magnesium stearate; mix well and compress using ½-in. flat-face beveled-edge punches.

## *Chewable Laxative Tablets*

| Ingredient | Quantity per tablet | Quantity per 10,000 tablets |
|---|---|---|
| Phenolphthalein | 64 mg | 0.64 kg |
| Powdered sugar | 750 mg | 7.5 kg |

| | | |
|---|---|---|
| Powdered cocoa (defatted) | 350 mg | 3.5 kg |
| Gelatin (10% solution) | q.s. | q.s. |
| Calcium stearate | 12 mg | 0.12 kg |
| Talc | 60 mg | 0.60 kg |

## *Procedure*

Mix the phenolphthalein, sugar, and cocoa, and moisten with the gelatin solution. Pass through an 8-mesh screen, and dry in a tray oven at 120 to 130°F. When dry, reduce granule size by passing through a 16-mesh screen. Mix the calcium stearate and talc, pass through a 100-mesh screen, add to the granulation, and compress to weight using ⅝-in. flat-face punches.

## *Ferrous Sulfate Tablets*

| Ingredient | Quantity per tablet | Quantity per 10,000 tablets |
|---|---|---|
| Ferrous sulfate (dried) | 300 mg | 3.0 kg |
| Corn starch | 60 mg | 0.60 kg |
| 20% Sugar solution | q.s. | q.s. |
| Explotab | 45 mg | 0.45 kg |
| Talc | 30 mg | 0.30 kg |
| Magnesium stearate | 4 mg | 0.04 kg |

## *Procedure*

Mix the ferrous sulfate and cornstarch; moisten with sugar syrup to granulate through a 12-mesh screen. Dry in a tray oven overnight at 140 to 150°F. Size through an 18-mesh screen; add the Explotab, talc, and magnesium stearate, and compress to weight using ⅜-in. deep cup punches in preparation for sugar-coating.

# APPENDIX B

## *Common Tablet Ingredients in Dry Granulation Formulas*

### *Aspirin Tablets (5-Grain)*

| Ingredient | Quantity per tablet | Quantity per 10,000 tablets |
|---|---|---|
| Aspirin (20-mesh) | 325.0 mg | 3.250 kg |
| Starch USP (dried | 32.5 mg | 0.325 kg |
| Cab-O-Sil | 0.1 mg | 0.010 kg |

### *Procedure*

Combine the aspirin, starch, and Cab-O-Sil, and mix in a P-K twin-shell blender for 10 min. Compress into slugs using 1-in. flat-face punches. Reduce the slugs to granulation by passing through a 16-mesh screen in a Stokes Oscillating Granulator or through a Fitzpatrick Mill with a #2B screen, at medium speed, and with knives forward. Transfer the granulation to a tablet machine hopper, and compress into tablets using ⅝-in. standard concave punches.

*Note:* All operations should be carried out in a dehumidified area at a relative humidity of less than 30% at 70°F.

### *Effervescent Aspirin Tablets (5-Grain)*

| Ingredient | Quantity per tablet | Quantity per 10,000 tablets |
|---|---|---|
| Sodium bicarbonate (fine granular) | 2.050 g | 20.500 kg |
| Citric acid (fine granular) | 0.520 g | 5.200 kg |
| Fumaric acid (fine granular) | 0.305 g | 3.050 kg |
| Aspirin (20-mesh, granular) | 0.325 g | 3.250 kg |

## *Procedure*

Mix the above ingredients in a P-K twin-shell blender for 20 min; transfer to a tablet machine equipped with 1¼-in. flat-face punches, and compress slugs to approximately ¾-in. thick. Grind the slugs through a 16-mesh screen. Mix for 5 min in a twin-shell blender, and compress into tablets using ⅞-in. flat-face beveled-edge punches.

*Note:* All operations should be carried out in a dehumidified area at a relative humidity of less than 30% at 70°F.

# APPENDIX C

## Common Tablet Ingredients in Direct Compression Formulas

### *Acetaminophen Tablets (USP)*

| Ingredient | Quantity per tablet | Quantity per 10,000 tablets |
|---|---|---|
| Acetaminohpen USP(granular or large crystal)* | 325.00 mg | 3.25 kg |
| Avicel PH 101 | 138.35 mg | 1.3835 kg |
| Stearic acid (fine powder) | 1.65 mg | 0.0165 kg |

* If smaller crystalline size acetaminophen is desired to improve dissolution, it is necessary to use a higher proportion of Avicel, to use pH 102 in place of pH 101, and to use a glidant. All lubricants should be screened before being added to blender

### *Procedure*

Blend the acetaminophen and Avicel pH 101 for 25 min. Screen in the stearic acid, and blend for an additional 5 min. Compress tablets using 7/16-in. standard concave or flat beveled tooling.

*Vitamin $B_1$ tablets*
*(thiamine hydrochloride USP; 100 mg)*

| Ingredient | Quantity per tablet | Quantity per 10,000 tablets |
|---|---|---|
| Thiamine hydrochloride USP | 100.00 mg | 1.0 kg |
| Avicel PH 102 | 83.35 mg | 0.8335 kg |
| Lactose (anhydrous) | 141.65 mg | 1.4165 kg |
| Magnesium stearate | 6.65 mg | 0.0665 kg |
| Cab-O-Sil | 1.65 mg | 0.0165 kg |

## *Procedure*

Blend all ingredients except the magnesium stearate for 25 min. Screen in the magnesium stearate and blend for an additional 5 min. Compress using 13/32-in. standard concave tooling.

*Note:* Anhydrous lactose could be replaced with Fast Flo lactose with no loss in tablet quality. This would reduce the need for a glidant (which is probably present in too high a concentration in most of these formulations). Usually, only 0.25% is necessary to optimize fluidity.

*Chlorpromazine tablets USP (100 mg)*

| Ingredient | Quantity per tablet | Quantity per 10,000 tablets |
|---|---|---|
| Chlorpromazine hydrochloride USP | 100.00 mg | 1.0 kg |
| Avicel PH 102 | 125.00 mg | 1.25 kg |
| Dicalcium phosphate (unmilled) or Emcompress | 125.00 mg | 1.25 kg |
| Cab-O-Sil | 1.74 mg | 0.0174 kg |
| Magnesium stearate | 5.25 mg | 0.0525 kg |

## *Procedure*

Blend all ingredients except the magnesium stearate for 25 min. Screen in the magnesium stearate, and blend for an additional 5 min. Compress into tablets using 11/32-in. tooling.

# PART II: TABLET COATING

No discussion on tablet coating would be complete without a brief historical review of pharmaceutical coating to provide an appropriate perspective to the evolutions in the coating process that have occurred over the past thousand years.

One of the earliest references to coated solid dosage forms appears in early Islamic drug literature, where coated pills were mentioned by Rhazes. The use of coating on drugs was probably an adaptation from early food preservation methods, and French publications in the 1600s described coating as a means of masking the taste of medicines. Sugar coating of pills was developed to a considerable extent by the French in the mid-1800s, and patents issued in 1837 and 1840 utilized sugar compositions for coated pills of cubeb and copaiba. Subsequently, there was rapid acceptance of sugar-coated pills as the preferred solid dosage form for both prescription and patent medicines in Europe and the United States. It soon was recognized that quality sugar coating on a large scale could be accomplished more readily in coating pans, and several early pharmaceutical companies in the United States were established, with coated pills as a major part of their product line.

Except for the substitution of compressed tablets for pills, the sugar coating equipment and process remained essentially unchanged for the next 75 years. In 1953, a dramatic change was made in tablet coating when Abbott Laboratories marketed the first film-coated pharmaceutical tablet. Concurrently, in the early 1950s, Dr. Dale Wurster, a professor at the University of Wisconsin, patented an air suspension coater that efficiently applied film coating compositions. This stimulated renewed interest in tablet coating technology, and for the next 12 to 15 years, several hundred patents and research papers on the subject were published. The invention by Dr. Wurster showed the merits of high airflow in the coating process-, and eventually, a series of perforated coating pans (Accela-Cota, Hi-Coater, Driacoater) were developed as replacements for the coating pans of the 30s and 40s (Figs 13.18 to 13.20).

In this chapter, the current state of tablet coating and the opportunities for continued improvement are presented.

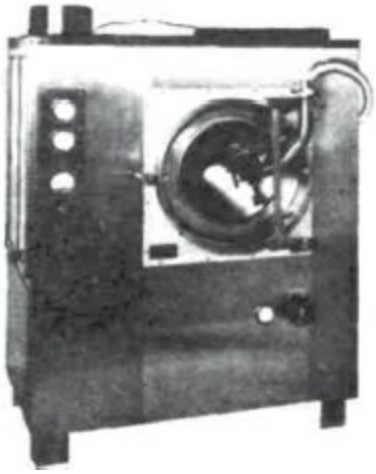


**Fig. 13.18:** Accela-cota system

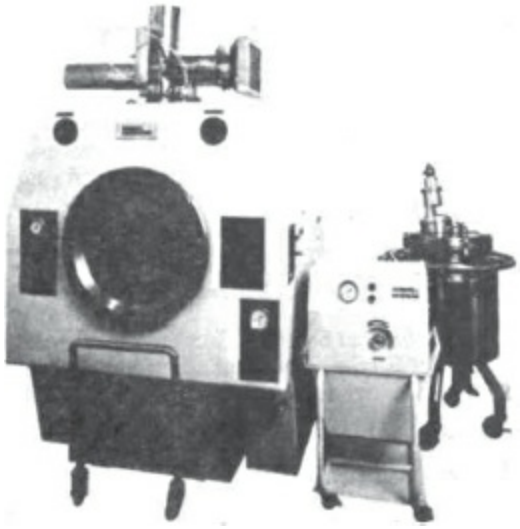


**Fig. 13.19:** Hi-coater system

**Fig. 13.20:** Driacoater system

## TABLET COATING PRINCIPLES

The principles of tablet coating are relatively simple. The application of coating to tablets, which is an additional step in the manufacturing process, increases the cost of the product; therefore, the decision to coat a tablet is usually based on one or more of the following objectives:

1. To mask the taste, odor, or color of the drug.
2. To provide physical and chemical protection for the drug.
3. To control the release of the drug from the tablet.
4. To protect the drug from the gastric environment of the stomach with an acid-resistant enteric coating.
5. To incorporate another drug or formula adjuvant in the coating to avoid chemical incompatibilities or to provide sequential drug release.
6. To improve the pharmaceutical elegance by use of special colors and contrasting printing.

The coating process can best be described by initially discussing the key factors that it comprises and then showing their complex interactions. There are three primary components involved in tablet coating:

1. Tablet properties.
2. Coating compositions.
3. Coating process.
    - Coating equipment.
    - Parameters of the coating process.
    - Facility and ancillary equipment.
    - Automation in coating processes.

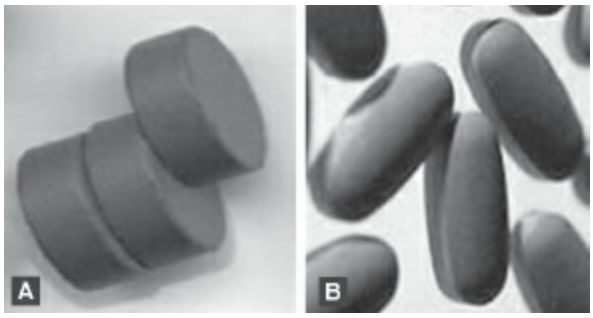
## TABLET PROPERTIES

Tablets that are to be coated must possess the proper physical characteristics. In the coating process, the tablets roll in a coating pan or cascade in the air stream of an air suspension coater as the coating composition is applied. To tolerate the intense attrition of tablets striking other tablets or walls of the coating equipment, the tablets must be resistant to abrasion and chipping. Tablet surfaces that are brittle, that soften in the presence of heat, or that are affected by the coating composition tend to become rough in the early phase of the coating process and are unacceptable for film coating. Film coatings adhere to all exposed surfaces, so that any surface imperfection is coated and not eliminated. The quality of thin film coatings applied to compressed tablets usually depends much more on the quality of the starting tablet than on the time at which sugar coatings are applied. Sugar coatings, with their high solids content, dry more slowly and can fill many of the minor tablet surface imperfections that may occur in the early phase of the coating process.

In addition to a smooth surface, the physical shape of the tablet is important. When a coating composition is applied to a batch of tablets in a coating pan, the tablet surfaces become covered with a tacky polymeric film. Before the tablet surface dries, the applied coating changes from a sticky liquid to a tacky semisolid, and eventually to a nontacky dry surface. The tablets must be in constant motion during the early drying phase or tablet agglomeration can occur. The ideal tablet shape for coating is a sphere, which allows tablets to roll freely in the coating pan, with minimal tablet-to-tablet contact. The worst shape is a square flat-faced tablet, in which case coating materials would collect between the surfaces to glue them together, like a stack of dominos or poker chips (Fig. 13.21A). For this reason, coated tablets have rounded surfaces (Fig 13.21B); the more convex the surface is, the fewer difficulties will be encountered with tablet agglomeration.

A compressed tablet formulation includes many ingredients besides the active drug to provide a readily compressible, resilient, and rapidly dissolving dosage form. The resulting surface properties of the tablet depend on the chemical nature of the ingredients utilized in the formulation. For the coating to adhere to the tablet, the coating composition must wet the surface. Hydrophobic tablet surfaces are difficult to coat with aqueous-based coatings

that do not wet the surface. The composition of the coating formulation can be adjusted, however, through the addition of appropriate surfactants to reduce the surface tension of the coating composition and improve coating adhesion.




**Figs 13.21A and B:** (A) Flat-faced tablets; (B) Rounded surface tablets

## COATING COMPOSITIONS

The coating materials may be a physical deposition of the material on the tablet substrate, or they may form a continuous film with a wide variety of properties depending upon the composition of the coating formulations. Examples of physical deposition of the coating materials are the techniques of sugar, shellac, and wax coatings. During the last 40 years, a wide variety of polymers have been evaluated and are being used commercially for tablet coating. Further discussion of coating materials in this chapter is limited to synthetic polymers, solvents, plasticizers, colorants, opaquant-extenders, and miscellaneous coating solution components.

An ideal film coating material should have the following attributes:

1. Solubility in solvent of choice for coating preparation.
2. Solubility required for the intended use, e.g. free water-solubility, slow water-solubility, or pH-dependent solubility (enteric coating)
3. Capacity to produce an elegant looking product.
4. Stability in the presence of heat, light, moisture, air, and the substrate being coated. The film properties should not change with aging.
5. Essentially no color, taste or odor.
6. Compatibility with common coating solution additives.
7. Nontoxicity with no pharmacologic activity, and ease of application to the particles or tablets.
8. Resistance to cracking, and provision of adequate moisture, light, odor, or drug sublimation barrier when desired.
9. No-bridging or filling of the debossed tablet surfaces by the film former.
10. Ease of printing procedure on high speed equipment.

No commercially available material fulfills all requirements of an ideal coating material. A pharmaceutical scientist usually formulates a coating solution to achieve certain desired properties for the film-coated product. The available film formers can be classified into nonenteric and enteric materials.

## Film Formers

### *Nonenteric Materials*

It is not possible to mention all polymers that have been investigated for filmcoating. The following discussion describes only some of the materials most commonly used by the pharmaceutical industry and is intended as a guide for the student or pharmaceutical scientist.

### *Hydroxypropyl Methylcellulose*

The polymer is prepared by reacting alkali-treated cellulose first with methyl chloride to introduce methoxy groups and then with propylene oxide to introduce propylene glycol ether groups. The resulting products are commercially available in different viscosity grades. This polymer is a material of choice for air suspension and pan-spray coating systems. The reasons for its widespread acceptance include (1) solubility characteristics of the polymer in gastrointestinal fluid, and in organic and aqueous solvent systems, (2) noninterference with tablet disintegration and drug availability, (3) flexibility, chip resistance, and absence of taste or odor, (4) stability in the presence of heat, light, air, or reasonable levels of moisture and (5) ability to incorporate color and other additives into the film without difficulty. The interaction of this polymer with colorants is rare. Hydroxypropyl methylcellulose closely approaches the desired attributes of an ideal polymer for film coating. When used alone, the polymer has the tendency to bridge or fill the debossed tablet surfaces. A mixture of hydroxypropyl methylcellulose with other polymers or plasticizers is used to eliminate bridging or filling problems. This polymer is also used considerably in glossing solutions.

### *Methyl Hydroxyethylcellulose*

This polymer is prepared by reacting alkali-treated cellulose first with methyl chloride and then with ethylene oxide. A wide variety of viscosity grades are available. Because of its structural similarity to hydroxypropyl methylcellulose, this polymer is expected to have similar properties. It is marketed in Europe, but because it is soluble in fewer organic solvents, it is not used as frequently as hydroxypropyl methylcellulose.

### *Ethylcellulose*

It is manufactured by the reaction of ethyl chloride or ethyl sulfate with cellulose dissolved in sodium hydroxide. Depending on the degree of ethoxy substitution, different viscosity grades are obtained and available commercially. This material is completely insoluble in water and gastrointestinal fluids, and thus cannot be used alone for tablet coating. It is usually combined with water-soluble additives, e.g. hydroxypropyl methylcellulose, to prepare films with reduced water solubility properties. A combination of ethylcellulose with water-soluble additives has been widely used in preparing sustained-release coatings of fine particles and tablets. The polymer is soluble in a wide variety of organic solvents and is nontoxic, colorless, odorless, tasteless, and quite stable to most environmental conditions. Unplasticized ethylcellulose films are brittle and require film modifiers to obtain an acceptable film formulation. Banker and co-workers from Purdue University have developed aqueous polymeric dispersions utilizing ethylcellulose. These pseudolatex systems are high-solids, low-viscosity compositions that have coating properties quite different from the regular ethylcellulose solutions. The material is commercially available through FMC Corporation as Aquacoat.

## *Hydroxyproylcellulose*

This material is manufactured by treatment of cellulose with sodium hydroxide, followed by a reaction with propylene oxide at an elevated temperature and pressure. It is soluble in water below 40°C (insoluble above 45°C), gastrointestinal fluids and many polar organic solvents. This polymer is extremely tacky as it dries from a solution system and may be desirable for a subcoat, but not for a color or gloss coat. The polymer yields very flexible films. It is usually not used alone, but it is used in combination with other polymers to improve the film characteristics.

## *Povidone*

Povidone is a synthetic polymer consisting of linear l-vinyl-2-pyrrolidinone groups. The degree of polymerization results in materials of various molecular weight range. Povidone is usually available in four viscosity grades identified by their K values, which approximate K-15, K-30, K-60, and K-90. The average molecular weight of these grades are 10,000, 40,000, 160,000, and 360,000 respectively. The most common uses of povidone in pharmaceuticals (frequently K-30) are as a tablet binder and a tablet coating.

It has excellent solubility in a wide variety of organic solvents, in water, and in gastric and intestinal fluids. When dry, povidone films are clear, glossy, and hard. The material is extremely tacky, but it is possible to modify the polymer properties by use of appropriate plasticizers, suspended powders, or other polymers. Although povidone is soluble in both acidic and basic fluids, it can be cross-linked with other materials to produce films with enteric properties. Povidone has been used to improve the dispersion of colorants in coating solutions to obtain a more uniformly colored film.

## *Sodium Carboxymethylcellulose*

This material is sodium salt of carboxymethyl-cellulose and is manufactured by the reaction of soda cellulose with the sodium salt of monochloroacetic acid. It is available in low, medium, high, and extra high viscosity grades. Sodium carboxymethylcellulose is easily dispersed in water to form colloidal solutions, but it is insoluble in most organic solvents, and therefore is not a material of choice for coating solutions based on organic solvents. Films prepared with sodium carbo-xymethylcellulose are brittle, but adhere well to tablets. Partially dried films are tacky, however, so coating compositions must be modified with additives. Conversion to aqueous-based film coating with high coating efficiency equipment probably increases the usefulness of this polymer in coating systems.

## *Polyethylene Glycols*

Polyethylene glycols (PEG) are manufactured by the reaction of ethylene glycol with ethylene oxide in the presence of sodium hydroxide at elevated temperature and under pressure. In addition to their other uses in formulations, they are used in film coating for which a wide variety of molecular weights are available. The materials with low molecular weights (200 to 600 series) are liquid at room temperature and are used as plasticizers for coating solution films. The materials with high molecular weights (series 900 to 8,000) are white, waxy solids at room temperature. These polymers are used in combination with other polymers to modify film properties. Combinations of polyethylene glycol waxes with cellulose acetate phthalate provide films that are soluble in gastric fluids. Such systems constituted one of the first commercially used nonenteric film coating processes. Coats produced with the use of high-molecular-weight PEGs can be hard, smooth, tasteless, and nontoxic, but are somewhat sensitive to elevated temperatures.

## *Acrylate Polymers*

A series of acrylate polymers is marketed under the trademark Eudragit. Eudragit E is a cationic copolymer based on dimethyl-aminoethyl methacrylate and other neutral methacrylic acid esters, and is the only Eudragit material that is freely soluble in gastric fluid up to pH 5, and expandable and permeable above pH 5. This material is available as (1) organic solution (12.5%) in isopropanol/acetone, (2) solid material, or (3) 30% aqueous dispersion. Eudragit RL and RS are copolymers synthesized from acrylic and methacrylic acid esters with a low content of quaternary ammonium groups. These are available only as organic solutions and solid materials. These polymers produce films for the delayed-action (pH-independent) preparations similar to ethylcellulose formulations. These materials are widely used in Europe, but have limited use so far in the United States.

## *Enteric Materials*

Enteric coating of pills and compressed tablets has existed for more than a century.[36] Some of the most important reasons for enteric coating are as follows:

1. To protect acid-labile drugs from the gastric fluid, e.g. enzymes and certain antibiotics.
2. To prevent gastric distress or nausea due to irritation from a drug, e.g. sodium salicylate.
3. To deliver drugs intended for local action in the intestines, e.g. intestinal antiseptics could be delivered to their site of action in a concentrated form and bypass systemic absorption in the stomach.
4. To deliver drugs that are optimally absorbed in the small intestine to their primary absorption site in their most concentrated form.
5. To provide a delayed-release component for repeat-action tablets.

An ideal enteric coating material should have the following properties:

1. Resistance to gastric fluids.
2. Ready susceptibility to or permeability to intestinal fluids.
3. Compatibility with most coating solution components and the drug substrates.

4. Stability alone and in coating solutions. The films should not change on aging.
5. Formation of a continuous (uninterrupted) film.
6. Non toxicity.
7. Low cost.
8. Ease of application without specialized equipment.
9. Ability to be readily printed or to allow film to be applied to debossed tablets.

Pharmaceutical formulators have a wide choice of materials for use in developing an enteric coated granule, pellet, or tablet product. These materials range from water-resistant films to pH-sensitive materials. Some are digested or emulsified by intestinal juices, and some slowly swell and fall apart when solvated. Many formulators use a combination of the actions just listed to achieve the desired objective. Most commercially available enteric materials fail to display two or more of the ideal properties of an enteric coating material. The following section discusses some of the difficulties encountered in enteric formulations.

The United States Pharmacopeia (USP) disintegration test for enteric coated tablets requires that the tablets tolerate agitation in simulated gastric fluid test solution at 37 ± 2°C (no discs). After 1 hour of exposure in simulated gastric fluid, tablets should show no evidence of disintegration, cracking, or softening. Then a disc is added to each tube, and the test is continued using simulated intestinal fluid maintained at 37 ± 2°C as the immersion fluid, for a period of time equal to 2 hours or to the time limit specified in the individual monograph. If all the tablets disintegrate, the product passes the test. If 1 or 2 tablets fail to disintegrate completely, the test is repeated on 12 additional tablets. To pass the disintegration test, at least 16 out of 18 tablets should disintegrate.

All enteric coated tablets must meet these requirements. Passing the USP enteric test does not guarantee optimal bioavailability of a particular dosage form. Several situations complicate the absorption of drug from enteric coated tablets. The pH of the stomach contents may vary from 1.5 to 4.0, with about 10% of the patients having achlorhydria. The amount of gastric fluid may vary between individuals, and for the same individual from time to time. Gastric residence time for the dosage form may range from less than

half an hour to more than 4 hours depending on the time of its administration, whether it was consumed with food, and if so, the type and quantity of food. The USP disintegration test does not require a qualitative or quantitative test for the active drug after agitation in artificial gastric fluid for 1 hour. Several commercially available enteric products passed the USP enteric test, but released varying amounts of drugs in simulated gastric fluid. Most acid-labile drugs need protection between pH values 1 and 5. The pH of material approaching pylorus is expected to be about 5. An ideal enteric polymer should dissolve or become permeable near and above pH 5.

A common problem associated with the retardant type of polymers (non-pH dependent solubility), which act by mechanical hydrophobicity, is that to provide enteric effect, the film might be so thick that if the dosage form travels too fast through the gastrointestinal tract, solubilization in intestinal fluids may never be achieved. Commercial products have failed the enteric test both for lack of gastric protection and for lack of solubility in intestinal fluids. Many others passed these in vitro tests, but failed to perform adequately when studied in vivo.

## *Cellulose Acetate Phthalate*

Cellulose acetate phthalate (CAP) has been widely used in the industry. It has the disadvantage of dissolving only above pH 6, and possibly delaying the absorption of drugs. It is also hygroscopic and relatively permeable to moisture and gastric fluids, in comparison with some other enteric polymers. CAP films are susceptible to hydrolytic removal of phthalic and acetic acids, resulting in a change of film properties. CAP films are brittle and usually formulated with hydrophobic-film forming materials or adjuvants to achieve a better enteric film. FMC Corporation has developed a patented aqueous enteric coating called Aquateric. Aquateric coating is a reconstituted colloidal dispersion of latex particles (not a solvent solution coating system). It is composed of solid or semisolid polymer spheres of cellulose acetate phthalate ranging in size from 0.05 to 3 microns with an average particle size of 0.2 micron. This material is currently being offered for potential industrial applications.

## *Acrylate Polymers*

Two forms of commercially available enteric acrylic resins are Eudragit L and Eudragit S. Both resins produce films that are resistant to gastric fluid. 

Eudragit L and S are soluble in intestinal fluid at pH 6 and 7, respectively. Eudragit L is available as an organic solution (Isopropanol), solid, or aqueous dispersion. Eudragit S is available only as an organic solution (Isopropanol) and solid.

## *Hydroxypropyl Methylcellulose Phthalate*

Shin-Etsu Chemical Company has made three enteric polymers available commercially. These are derived from hydroxypropyl methyl-cellulose, N.F., by esterification with phthalic anhydride, and are marketed as HPMCP 50, 55, and 55S (also known as HP-50, HP-55, and HP-55-S). HPMCP is the trade name for hydroxypropyl methylcellulose phthalate. These polymers dissolve at a lower pH (at 5 to 5.5) than CAP or acrylic copolymers, and this solubility characteristic may result in higher bioavailability of some specific drugs.[43] For general enteric preparations, HP-55 is recommended by Shin-Etsu; HP-50 and HP-55S are recommended for special situations. These polymers are quite stable compared with CAP because of their absence of labile acetyl groups.

## *Polyvinyl Acetate Phthalate*

Polyvinyl acetate phthalate (PVAP) is manufactured by the esterification of a partially hydrolyzed polyvinyl acetate with phthalic anhydride. This polymer is similar to HP-55 in stability and pH-dependent solubility. It is supplied as ready-to-use or ready-to-disperse enteric systems.

## Solvents

The primary function of a solvent system is to dissolve or disperse the polymers and other additives and convey them to the substrate surface. All major manufacturers of polymers for tablet coating provide basic physical-chemical data on their polymers. These data are usually helpful to a formulator. Some important considerations for an ideal solvent system are as follows:

1. It should either dissolve or disperse the polymer system.
2. It should easily disperse other coating solution components into the solvent system.
3. Small concentrations of polymers (2 to 10%) should not result in an extremely viscous solution system (>300 cps), creating processing problems.
4. It should be colorless, tasteless, odorless, inexpensive, nontoxic, inert, and nonflammable.
5. It should have a rapid drying rate (the ability to coat a 300 kg load in 3 to 5 hours).
6. It should have no environmental impact.

The most widely used solvents, either alone or in combination are water, ethanol methanol, isopropanol, chloroform, acetone, methy-lethylketone, and methylene chloride. Because of environmental and economic considerations, water is the solvent of choice; however, several polymers cannot be applied from aqueous systems. Drugs that readily hydrolyze in the presence of water can be more effectively coated with non-aqueous-solvent-based coatings. Such a process might require applying an initial sealing coat from an organic-based subcoating, followed by aqueous color and gloss coating. The use of organic-solvent-based film coatings will undoubtedly decrease as better aqueous systems are developed. It is unlikely, however, that organic solvents will be entirely supplanted.

## Plasticizers

The quality of a film can be modified by the use of “internal” or “external” plasticizing techniques. Internal plasticizing pertains to the chemical modification of the basic polymer that alters the physical properties of the polymer. By controlling the degree of substitution, the type of substitution, and the chain length, polymer properties can be altered significantly. Most often, the formulator uses external plasticizers as additives to the coating solution formula so that the desired effects are achieved for the film. An external plasticizer can be a nonvolatile liquid or another polymer, which when incorporated with the primary polymeric film former, changes the flexibility, tensile strength, or adhesion properties of the resulting film.

As the solvent is removed, most polymeric materials tend to pack together in threedimensional honeycomb arrangements. The choice of plasticizer depends upon the ability of plasticizer material to solvate the polymer and alter the polymer-polymer interactions. When used in correct proportion to the polymer, these materials impart flexibility by relieving the molecular rigidity. The type of plasticizer(s) and its ratio to the polymer can be optimized to achieve the desired film properties. One should also consider the viscosity of the plasticizer; its influence on the final coating solution; its effect on film permeability, tackiness, flexibility, solubility, and taste; and its toxicity, compatibility with other coating solution components, and stability of the film and the final coated product.

A combination of plasticizers may be needed to achieve the desired effect. The concentration of the plasticizers depends on many factors, including the polymer chemistry, method of application, and the other components present in the system. Even changes in the drying rate or use of elevated temperatures may alter the influence of the plasticizer in the coating process. The presence of titanium dioxide, colorants, flavors, and other additives also affect the film former. Most film formers tolerate only a certain additive load, and beyond that limit, the film properties are adversely affected.

The amount and type of plasticizers to be used for any given polymer can be based on the polymer manufacturer’s recommendations. Optimization of the plasticizer concentration must be based on the presence of the other additives. Concentration of a plasticizer is expressed in relation to the

polymer being plasticized. Recommended levels of plasticizers range from 1 to 50% by weight of the film former. Some of the commonly used plasticizers are castor oil; propylene glycol; glycerin; low-molecular-weight polyethylene glycols of 200 and 400 series; and surfactants, e.g. polysorbates (Tweens), sorbitan esters (Spans), and organic acid esters. With the increasing interest in aqueous coating, water-soluble plasticizers, e.g. polyethylene glycols and propylene glycol, are used. Conversely, castor oil and Spans are used primarily for organic-solvent-based coating solutions. For an external plasticizer to be effective, it should be soluble in the solvent system used for dissolving the film former and plasticizer. The plasticizer and the film former must be at least partially soluble or miscible in each other.

## Colourants

Coating solution formulations may contain a wide variety of components in addition to the film former, solvents, and plasticizers. Colorants may be soluble in the solvent system or suspended as insoluble powders. They are used to provide distinctive color and elegance to a dosage form. To achieve proper distribution of suspended colorants in the coating solutions requires the use of fine-powdered colorants (<10 microns). Repetitive production of colored coating solutions from different lots of the same colorant can be particularly difficult if colorant lots have different dye content, crystal form of dye, or particle size distribution. In general, the suspended colorants must be milled in the coating solvent or solution to attain a uniform dispersion of the colorants. Color variation in a product can be readily detected by the pharmacist and patient; therefore, the colors must be reproducible and stable.

The most common colorants in use are certified Food Drug and Cosmetic (FD&C) or Drug and Cosmetic (D&C) colorants. These are synthetic dyes or lakes of dyes. Lakes are derived from dyes by precipitating with carriers, e.g. alumina or talc. Lakes have become the colorants of choice for sugar or film-coating systems, as more reproducible tablet colors are attainable. Most commercially available lakes contain 10 to 30% of the pure dye content, but some lakes approach up to 50%. An occasional problem with the lake system might be the use of a solvent system that dissolves the dye, thereby establishing a time- and temperature-dependent equilibrium for leaching the dye from the lake system. Use of pure dye is recommended in such cases.

The concentration of colorants in the coating solutions depends on the color shade desired, the type of dye (i.e. dye versus the lake of the dye), and the concentration of the opaquant-extenders. If a very light shade is desired, a concentration of less than 0.01% may be adequate. On the other hand, if a dark color is desired, a concentration of more than 2.0% may be required. Since lakes contain less colorant, a larger concentration in solution is generally required.

The inorganic materials (e.g. iron oxides) and the natural coloring materials (e.g. anthocyanins, caramel, carotenoids, chlorophyll, indigo, flavones, turmeric, and carminic acid) are also used to prepare coating solutions.

A new line of colorants is being developed. These colorants are

nonabsorbable in the biologic system. This is accomplished by attaching dyes to polymers that are too large to be absorbed in the gastrointestinal tract and yet are resistant to degradation in the gastrointestinal tract. A magenta red dye is projected to be the first dye to be cleared for use.

A variety of products that are commercially available permit preparation of coating solution without additional milling equipment. Some examples are:

Opalux—Opaquant color concentrate for sugar coating.

Opaspray—Opaque color concentrate for film coating.

Opadry—Complete film coating concentrate.

All of these concentrates are promoted as achieving less lot-to-lot color variation.

## Opaquant-extenders

These are very fine inorganic powders used in the coating solution formulations to provide more pastel colors and increase film coverage. These opaquants can provide a white coating or mask the color of the tablet core. Colorants are much more expensive than these inorganic materials, and effectively less colorant is required when opaquants are used. The most commonly used material for this purpose is titanium dioxide. Some other materials are silicates (talc, aluminum silicate), carbonates (magnesium carbonate), sulfates (calcium sulfate), oxides (magnesium oxide), and hydroxides (aluminum hydroxide).

Rowe and associates have observed differences in the refractive indices of polymer film formers, pigments, and other additives commonly used for film coating. These observations have implications for the use of pigments and additives in the production of opaque films with good hiding power and film-coated tablets with highlighted intagliations.

## Miscellaneous Coating Solution Components

To provide a dosage form with a unique characteristic, special materials may be incorporated into the coating solution. Flavors or sweeteners are added to mask objectionable odors or to enhance a desired taste. Surfactants are used to solubilize otherwise immiscible or insoluble ingredients, or to facilitate faster dissolution of the coating. Antioxidants are incorporated to stabilize a dye system to oxidation and color change. Antimicrobials are added to prevent microbial growth in the coating composition during its preparation and storage, and on the coated tablets. Some aqueous cellulosic coating solutions are particularly prone to microbial growth, and prolonged storage of the coating composition should be avoided.

## COATING PROCESS

In most cases, the coating process is the last critical step in the tablet production cycle. Tablet coating is the application of a coating composition to a moving bed of tablets with the concurrent use of heated air to facilitate evaporation of the solvent. The distribution of the coating is accomplished by the movement of the tablets either perpendicular (coating pan) or vertical (air suspension coater) to the application of the coating composition._The successful application of the coating solution formula to a tablet provides the visual characteristics for the product; thus, the quality of the product may be judged on this final production step. The coating process can be divided in to two types; i.e. sugar coating and film coating. The type of process chosen depends on the type of coating that is to be applied, the durability (toughness) of the tablet core, and the economics of the process. Because of the ever-increasing cost of energy and labor, the cost of organic solvents, and the associated environmental constraints, the economics of the process is receiving greater emphasis. Sugar coating is still a widely used coating process because of the excellent tablet appearance it achieves.

## Coating Equipments

Most coating processes use one of three general types of equipment (1) the standard coating pan, (2) the perforated coating pan, or (3) the fluidized bed (air suspension) coater. The general trend has been toward energy-efficient, automated systems to shorten the total coating time and reduce operator participation in the coating process. In addition, several pharmaceutical companies have developed their own coating equipment or made modifications in standard equipment to accommodate their particular coating processes. Most of the systems, however, are based on three basic designs.

### *Conventional Pan System*

The standard coating pan system consists of a circular metal pan mounted somewhat angularly on a stand. The pan is 8 to 60 inches in diameter and is rotated on its horizontal axis by a motor (Fig. 13.22). Heated air is directed into the pan and onto the tablet bed surface, and is exhausted by means of ducts positioned through the front of the pan (Fig. 13.23). Coating solutions are applied to the tablets by ladling or spraying the material onto the rotating tablet bed. Use of atomizing systems to spray the liquid coating material onto the tablets produces a faster, more even distribution of the solution or suspension. Spraying can significantly reduce drying time between solution applications in sugar coating processes and allows for continuous application of the solution in film coating.

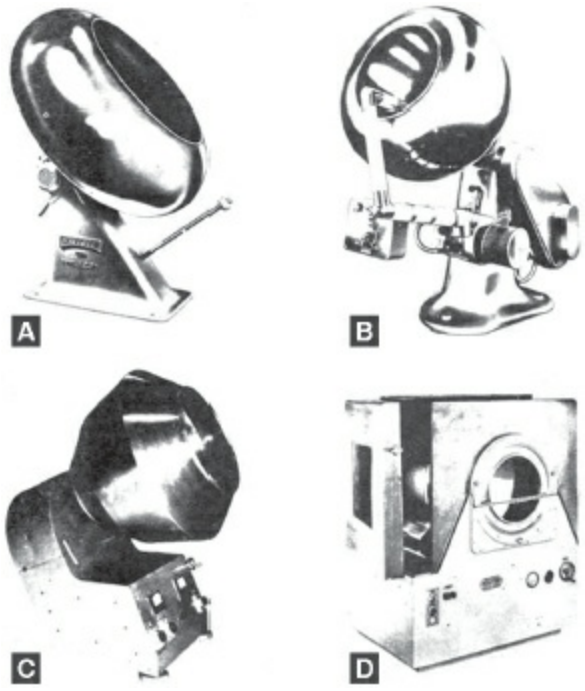



**Fig. 13.22:** Standard coating pans

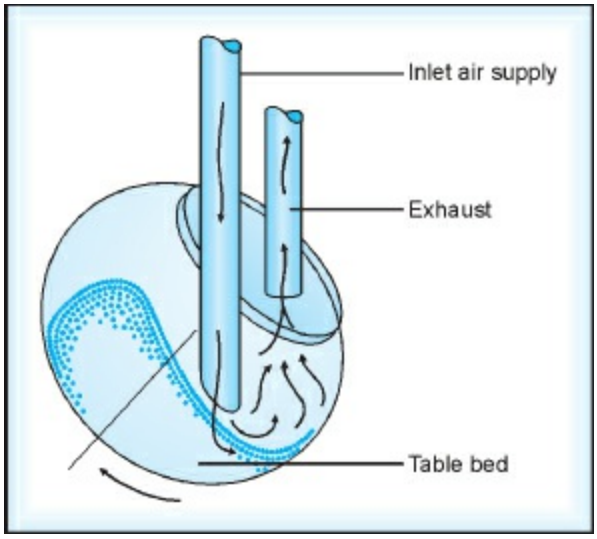



**Fig. 13.23:** Line diagram of standard coating pans

A significant improvement in the drying efficiency of the standard coating pan is achieved by the Pellegrini pan (Fig. 13.24), the immersion-sword (Fig. 13.25), and the immersion-tube systems (Fig. 13.26). The Pellegrini system has a baffled pan and a diffuser that distributes the drying air uniformly over the tablet bed surface. Newer models are completely enclosed, which further increases their drying efficiency and facilitates

automated control. With the immersion-sword system, drying air is introduced through a perforated metal sword device that is immersed in the tablet bed. The drying air flows upward from the sword through the tablet bed. Since the air is more intimately mixed with the wetted tablets, a more efficient drying environment is provided. Coating solutions are applied by an atomized spray system directed to the surface of the rotating tablet bed. With the immersion-tube system, a tube is immersed in the tablet bed. The tube delivers the heated air, and a spray nozzle is built in the tip of the tube. During this operation, the coating solution is applied simultaneously with the heated air from the immersed tube. The drying air flows upward through the tablet bed and is exhausted by a conventional duct. Relatively rapid processing times have been reported for both film and sugar coating with this system.

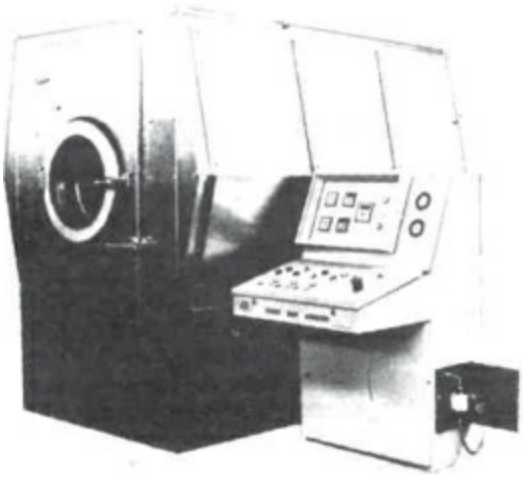


**Fig. 13.24:** Pellegrini pan (enclosed) system

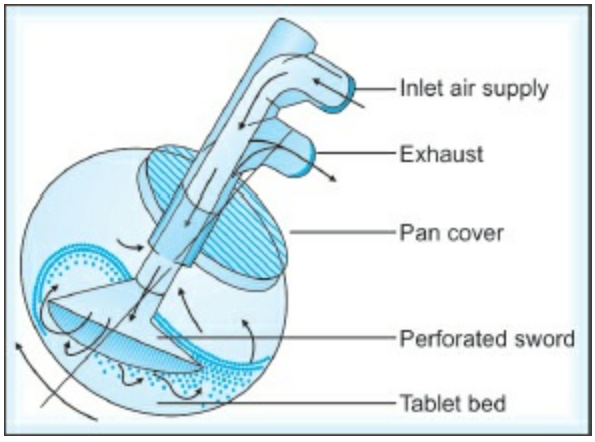



**Fig. 13.25:** Line diagram of Glatt immersion-sword system

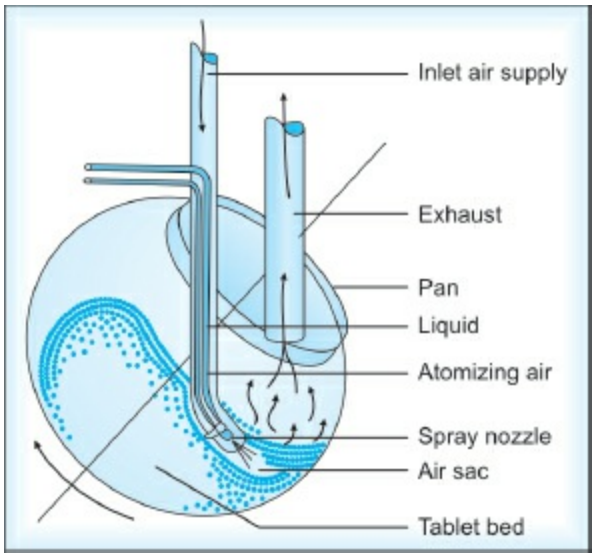



**Fig. 13.26:** Line diagram of immersion-tube system

## *Perforated Pan Systems*

In general, all equipment of this type consists of a perforated or partially perforated drum that is rotated on its horizontal axis in an enclosed housing. In the Accela-Cota and Hi-Coater systems, drying air is directed into the drum, is passed through the tablet bed, and is exhausted through perforations in the drum (Figs 13.27 and 13.28). The Driacoater introduces drying air through hollow perforated ribs located on the inside periphery of the drum (Fig. 13.29). As the coating pan rotates, the ribs dip into the tablet bed, and drying air passes up through and fluidizes the tablet bed. Exhaust is from the back of the pan.

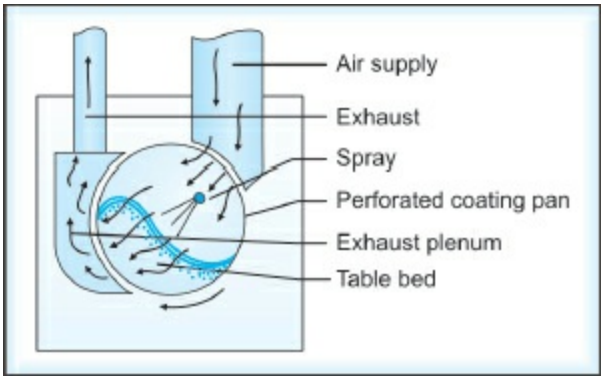



**Fig. 13.27:** Line diagram of Accela-Cota system

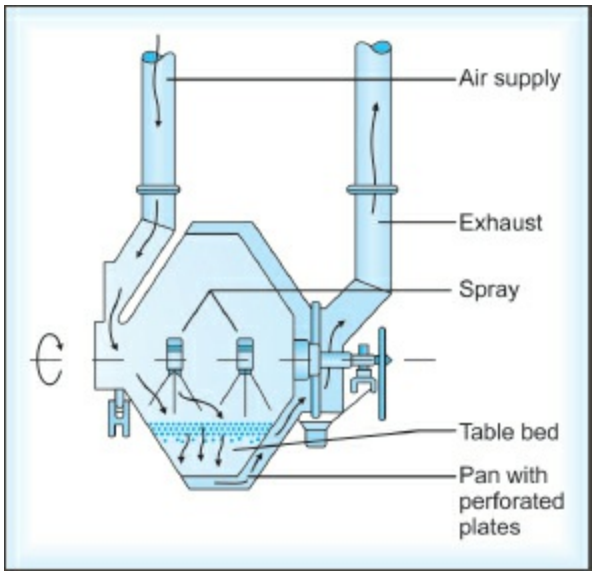



**Fig. 13.28:** Line diagram of Hi-Coater system

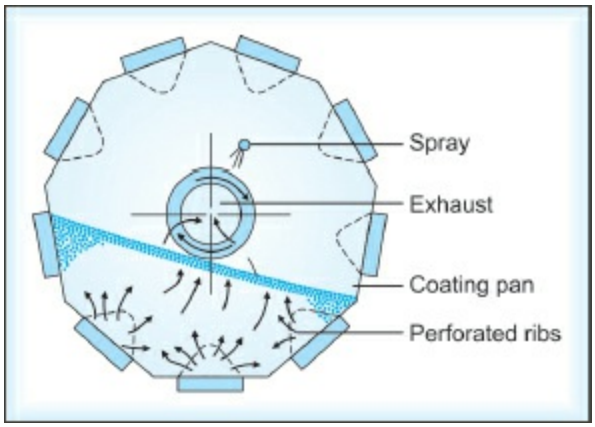



**Fig. 13.29:** Line diagram of driacoater pan

The Glatt coater is the latest perforated pan coater to be introduced in the industry (Fig. 13.30). In the Glatt coater, drying air can be directed from inside the drum through the tablet bed and out an exhaust duct; alternatively, with an optional split-chambered plenum, drying air can be directed in the reverse manner up through the drum perforations for partial fluidization of the tablet bed. Several airflow configurations are possible.

In all four of these perforated pan systems, the coating solution is applied to the surface of the rotating bed of tablets through spraying nozzles that are positioned inside the drum.

**Fig. 13.30:** Photograph of Glatt coater

Perforated pan coaters are efficient drying systems with high coating capacity, and can be completely automated for both sugar coating and film coating processes.

## *Fluidized Bed (Air Suspension) Systems*

Fluidized bed coaters are also highly efficient drying systems. Fluidization of the tablet mass is achieved in a columnar chamber by the upward flow of drying air (Fig. 13.31). The airflow is controlled so that more air enters the center of the column, causing the tablets to rise in the center. The movement of tablets is upward through the center of the chamber. They then fall toward the chamber wall and move downward to re-enter the air stream at the bottom of the chamber. In some units, a smaller column(s) is used to direct tablet movement within the main column. Coating solutions are continuously applied from a spray nozzle located at the bottom of the chamber or are sprayed onto the top of the cascading tablet bed by nozzles located in the upper region of the chamber.

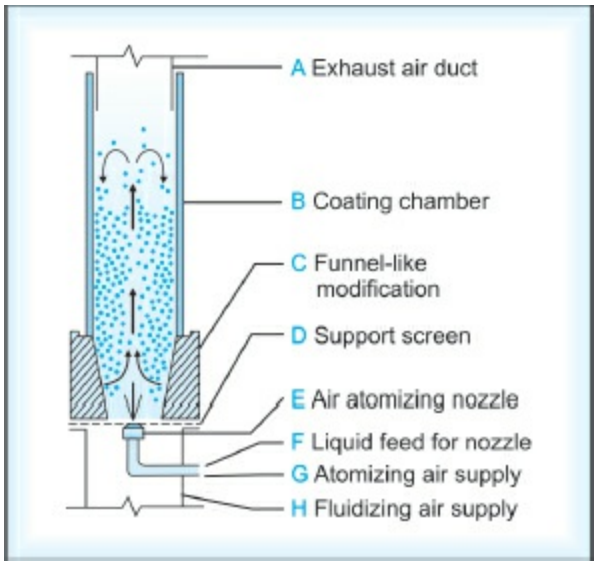



**Fig. 13.31:** Line diagram of a fluidized bed coater

Tablet cores that are friable and prone to chipping and edge abrasion may be difficult to coat even under optimum conditions in the fluidized bed systems, owing to the relatively rough tablet-to-tablet impact and tablet-chamber contact.

## *Spray Application Systems*

The two basic types of systems used to apply a finely divided (atomized) spray of coating solutions or suspensions onto tablets are (1) high-pressure, airless and (2) low-pressure, air-atomized. The principal difference in the two types is the manner in which atomization of the liquid is achieved.

In the airless spray system, liquid is pumped at high pressure (250 to 3,000 pounds per square inch gauge (psig) through a small orifice (0.009 inch to 0.020 inch id) in the fluid nozzle (Fig. 13.32), which results in a finely divided spray. The degree of atomization and the spray rate are controlled by the fluid pressure, orifice size, and viscosity of the liquid. Because of the small orifice, suspended solids in the coating composition must be finely milled or filtered to prevent orifice blockage.

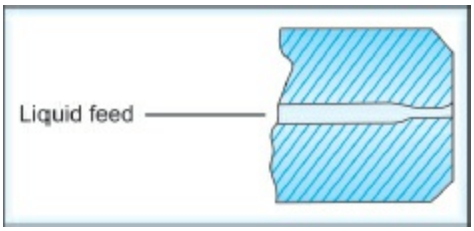



**Fig. 13.32:** Simplified diagram of a high-pressure, airless nozzle

In the low-pressure air-atomized system, liquid is pumped through a somewhat larger orifice (0.020 inch to 0.060 inch id) at relatively low pressures (5 to 50 psig) (Fig. 13.33). Low-pressure (10 to 100 psig) air contacts the liquid stream at the tip of the atomizer, and a finely divided spray is produced. The degree of atomization is controlled by the fluid pressure, fluid cap orifice, viscosity of the liquid, air pressure, and air cap design.

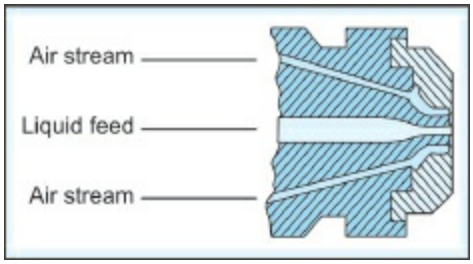



**Fig. 13.33:** Simplified diagram of a low-pressure, air-atomized nozzle

Both airless and air atomizing systems can be used effectively. Originally, airless systems were primarily used in air suspension coaters, but now the choice depends on the coating solution formula and on the process developed for a particular product.

## Facility and Ancillary Equipment

The facility required for any coating operation should be designed to meet the requirements of current Good Manufacturing Practices (GMPs) as set forth in the latest revision of the Code of Federal Regulations, Title 21, Part 211. An execution of coating process in a pharmaceutical set up is shown in Fig. 13.34. Adequate space is needed not only for the coating equipment, but also for solution preparation and in-process storage. The specific safety requirement for coating areas depends on the nature of the solvent. Where explosive or toxic concentrations of organic solvent could occur, during either solution preparation or the coating operation, electrical explosion-proofing and specialized ventilation are required.

Treatment of the exhaust air from the coating operation may be desired to recover expensive organic solvents or to prevent solvents and particulate from entering the atmosphere. Local, state, and federal Environmental Protection Agency (EPA) regulations define the limits of organic solvent and particulate allowed in the atmosphere.

**Fig. 13.34:** Coating of tablets in a pharmaceutical setup (*Courtesy* of Unicure)

Compliance with the regulations can be extremely expensive, and this cost factor should be considered in developing a new coating. A major

advantage of totally aqueous-based film coating is that all direct and indirect expenses relating to the purchase, handling, and environmentally acceptable removal of the organic solvent are circumvented.

Other equipment is needed to support the coating operation. Solution preparation requires tanks, filters, and mixers. A colloid mill or ball mill may be needed for the homogeneous dispersion of insoluble solids in the liquid coating mixture. Jacketed tanks may be needed for keeping some solutions at an elevated temperature.

The coating liquid can be supplied to the nozzle system of the coating equipment by means of portable pressure tanks or various pumping systems.

## Automation

There is little published information providing details of automated coating processes. A review article by Thomas discusses details of a programmable controller for pan coating systems. Within the last 6 to 8 years, automation has been achieved in sugar coating and film coating (nonaqueous and aqueous) systems. Through a series of sensors and regulating devices for temperature, airflow, spray rate and pan speed, a feedback control of the process is maintained. Precise automated control of such a dynamic process is possible only with the help of the programmable controller. As in all automated processes, a manual bypass should be built into the system to accommodate any special applications or equipment malfunctions. For process automation, the perforated pans are preferred over the old conventional coating pans because of their better efficiency. Figure 13.35 represents a completely automated system used for film or sugar coating. As new tablet manufacturing plants are built by major pharmaceutical companies, varying degrees of automation are built into the tablet coating process. These automated coating systems are either designed by coating equipment manufacturers or developed by individual companies and tailored to their specific equipment and/or products.

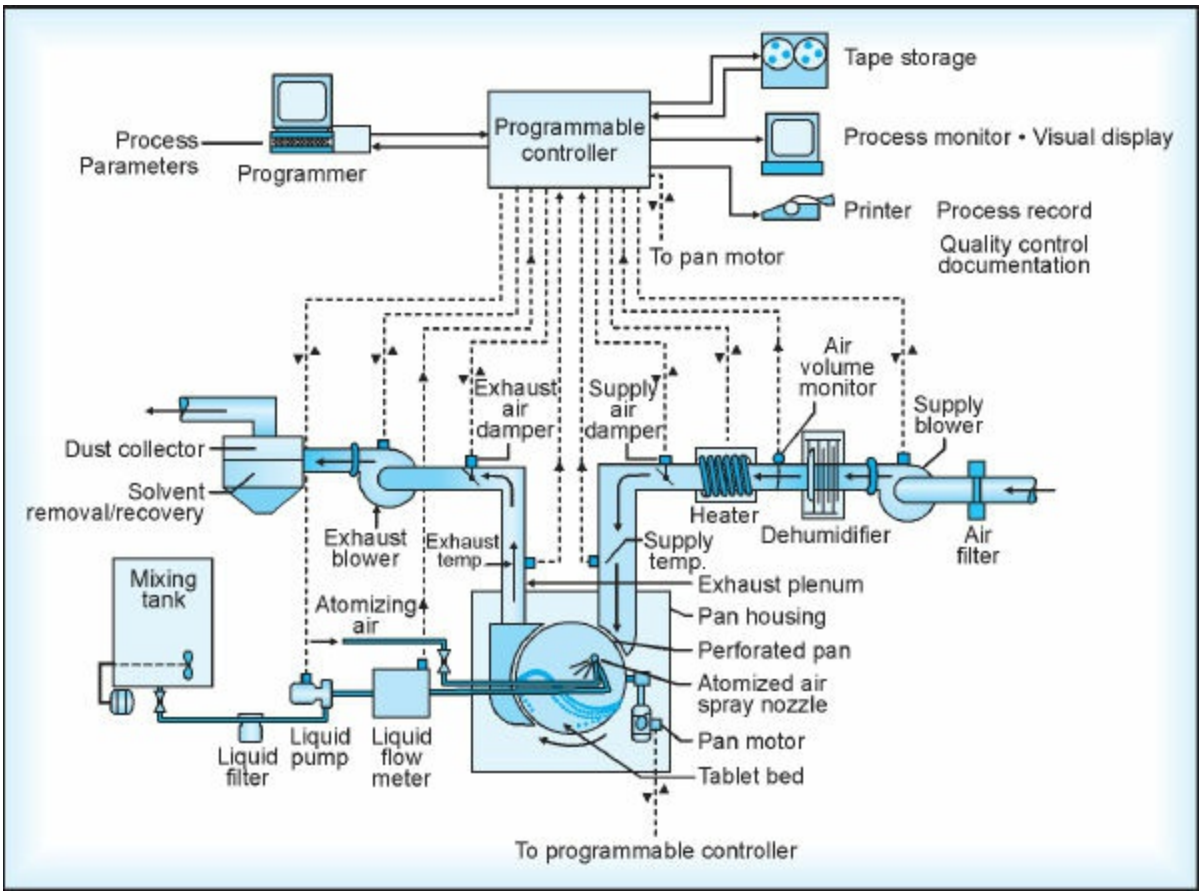



**Fig. 13.35:** Diagram of an automated coating system

## Sugar Coating Process

The sugar coating process involves several steps, the duration of which ranges from a few hours to a few days. A successful product greatly depends on the skill of the coating operator. This is especially true in the pan-ladling method, in which the coating solutions are poured over the tablet cores. The operator determines the quantity of solution to add, the method and rate of pouring, when to apply drying air, and how long or how fast the tablets should be tumbled in the pan. Newer techniques utilize spraying systems and varying degrees of automation to improve coating efficiency and product uniformity. Regardless of the methods used, a successful sugar coating process yields elegant, highly glossed tablets.

The basic sugar coating process involves the following steps (1) sealing, (2) subcoating, (3) syruping (smoothing), (4) finishings and (5) polishing.

The tablet cores preferably have deep convex surfaces with thin rounded edges to facilitate sugar coating. Since sugar coating tends to be long and vigorous, the cores should be relatively resistant to breakage, chipping, and abrasion.

### *Seal Coating*

To prevent moisture penetration into the tablet core, a seal coat is applied. This is especially needed in pan-ladling processes, in which localized overwetting of a portion of the tablet bed occurs. Without a seal coat, the overwetted tablets would absorb excess moisture, leading to tablet softening or disintegration and affecting the physical and chemical stability of the finished product. In spray processes, it is possible to adjust the application of the subcoats and further coats so that localized overwetting does not occur. This adjustment thus eliminates the seal coating step. Shellac is an effective sealant, but tablet disintegration and dissolution times tend to lengthen on aging because of the polymerization of the shellac. Zein is an alcohol-soluble protein derivative from corn that has also been used as an effective sealant. Lengthening dissolution times have not been reported on aging of zein seal coated tablets.

### *Subcoating*

The subcoating is applied to round the edges and build up the tablet size.

Sugar coating can increase the tablet weight by 50 to 100%. The subcoating step consists of alternately applying a sticky binder solution to the tablets followed by a dusting of subcoating powders and then drying. Subsequent subcoats are applied in the same manner until the tablet edges have been covered and the desired thickness is achieved. For spray processes, a subcoating suspension containing both the binder and the insoluble powder is sprayed intermittently on the tablet bed. With both methods of application, control of the drying rate is critical to obtaining a rapid application of the subcoat.

## *Syrup (Smoothing/Color) Coating*

The purpose of this step is to cover and fill in the imperfections in the tablet surface caused by the subcoating step, and to impart the desired color to the tablet. This step perhaps requires the most skill. The first syrup coats usually contain some suspended powders and are called “grossing syrups.” Dilute colorants can be added to this phase to provide a tinted base that facilitates uniform coloring in later steps. In general, no color is added until the tablets are quite smooth; premature application to rough tablets can produce a mottled appearance in the final coated tablets. In subsequent syruping steps, syrup solutions containing the dye are applied until the final size and color are achieved. In the final syruping or finishing step, a few clear coats of syrup may be applied.

## *Polishing*

The desired luster is obtained in this final step of the sugar coating process. The tablets can be polished in clean standard coating pans, or canvas-lined polishing pans (Fig. 13.36), by carefully applying powdered wax (beeswax or carnauba) or warm solutions of these waxes in naphtha or other suitable volatile solvents.

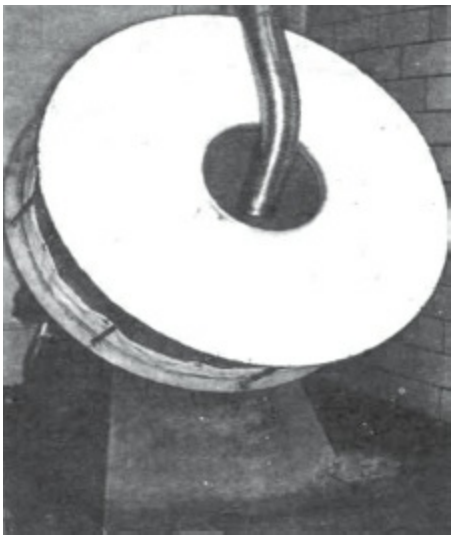


**Fig. 13.36:** Canvas-lined polishing pan

The basic sugar coating process is illustrated in this example. An infinite number of variations in the materials and processes are possible; however, the complexity of the process can be appreciated by the following example.

## I. Materials and Equipment

a. *Coating pans*—Stainless steel, 40 inches in diameter, with variable speed control; or 48-inch Accela-Cota, with 2 to 3 air atomizing nozzles. Nozzles should have a fluid orifice of .040 to .060 inch. Set atomizing air to 30 to 40 psig.

b. *Tablet cores*—55 to 70 kg of 3/8-inch standard convex tablets.

If desired, coating solutions may be poured or ladled onto the tablets. If this method is chosen, apply the solutions in a steady flow, with even distribution over the rotating tablet bed.

## II. Process

a. *Seal coat*—The specific advantage of using a spray system described in this example is that a faster and more even distribution of the coating materials is obtained. Start the tablets rolling (pan speed: 10 rpm). Set supply air to 30°C. Apply 3 applications of zein solution (Table 13.10), 800 ml per application. Allow 15 to 20 min between applications to ensure that the tablets are dry. If tablets become tacky between applications, apply just enough talc to prevent sticking to the pan and to each other. Make sure that the solution is well distributed. Additional mixing by hand may be necessary to achieve this if pan design and baffling are inefficient.

b. *Subcoat*—Use any of the gelatin/acacia solutions and subcoating powders listed in Table 13.10.

   1. Turn heat and inlet air off. Use exhaust only. Start pan speed at 10 rpm.
   2. Apply 3 to 9 coats. Use 1.5 L of warm-gelatin/acacia solution for the first coat. Reduce subsequent amounts accordingly to obtain the correct thickness. Be sure that edges are covered. Thickness is checked volumetrically.
   3. Allow at least 20 min between coats to permit adequate drying. Be sure the solution is rapidly and uniformly dispersed in the tablet bed. Dust with subcoating powder when tackiness develops. Apply subcoating powder until tablets roll freely and show no signs of tackiness.
   4. After the last coat, jog the pan periodically for at least 2 to 4 hours to ensure dryness.

c. *Syrup (Smoothing/Color) Coat*—The syruping coat usually involves three basic phases: grossing syrup (a syrup solution with subcoating powders dispersed in it), heavy syrup, and regular syrup. Apply each step in the sequence outlined.

   1. Remove excess dust in pan before starting. Turn on exhaust outlet air. Set inlet air temperature to provide an exhaust temperature of 45 to 48°C. Set pan speed at 12 rpm.
   2. Apply 5 to 15 coats of the grossing syrup, just enough to wet the entire

bed. Because this solution dries relatively quickly, uniform, rapid distribution must be provided. Apply successive amounts of grossing syrup immediately after each preceding application is drying and slightly dusty.

3. Apply several heavy-colored syrup coats in a similar manner until a specific tablet volume is attained.
4. Turn off heat, and reduce inlet and exhaust air.
5. Apply several coats of the regular-colored syrup solution to achieve a final smoothness, size, and color development.
6. Each coat of regular-colored syrup is applied as soon as the tablets exhibit a slightly frosted appearance. Do not allow them to become dusty.

d. *Finishing*

1. Make sure that the pan is clean.
2. Operate pan with the heat turned off, no supply, and greatly reduced exhaust air. Set pan speed at 12 rpm.
3. Apply 3 or 4 coats of regular-colored syrup rapidly, without permitting the tablet bed to frost or become dusty.

**Table 13.10:** Formulations used in sugar coating

| Seal coating solutions | Formula variation | |
|---|---|---|
| | I | II |
| Cellulose acetate phthalate | | 175 g |
| Zein | 480 g | |
| Oleic acid, USP | 60 g | |
| Propylene glycol, USP | | 52.5 g |
| Polyethylene glycol 4000 | 144 g | |
| Methylene chloride | 480 ml | 840 ml |
| Alcohol SD 3A 200-proof | q.s. to 2.4 L | q.s. to 1.75 L |

| Subcoating solutions | Formula variation | | | |
|---|---|---|---|---|
| | I | II | III | IV |
| Gelatin | 60 g | 5.4 kg | | 60 g |
| Acacia | 60 g | 2.7 kg | 450 g | 60 g |
| Sugar, cane | 1500 g | 53.7 kg | | |
| Syrup, corn | | | 450 g | 1500 g |
| Syrup, USP | | | 3.785 L | |
| Water, distilled | 1.0 L | 44.3 kg | | 1.0 L |

| Subcoating powders | Formula variation | | | | | | |
|---|---|---|---|---|---|---|---|
| | I | II | III | IV | V | VI | VII |
| Kaolin | | 225 kg | | | | | |
| Dextrin | | 112 kg | 185 kg | | | | |
| Cocoa powder | | 60 kg | | | | | |
| Calcium carbonate, pptd | | | 480 kg | | 7.72 kg | | |
| Sugar, cane, powdered | 4.1kg | 112 kg | 240 kg | 40 kg | 0.9 kg | 180 g | 8.62 kg |
| Acacia, powdered | 0.12 kg | | | 6 kg | 1 g | 0.86 kg | |
| Starch, corn | 1.35 kg | | | | | 0.9 kg | 60 g |
| Talc, USP | 0.23 kg | | | | | | 1 g |
| Calcium sulfate, NF | | | | | | | 8.62 kg |

| Syrup solutions | Grossing syrups | | | Heavy syrup | Regular syrups | | |
|---|---|---|---|---|---|---|---|
| | I | II | III | IV | V | VI | VII |
| Colorant | q.s. ad | q.s. ad | q.s. ad | q.s. ad | q.s. ad | q.s. ad | q.s. ad |
| Subcoating powder | 22.7 kg | | | | | | |
| Calcium carbonate, light | | 7.75 kg | 69 g | | | | |
| Sugar, cane, powder | 136 kg | 22.7 kg | 572 g | 2.73 kg | 181 kg | 85 g | 1.2 kg |
| Starch, corn | | 136 kg | 69 g | | | | |
| Syrup, USP | | 22.7 L | 3.785 L | 256 | 256 kg | | |
| Water, distilled | 76 kg | | 290 ml | | | q.s. 100 ml | 1.0 L |

| Polishing solutions | Formula variation | |
|---|---|---|
| | I | II |
| Wax, carnauba, yellow | 0.09 kg | 10 g |
| Beeswax, white | 0.09 kg | 90 g |
| Wax, paraffin | 0.02 kg | |
| Naphtha | 3.785 L | 1.0 L |

4. The last coats of regular syrup can be applied without colorant. This gives “depth” to the color and enhances the elegance of the coat.
5. Shut off exhaust air before applying the last coat. Apply coat; mix uniformly and shut off pan while the tablets are still damp. A quick jog every few minutes prevents sticking. After 15 to 30 min, stop jogging and leave tablets in pan to dry slowly overnight.

e. *Polishing:* Polishing can be done in the same pan as the sugar coating, but better results are obtained in canvas-lined pans.
   1. Supply air, exhaust air, and heat should be turned off. Pan speed 12 rpm.
   2. Apply 3 to 4 coats of warm polishing solution, approximately 300 ml per application.
   3. Let solvent evaporate completely between coats.

Tablet coatings achieve their luster during the polishing phase. In the canvas-lined pans, the lining is used to transfer the waxes to the tablet surface and to provide a buffing action. The wax polishing solutions are usually poured onto the canvas, and the tablets pick up the wax shine as they tumble in the pan. The waxes can also be dusted onto the tablets. Care must be exercised to distribute the wax evenly to avoid wax spots on some tablets. Application of warm air can facilitate distribution.

The techniques used to obtain the desired product, especially in the pan-ladling process, are complex and can only be learned through practice. The beginner is advised to consult the listed literature for specific techniques, materials, and precautions of the sugar coating processes. The use of modern efficient automated systems is rapidly making manual techniques obsolete. Several automated processes have been described in the literature.

Through the addition of cellulosic polymers, and other coating ingredients normally associated with film coating, much thinner sugar coatings have been attained.

## Film Coating Process

Since film coating originated from the pan sugar coating era, it is not surprising that the film coating process today still contains some of the process features associated with the early work. With the possible exception of the air suspension coater, film coating and sugar coating share the same equipment and process parameters.

### *Pan-pour Methods*

Pan-pour methods have been used for many years for film coating, but they have been supplanted by newer coating techniques that are faster and more reproducible. Coating compositions used in the earlier pan-pour methods were usually too viscous to be sprayed effectively. Tablets coated by panpour methods are subjected to alternate solution application, mixing and drying steps similar to pan-pour sugar coating. The method is relatively slow and relies heavily on the skill and technique of the operator to balance the steps to produce an acceptable product. Tablets that are film coated by pan-pour processes almost always require additional drying steps to remove latent solvents. Aqueous-based film coatings are not suitable for this method of application because localized over-wetting inherent with the panpour process causes numerous problems ranging from surface erosion to product instability due to unacceptably high latent moisture content in the cores.

### *Pan-spray Methods*

The introduction of spraying equipment was the next evolution in improving the efficiency of film coating processes. Spraying lends versatility to the process and allows for automated control of liquid application. Spray patterns are selected to provide a continuous band across the tablet bed surface. Broad, flat spray patterns are usually chosen by selection of appropriate nozzle systems so that the entire width of the tablet bed can be covered by the spray from 1 to 5 nozzles.

## Process Variables

Whether the coating process is in a conventional pan system or in one of the perforated pan systems previously described, certain elements of the process need to be controlled to ensure consistent product quality. The process is as important as the coating solution formulation; consequently development of a well-defined and well-controlled process should be a major concern of the formulator.

The variables to be controlled in pan-spray film coating processes are:

1. *Pan variables*

   Pan design/baffling

   Speed

   Pan load

2. *Process air*

   Air quality Temperature

   Airflow rate/volume/balance

3. *Spray variable*

   Spray rate

   Degree of atomization

   Spray pattern

   Nozzle-to-bed distance

Since each listed variable is important to the overall success of the coating, further discussion is warranted.

## *Pan Variables*

Pan shape, baffling, rotational speed, and loading all affect the mixing of the tablet mass. Uniform mixing is essential to depositing the same quantity of film on each tablet. The tablet coating adds an approximate increase in weight of only 2 to 5% to the tablet. Unacceptable color uniformity or enteric film integrity is encountered if the tablets are inadequately coated because of poor tablet movement in the coating pan.

Tablet shape can also affect mixing. Some tablet shapes may mix freely while other shapes may require a specific baffling arrangement to ensure 

adequate mixing. Baffles, however, provide a source for chipping and breakage if they are not carefully selected and used.

Pan speed affects not only mixing, but also the velocity at which the tablets pass under the spray. Speeds that are too slow may cause localized overwetting, resulting in the tablets sticking to each other or to the pan. Speeds that are too high may not allow enough time for drying before the same tablets are reintroduced to the spray; again, this results in a rough coating appearance on the tablets. Pan speeds of 10 to 15 rpm are commonly used in the large pan coaters for nonaqueous film coating. Slower pan speeds (3 to 10 rpm) are used for aqueous film coating primarily to accommodate slower application rate and drying of the coating liquid. Selection of pan operating conditions depends on the equipment availability, type of tablets being coated, and the characteristics of the coating solution.

## *Spray Variables*

The spray variables to be controlled are the rate of liquid application, the spray pattern, and the degree of atomization. These three variables are interdependent. In the airless, high-pressure system, all three variables are directly affected by fluid pressure and nozzle design. In the air-atomized, low-pressure system, the rate of liquid flow is most directly affected by the liquid pressure and liquid orifice size. The degree of atomization and spray pattern are most directly affected by atomizing air pressure, air volume, and the shape and design of the air jets in relation to the fluid stream.

The proper rate at which the coating solution should be applied depends on the mixing and drying efficiency of the system, in addition to the coating formula and core characteristics. There is a range in which the coating rate must operate to achieve the desired product quality or processing time. Overwetting and underwetting must be avoided in all coating operations.

A band of spray should be spread evenly over the tablet mass. In larger pans, more nozzles must be added to cover die tablet bed width. A spray pattern that is too wide could result in the application of coating directly to the pan surface, producing lower coating efficiency and wasted material. If the spray pattern is too narrow, localized overwetting may result, and the tablet-to-tablet coating uniformity will be poor. Thus, tablets need to make many more passes through the spraying area to be adequately coated. During the coating operation, the spray width can be adjusted by moving the nozzles

closer or farther away from the tablet bed. In the air-atomized, low-pressure systems, adjusting the air pressure and/or direction accomplishes the same effect. The distance that the nozzle is from the tablet bed affects not only the spray width, but also the quantity of coating applied to individual tablets per pass under the spray.

Atomization is the process whereby the liquid stream is finely subdivided into droplets. The degree of atomization—the size and size distribution of the droplets—obtained from the spray nozzle is not an easily controllable parameter.

The relationships between the orifice size, nozzle configuration, fluid pressure, atomizing air pressure, air volume, and fluid viscosity vary with each coating formulation. Manufacturing literature may provide the droplet size range expected from a particular nozzle type based on water; however, this type of data is inadequate for optimizing the nozzle performance in relation to the variety of solutions and suspensions' used to coat tablets.

The degree of atomization, at present, can only be controlled empirically. Adjustments of either the fluid pressure on the airless high-pressure systems or the atomizing air pressure and air volume on the low-pressure systems change the degree of atomization. Higher pressures yield greater atomization. Atomization that is too fine causes some droplets to dry before reaching the tablet bed. This "spray-drying" effect can be readily detected as roughness on the tablet surface, especially in intagliations or as excess dust in the pan. Insufficient atomization may result in droplets that are too large reaching the tablet surface and causing localized overwetting, which could lead to sticking, picking, or a rough "orange-peel" effect.

## *Process Air Variables*

The temperature, volume, rate, quality, and balance are parameters of the process air that need to be controlled to obtain an optimum drying environment for a particular coating process. The sensitivity of the film former and product core to heat largely determines the upper temperature at which the coating process is successful. In general, higher tablet bed and coating chamber temperatures are more conducive to rapid solvent evaporation, and consequently to faster coating rate. The limits to the air volume and rate depend on the overall design of the air-handling system and coating equipment. The upper end of the system's range is used most often.

The more efficient the equipment design, the less air volume is needed for drying.

Supply air should have some degree of dehumidification. Seasonal fluctuations in the moisture content of incoming air can alter coating and drying conditions and possibly have adverse affects on the quality of the coating.

The balance between supply and exhaust airflow should be such that all dust and solvent are contained within the coating system.

## Fluidized Bed Process

The fluidized bed systems have been successfully used for rapid coating of tablets, granules, and capsules. The coating solution formulations used with these processes are similar to those used for the pan processes. Since air is used to move the tablets in the coating process, there are some specific process controls unique to air suspension coaters.

The chamber design, together with the process air, controls the fluidization pattern. Tablet shape, size, density, and quantity of load affect the ability of the tablet mass to be fluidized.

Adequate fluidization and drying depend on the volume and rate of the process air. Control of the process air is achieved by adjusting a variable speed blower or by using dampers to keep the tablet mass in a constant “fluid” motion inside the chamber. Too high an airflow results in excess tablet attrition and breakage. If the airflow rate is too low, the mass does not move fast enough through the spray region, and overwetting may occur. Fluidization may also be affected by the increase in weight or by changes in the factional characteristics of the tablets during coating application. Consequently, periodic adjustment of the rate and volume will be necessary to maintain optimum fluidization.

During the coating operation, both the inlet and exhaust air temperatures are monitored. Evaporation of the solvent causes the exhaust air temperature to be cooler than the inlet. Any change in the rate of application of the coating solution can be monitored by the difference between the inlet and exit air temperatures.

## *Examples*

The literature available from pharmaceutical polymer manufacturers provides numerous coating formulas utilizing their particular polymers. These formulas require the effective utilization of the favorable properties of the various polymers, plasticizers, and additives in the final coating composition to acquire a quality coated tablet. Representative examples of organic and aqueous-based film coating formulations for use in laboratory perforated coating pan are provided.

## I. Materials and equipment

Standard 24-inch Accela-Cota with 2 baffles. Spray system—Air-atomized spray nozzle with a .040-inch fluid orifice and a flat spray air cap.

Pumping system—Pressure tanks.

Coating materials—Description to follow (example 1).

Pan load—12 kg of ½-inch standard convex tablets.

Operating conditions—Set pan speed at 12 to 15 rpm. Adjust supply air temperature to give an exhaust air temperature of 30°C during spray application. Use 40 to 50°C for the aqueous coating systems. Atomizing air pressure should equal 30 to 50 psig.

## II. Process

1. Load tablets into pan. Attach and adjust the spray nozzle to spray on upper half of tablet bed.
2. Turn on heat, drying air, exhaust, and atomizing air.
3. Intermittently jog the pan while tablets are warming,
4. When exhaust temperature reaches 30°C, start spraying.
5. Apply 3.0 to 4.0 L of color solution at a rate of 70 to 100 ml/min. Adjust rate downward if tablets become tacky.
6. Apply 1.5 to 2.5 L of clear solution at a rate of 70 to l00 ml/min. Adjust rate downward if tablets become tacky. Allow tablets to dry in pan with air and heat on for 5 to 10 min.

## Non-aqueous Film Coating Formulas

### *Example 1: Hydroxypropyl Methylcellulose Nonaqueous Formula*

(This formula in which all the ingredients are solubilized in the solvents, can be applied by spraying or pouring systems.)

| | |
|---|---|
| HPMC, 15 cps | 4% |
| Propylene glycol, USP | 1.2% |
| Ethyl alcohol 200-proof | 45% |
| Methylene chloride | q.s. ad 100% |

The polymer is gradually added to the ethyl alcohol while the solvent is continuously agitated. A portion of the methylene chloride is added to this suspension, to solubilize the polymer. The propylene glycol is then added, and the remainder of the methylene chloride is added to obtain the proper volume.

The addition of insoluble colorants, opaquants, or flavors requires a milling step to facilitate their adequate dispersal.

### *Example 2: Cellulose Acetate Phthalate/Carbowax Nonaqueous Formula—Nonenteric*

(This formula should be poured, or diluted with appropriate solvents for spraying.)

| | |
|---|---|
| Cellulose acetate phthalate, NF | 5.0% |
| Polyethylene glycol 8000, NF | 15.0% |
| Sorbitan monooleate, NF | 0.3% |
| Dye yellow, D and C Lake # 5 | 0.05% |
| Titanium dioxide | 0.5% |
| Vanillin | 0.1% |
| Castor oil | 0.25% |
| Ethyl alcohol 200-proof | 12.0% |
| Acetone | q.s. ad 100.0% |

The cellulose acetate phthalate is dissolved in the ethyl alcohol, sorbitan monooleate, and part of the acetone. To ensure proper dispersion, the dye, titanium dioxide, and vanillin are milled in a ball or high-energy mill, or are dispersed in acetone using a colloid mill. After the particle size reduction or dispersion has occurred, the colorants are added to the solution containing the polymer. The polyethylene glycol 8000 is melted and added with the castor oil to the polymer dispersion. The composition is brought to proper volume with acetone. This preparation must be kept slightly warm and must be properly agitated to assure proper distribution of the polyethylene glycol 8000 and the colorants in suspension.

### *Example 3: Cellulose Acetate Phthalate Enteric Solution*

(This system could be sprayed or poured.)

| | |
|---|---|
| Cellulose acetate phthalate, NF | 12.0% |
| Propylene glycol | 3.0% |
| Sorbitan monooleate, NF | 1.0% |
| Ethyl alcohol 200-proof | 45.0% |
| Acetone | q.s. ad 100.0% |

This solution is prepared in a manner similar to example 2. The cellulose acetate phthalate is dissolved in solvent mixture, and acetone is added to obtain the proper volume after polymer solvation is obtained. Cellulose acetate phthalate was the primary synthetic enteric polymer used in the industry for many years. Now enteric acrylic resins and phthalate derivatives of polyvinyl acetate or hydroxypropyl methylcellulose are also available.

## Aqueous Film Coating Formulas

*Examples 4, 5, 6, and 7*: Cellulose Aqueous Formula

(Aqueous systems should be sprayed.)

| | 4 | 5 | 6 | 7 |
|---|---|---|---|---|
| HPMC, 15 cps | 4.0% | | 2.0% | 5.0% |
| HPMC, 6 cps | | 6.0% | 4.0% | |
| HPC | | | 1.0% | |
| PG | 1.0% | 1.0% | | |
| PEG 400 | | 0.5% | 2.0% | 1.0% |
| Water | q.s. 100.0% | 100.0% | 100.0% | 100.0% |

These polymers are soluble in water. Slowly add the polymer(s) to vigorously stirred water. Continue the agitation until the polymer(s) are solubilized, add propylene glycol or polyethylene glycol or both, then bring to proper volume with water. Colorants and pigments may be added after milling or dispersion in water.

## COATING FORMULA OPTIMIZATION

Optimization is usually associated with minor modifications in a basic formula. As discussed earlier, the basic or starting formula is obtained from past experience or from various sources in the literature. Modifications on this basic formula may be necessary to improve adhesion of the coating to the core; to decrease bridging of intagliations; to increase coating hardness; or to improve any property of the coating that the formulator deems deficient. Colorant and opaquant concentrations are usually fixed to achieve a predetermined shade. Changes of the polymer(s)-to-plasticizer ratio, however, or the addition of different plasticizers or polymers, are common modifications made in optimization of the coating.

This type of experimentation can be best achieved by conducting a fractional factorial type of study, in which the concentration of a few plasticizers or polymers are evaluated in the same general coating formulation. Factorial studies allow evaluation of more variables with fewer experiments. The evaluation of each coating composition, however, must be conducted by a readily quantifiable criterion. For example, if the coating compositions are to be applied to tablets, can the coating conditions be effectively repeated? Can the properties of the film be measured by an objective testing system? The conditions used in the coating process frequently have as great an effect on the quality of the tablet coating as the coating composition. Studies on free films are much easier to conduct because there are test methods that can be used to evaluate changes in film properties with modifications in coating composition. Bonding of a film to a tablet surface or bridging of an intagliation can be measured, but the experimental error is much higher. Optimization of a particular property in free films should always be confirmed by the performance and appearance of the coated tablet.

The literature and patents cite numerous film coating compositions. The selection of a specific formulation depends on the coating equipment and conditions available, the intended purpose of the coating, and total solid load desired in the coating.

## COATED TABLET EVALUATION/QUALITY CONTROL

After coating, the tablets should be inspected and tested for appearance and performance. Inspection should include checks for color (both hue and continuity), size, appearance, and any physical defects in the coating, which could affect the performance or stability of the product.

The drug release characteristics and longterm stability of coated products may be predicted by evaluating isolated free films for permeability and mechanical strength. Methods commonly used to prepare isolated films include casting and spraying techniques. Cast films can be prepared by spreading the coating composition on a teflon, glass, or aluminum foil surface using a spreading bar to get a uniform film thickness. Many cast films adhere so well to glass that the film cannot be removed intact, but glass is certainly suitable for evaluating the appearance of the film. Many investigators who conducted water vapor permeability studies prepared their films by pouring their coating composition on mercury in petri dishes. This is convenient, as the surface area is constant, and the film can be readily removed from the liquid surface.

Sprayed films can be obtained by mounting a plastic-coated surface in a spray hood or coating pan. Care must be used in spraying the film to obtain a uniform film representative of the type achieved in tablet coating.

The physical appearance of these films can provide evidence of potential colorant or opaquant separation. Lack of color uniformity within the film could suggest that the insoluble additives have not been properly suspended or that some interaction has occurred between the ingredients. In addition, the films can be submitted for the following tests.

## Water Vapor Permeability

If the coating is going to be used as a seal coat or to provide some physical protection for a tablet containing a water-unstable drug, then knowledge of the film's water vapor permeability should be assessed. Water vapor permeability is commonly evaluated using water vapor transmission cells sealed with polymeric free film (Fig. 13.37). A water vapor transmission cell consists of a glass cylinder containing a saturated salt solution to produce a specific internal vapor pressure. Polypropylene washers, rubber washers and aluminum caps are often used to prevent water vapor egress through any areas other than the film. These transmission cells are kept in a humidity controlled chamber and weighed periodically over a specified time period. These transmission cells rely on a vapor pressure gradient to achieve a linear weight gain. From the daily weight gains, the water vapor permeability constant can be calculated using following equation.

$$P_{erm} = \frac{WL}{A\Delta P} \qquad \ldots (2)$$

where, $P_{erm}$ is the moisture permeability constant, $W$ is the amount of moisture transmitted per unit time, $L$ is the thickness of the film, A is the area of the film exposed, and $\Delta P$ is the vapor pressure gradient.

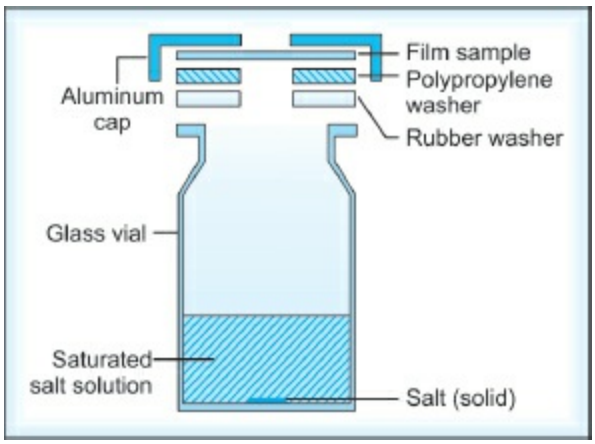



**Fig. 13.37:** Schematic of a water vapor transmission cell

## Film Tensile Strength

Mechanical properties of polymeric films can be evaluated by (1) microindentor probe analysis, (2) puncture and shear tests, and (3) stress relaxation. The stress-strain testing is the most widely used technique, in which strips of the film are tested on a tensile-strength tester by applying a known force (axial load) at a constant rate and measuring the load and deformation simultaneously. The stress-strain test will provide a generalized curve (Fig. 13.38) from which several useful properties can be determined. During the tensile test, initially there is a linear portion of the curve where the elongation of the polymeric film is directly proportional to the applied stress, and the slope of this linear region is known as *Young's modulus*. Young's modulus is a measure of the stiffness of the film or the ability of the film to withstand high stress while undergoing little elastic deformation. The greater the slope of the line, the higher the modulus and the greater the stiffness of the polymeric material. The plateau area where the film first undergoes a marked increase in strain without a corresponding increase in stress is known as *plastic deformation*. During plastic deformation, the structure of the polymeric film changes as the polymer chains orient themselves parallel to the direction of flow. As the film continues to be elongated, a point is reached at which the film breaks, known as the *elongation at break*. The stress value at this point is referred to as the *ultimate tensile strength*, which is a measure of the ability of the solid to withstand fracture. The area under the curve represents the work done and is referred to as the *toughness of the polymer*.

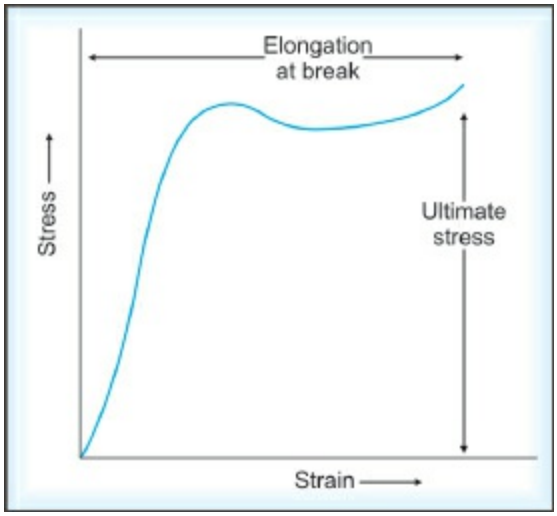



**Fig. 13.38:** Example of a stress-strain curve obtained in the tensile testing of

strips of free films

This test is particularly good when the effect of varying the concentration of a series of plasticizers or additives is being evaluated. Coating compositions that yield brittle films must be plasticized to obtain a more flexible film that is acceptable for tablet coating. Tensile-strength testing is one of the better ways to optimize the level of additives in the formulation.

Evaluation of the quality of coating on a tablet involves studying not only the film per se, but also the film-tablet surface interactions. A number of test methods can be employed.

1. *Adhesion tests* with tensile-strength testers have been used to measure the force required to peel the film from the tablet surface.Rowe has been a prolific investigator in the area of film coating evaluation and the factors affecting film strength.
2. *Diametral crushing strength* of coated tablets can be determined with a tablet hardness tester. Obviously, the resistance of the uncoated tablet to crushing will be a major factor in the test results. With this test, one is seeking information on the relative increase in crushing strength provided by the film and the contribution made by changes in the film composition.
3. The in vitro performance of the coated product is evaluated by *disintegration and dissolution testing*. Unless the coating is intended to control drug release, the coating should have a minimal effect on tablet disintegration or dissolution. A description of the standard equipment and methods used for each test is given in the USP/NF. Standards of acceptance for drug products can be found in the USP/NF or the Code of Federal Regulations, Title 21.
4. *Stability studies* must be conducted on coated tablets to determine if temperature and humidity changes will cause film defects. Exposure of coated tablets to elevated humidity and measurement of tablet weight gain provide relative information on the protection provided by the film. Stability testing is conducted to determine the effect of time and storage conditions on the physical and chemical stability of the coated product. The stability program should be designed to determine the shelf-life or expiry dating of the coated product under normal storage conditions in its intended package. As a rule, tablet products should have at least a two-year expiration date, which means that the product must conform to all

standards of performance and potency for at least two years after manufacture.

5. Some investigators have attempted to quantify film surface roughness, hardness, and color uniformity through instrumental means, but in general, visual inspection is sufficient to define relative coated tablet quality. A practical qualitative measure of the resistance of a coated tablet to abrasion can be obtained by merely rubbing the coated tablet on a white sheet of paper. Resilient films remain intact, and no color is transferred to the paper; very soft coatings are readily “erased” from the tablet surface to the paper.
6. Additional testing of coated tablets may also include *tests for resistance to chipping and cracking* during handling. Methods and devices for these tests are similar to those used for uncoated tablets.

## Film Defects

Variations in formulation and processing conditions may result in unacceptable quality defects in the film coating. The source of these defects and some of their probable causes are described in the following sections.

### *Sticking and Picking*

Overwetting or excessive film tackiness causes tablets to stick to each other or to the coating pan. On drying, at the point of contact, a piece of the film may remain adhered to the pan or to another tablet, giving a “picked” appearance to the tablet surface and resulting in a small exposed area of the core. A reduction in the liquid application rate or increases in the drying air temperature and air volume usually solve this problem. Excessive tackiness may be an indication of a poor formulation.

### *Roughness*

A rough or gritty surface is a defect often observed when the coating is applied by a spray. Some of the droplets may dry too rapidly before reaching the tablet bed, resulting in deposits on the tablet surface of “spray-dried” particles instead of finely divided droplets of coating solution. Moving the nozzle closer to the tablet bed or reducing the degree of atomization can decrease the roughness due to “spray-drying.” Surface roughness also increases with pigment concentration and polymer concentration in the coating solution.

### *Orange-Peel Effects*

Inadequate spreading of the coating solution before drying causes a bumpy or “orange-peel” effect on the coating. This indicates that spreading is impeded by too rapid drying or by high solution viscosity. Thinning the solution with additional solvent may correct this problem.

### *Bridging and Filling*

During drying, the film may shrink and pull away from the sharp corners of an intagliation or bisect, resulting in a “bridging” of the surface depression. This defect can be so severe that the monogram or bisect is completely obscured. This mainly represents a problem with the formulation. Increasing

the plasticizer content or changing the plasticizer can decrease the incidence of bridging. Filling is caused by applying too much solution, resulting in a thick film that fills and narrows the monogram or bisect. In addition, if the solution is applied too fast, overwetting may cause the liquid to quickly fill and be retained in the monogram. Judicious monitoring of the fluid application rate and thorough mixing of the tablets in the pan prevent filling.

## *Blistering*

When coated tablets require further drying in ovens, too rapid evaporation of the solvent from the core and the effect of high temperature on the strength, elasticity, and adhesion of the film may result in blistering. Milder drying conditions are warranted in this case.

## *Hazing/Duii Film*

This is sometimes called *bloom*. It can occur when too high a processing temperature is used for a particular formulation. Dulling is particularly evident when cellulosic polymers are applied out of aqueous media at high processing temperatures. It can also occur if the coated tablets are exposed to high humidity conditions and partial solvation of film results.

## *Color Variation*

This problem can be caused by processing conditions or the formulation. Improper mixing, uneven spray pattern, and insufficient coating may result in color variation. The migration of soluble dyes, plasticizers, and other additives during drying may give the coating a mottled or spotted appearance. The use of lake dyes eliminates dye migration. A reformulation with different plasticizers and additives is the best way to solve film instabilities caused by the ingredients.

## *Cracking*

Cracking occurs if internal stresses in the film exceed the tensile strength of the film. The tensile strength of the film can be increased by using higher-molecular-weight polymers or polymer blends. Internal stresses in the film can be minimized by adjusting the plasticizer type and concentration, and the pigment type and concentration.

## *Twinning*

Twinning is a form of overwetting whereby two or more of the tablet cores are stuck together. This situation is prevalent when attempting to coat flat-faced or caplet-shaped tablets and may arise due to any of the possible causes for overwetting. Placing even a very subtle amount of curvature on an otherwise flat surface can minimize twinning problems.

## SPECIALIZED COATINGS

### Compression Coating

This type of coating requires a specialized tablet machine. The finished product is a tablet within a tablet, as described in Chapter 11. Compression coating is not widely used, but it has advantages in some cases in which the tablet core cannot tolerate organic solvents or water and yet needs to be coated for taste masking, or to provide delayed or enteric properties to the product. In addition, incompatible ingredients can be conveniently separated by this process.

## Electrostatic Coating

Electrostatic coating is an efficient method of applying coating to conductive substrates. A strong electrostatic charge is applied to the substrate. The coating material containing conductive ionic species of opposite charge is sprayed onto the charged substrate. Complete and uniform coating of corners and intagliations on the substrate is achieved. The adaptability of this method to such relatively nonconductive substrates as pharmaceutical tablets is limited, although one process has been proposed.

## Dip Coating

Coating is applied to the tablet cores by dipping them into the coating liquid. The wet tablets are dried in a conventional manner in coating pans. Alternate dipping and drying steps may be repeated several times to obtain the desired coating. This process lacks the speed, versatility, and reliability of spraycoating techniques. Specialized equipment has been developed to dip-coat tablets, but no commercial pharmaceutical application has been obtained.

## Vacuum Film Coating

Vacuum film coating is a new coating procedure that employs a specially designed baffled pan. The pan is hot water jacketed, and it can be sealed to achieve a vacuum system. The tablets are placed in the sealed pan, and the air in the pan is displaced by nitrogen before the desired vacuum level is obtained. The coating solution is then applied with an airless (hydraulic) spray system. The evaporation is caused by the heated pan, and the vapors are removed by the vacuum system. Because there is no high-velocity heated air, the energy requirements are low and coating efficiency is high. Organic solvents can be effectively used with this coating system with minimal environmental or safety concerns.

## FUTURE TABLET DEVELOPMENTS

Much progress has been achieved in tablet coating through significant equipment improvements. The expanded use of microprocessors for process control and improved spraying systems are likely.

The advent of the pseudo-latex coating systems has shown that high solids containing coating formulations are attainable. Most film coating formulations still are primarily volatile solvents. Reducing the quantity of solvent that must be removed in the coating process would dramatically improve coating efficiency.

With the exception of the polymers used for enteric coating, most polymers available for pharmaceuticals were primarily developed for nonpharmaceutical use. There is a need for polymers developed specifically to meet the pharmaceutical coating requirements. Polymers that can be applied in high concentration from aqueous-based solvents would be desirable. Their films should be flexible, adhere well to all tablet surfaces, and be nontacky.

# PART III: TABLET PUNCH FORCES

## TABLET FORMATION PROCESS

The mechanical criteria for a successful tablet formulation are good flowability for powders and adequate strength without fracture for compacts. The quality of a compressed tablet is determined by powder characteristics and compression behavior. Powder characteristics depend upon flow properties that are primarily affected by particle size and shape. Additionally, high interparticle friction can have a detrimental effect on powder characteristics due to bridging and non-uniform flow. A non-uniform particle size distribution may also lead to material segregation resulting in uniformity problems. The tableting characteristics of powders depend on the viscoelastic properties of the material. The process of compaction has been defined as "the compression and consolidation of a two-phase system due to an applied load." The quality of the compact depends on the compression and consolidation of the powder mass, decompression of the compact, ejection from the die, and subsequent scrape-off from the lower punch. A schematic representation of the compression process is shown in Fig. 13.39. Tablet quality is affected by both the magnitude and the rate of application of the compression force since viscoelastic properties are time dependent.

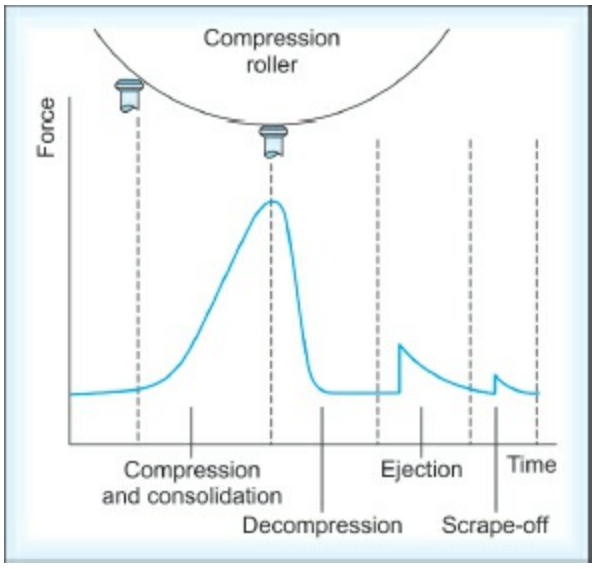



**Fig. 13.39:** The schematic representation of the compression process

## Compression

When external mechanical forces are applied to a powder mass, there is normally a reduction in its bulk volume as a result of one or more of the following effects. The onset of loading is usually accompanied by closer repacking of the powder particles, and in most cases, this is the main mechanism of initial volume reduction, as shown diagrammatically in Fig. 13.40.

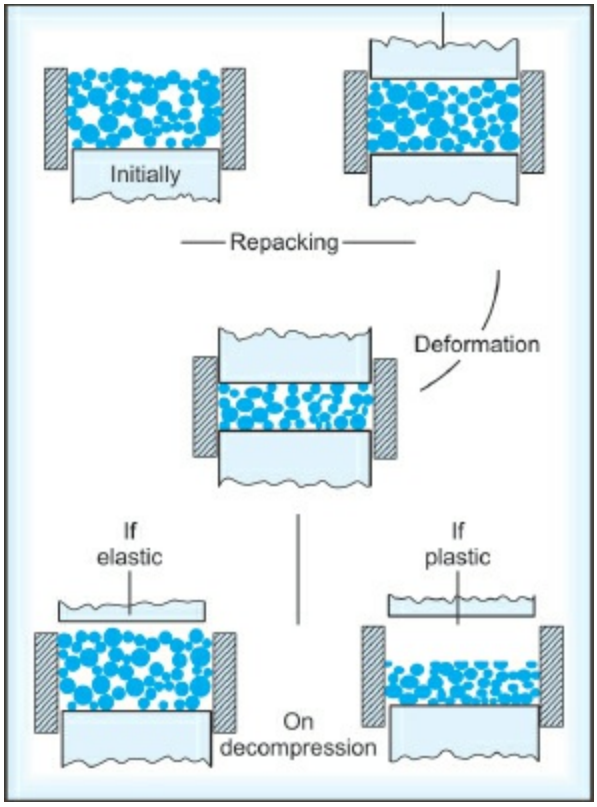



**Fig. 13.40:** Diagram of the effect of compressional force on a bed of powder

As the load increases, however, rearrangement becomes more difficult, and further compression involves some type of particle deformation as mentioned here under:

### *Elastic Deformation*

If on removal of the load, the deformation is to a large extent spontaneously reversible, i.e. if it behaves like rubber, then the deformation is said to be *elastic*. All solids undergo some elastic deformation when subjected to

external forces. With several pharmaceutical materials, such as acetylsalicyclic acid and microcrystalline cellulose, elastic deformation becomes the dominant mechanism of compression within the range of maximum forces normally encountered in practice.

## *Plastic Deformation*

In other groups of powdered solids, an elastic limit, or *yield point*, is reached, and loads above this level result in deformation not immediately reversible on removal of the applied force. Bulk volume reduction in these cases results from plastic deformation and/or viscous flow of the particles, which are squeezed into the remaining void spaces, resembling the behavior of modeling clay. This mechanism predominates in materials in which the shear strength is less than the tensile or breaking strength (Fig. 13.40).

## *Brittle Fracture*

Conversely, when the shear strength is greater, particles may be preferentially fractured, and the smaller fragments then help to fill up any adjacent air space. This is most likely to occur with hard, brittle particles and in fact is known as brittle fracture; sucrose behaves in this manner. The predisposition of a material to deform in a particular manner depends on the lattice structure, in particular whether weakly bonded lattice planes are inherently present.

## *Microsquashing*

Irrespective of the behavior of large particles of the material, small particles may deform plastically, a process known as microsquashing, and the proportion of fine powder in a sample may therefore be significant. Asperities that are sheared off larger, highly irregular particles could also behave in this way, thus, particle shape is another important factor.

The above account describes all of the possible mechanisms that can contribute to a reduction in the bulk volume of a bed of powder, when subjected to external mechanical forces. The chemicophysical characteristics of the material being studied determine the contribution each effect makes as the compressional load is increased. All of the deformation effects may be accompanied by the breaking and formation of new bonds between the particles, which gives rise to consolidation as the new surfaces are pressed together.

The packaging of bulk powders and the filling of hard gelatin capsules mostly involve bulk volume reductions achievable by repacking, and possibly a minimal amount of deformation. At the other end of the scale, in the tabletting process—or in such specialized techniques as roll compacting or extruding, which involve high levels of compressive force—repacking, elastic deformation, plastic deformation, and brittle fracture may all take place.

Some deformation processes (plastic deformation, for example) are time-dependent and occur at various rates during the compaction sequence, so that the tablet mass is never in a state of stress/strain equilibrium during the actual tabletting event. This means that the rate at which load is applied and removed may be a critical factor in materials for which dependence on time is significant. More specifically, if a plastically deforming solid is loaded (or unloaded) too rapidly for this process to take place, the solid may exhibit brittle fracture. This is a contributing factor to the well-known problem of structurally failed tablets of some drugs as tablet machine speed is raised. Conversely, if the dwell time under the compressive load is prolonged, then plastic deformation may continue, leading to more consolidation. This phenomenon has recently been studied using a compaction simulator, whereby it was shown that the expansion of acetaminophen tablets (a material with known laminating tendency) during decompression was particularly sensitive to dwell time under a maximum load. For this reason, relatively slower machine speeds and compression rolls of large diameter sometimes help with troublesome tablet formulations.

## Consolidation

When the surfaces of two particles approach each other closely enough (e.g. at a separation of less than 50 nm), their free surface energies result in a strong attractive force, a process known as *cold welding*. The nature of the bonds so formed are similar to those of the molecular structure of the interior of the particles, but because of the roughness of the particle surface (on a molecular scale), the actual surface area involved may be small. This hypothesis is favored as a major reason for the increasing mechanical strength of a bed of powder when subjected to rising compressive forces.

On the macroscale, most particles encountered in practice have an irregular shape, so that there are many points of contact in a bed of powder. Any applied load to the bed must be transmitted through these particle contacts; under appreciable forces, this transmission may result in the generation of considerable frictional heat. If this heat is not dissipated, the local rise in temperature could be sufficient to cause melting of the contact area of the particles, which would relieve the stress in that particular region. In that case, the melt solidifies, giving rise to *fusion bonding*, which in turn results in an increase in mechanical strength of the mass.

Many pharmaceutical solids possess a low specific heat and poor thermal conductivity, so that heat transfer away from the contact points is slow. This behavior was quantified by Rankell and Higuchi, who were able to estimate, from heat transfer kinetics, that temperatures high enough to fuse typical organic medicinal substances were theoretically possible. The differences between this form of bond formation and cold welding are somewhat pedantic; the end results are essentially the same.

In both “cold” and “fusion” welding, the process is influenced by several factors, including:

1. The chemical nature of the materials
2. The extent of the available surface
3. The presence of surface contaminants
4. The intersurface distances

The type and degree of crystallinity in a particular material influences its consolidative behavior under appreciable applied force. One of the earliest reports of such influence was that of Jaffe and Foss, who demonstrated that

substances possessing the cubic lattice arrangement were tabletted more satisfactorily than those with a rhombohedral lattice, for example. The isotropic nature of the former group might be expected to contribute to better tabletting because no alignment of particular lattice planes is required. In addition they provide three equal planes for stress relief at right angles to each other. Lattice planes with the greatest separation undergo plastic deformation more readily, since such planes are more weakly bonded. The particles of most pharmaceutical powders consist of small crystallites, or grains, aggregated in a random manner so that their crystal planes are not aligned with one another. Such an arrangement adds to the material's resistance to plastic deformation.

Of course, there are many exceptions to generalizations of this kind. For example, of the two chemically similar organic materials methacetin and phenacetin, apparently only the former can be tabletted without the tendency to laminate. More importantly, perhaps, different polymorphic forms and crystal habits of the same compound may not behave in the same way in terms of compaction properties. Detection and evaluation of these seemingly unavoidable changes in the materials from bulk chemical plants, which are responsible for many of our unanticipated tabletting problems with established products, are important. Routine testing of compaction characteristics in some type of instrumented tabletting machine constitutes a desirable and informative part of such procedures.

One interesting observation on the more successful direct compression excipients that are commercially available concerns this "material structure" aspect of tabletting. Without exception, these products may be described as microgranulations, since they consist of masses of small crystallites randomly embedded in a matrix of some gluelike (often amorphous) material. Such a combination imparts the desired overall qualities, which result in (1) strong tablets by providing a plastically deforming component (the matrix) to relieve internal stresses and (2) strongly bonding surfaces (the faces of the crystallites) to enhance consolidation.

The compressional process is affected by the extent of the available surface, the presence of surface contaminants, and the intersurface distances; if large, clean surfaces are brought into intimate contact, then bonding should occur. Brittle fracture and plastic deformation should generate clean surfaces, which the compressional force ensures are kept in close proximity. This is an

important rationale when considering the tabletting process, since many of the problems that arise can be traced to interference in these mechanisms. For example, such lubricants as magnesium stearate form weak bonds, so that overlubrication, or even overmixing of lubricant into the tabletting mass, results in a continuous coating of the latter, and hence in some cases, weak tablets.

Higuchi and his co-workers were among the first to report experimental data in support of these mechanisms for a pharmaceutical material. They interpreted the plot of specific surface area versus compressional force shown in Fig. 13.41 in terms of an initial increase in surface area (region O to A) due to

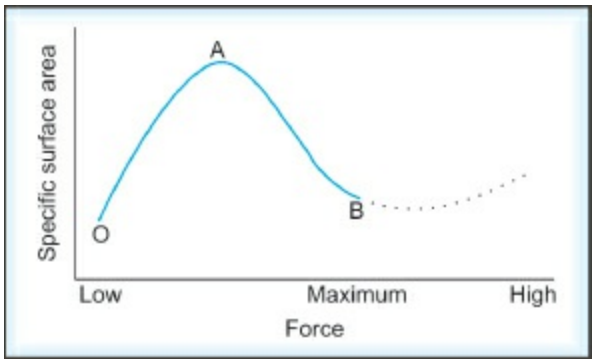



**Fig. 13.41:** The effect of increasing compressional force on the specific surface area of a powder mass. When a powder mass is subjected to increasing compressional force, there is initial particle fracture, which gives rise to increased surface area (O to A on graph). At some point (A), particle rebonding becomes the dominant factor, and from then on, surface area decreases (region A to B) unless tablet lamination begins particle breakdown. At high loads (A to B), rebonding of surfaces became dominant, with a resultant decrease in specific surface areas.

Years later, Armstrong and co-workers described similar curves, but at high compressional forces (*see* broken line in Fig. 13.41), they showed that some materials were subject to an increase in surface area. This increase was due to incipient failure or lamination of the tablet structure that resulted from considerable elastic recovery on decompression.

The actual solubility of solids also depends somewhat on the applied pressure, so that if a film of moisture is present on the solid surface, then the

high pressures at points of solid contact could force more material into solution. This dissolved solid would crystallize on relief of the applied stress to form a solid bridge whose strength would partly depend on the rate of this recrystallization. In general, slow rates should produce a more perfect crystal structure with consequent higher strength.

The observation that tablets containing at least a proportion of water-soluble component are often more readily formed than those without, may indicate this mechanism. The finding that some overdried mixtures, in which moisture residues are extremely low, have inferior tabletting qualities may be further evidence. The known internal lubricant property of water, however, provides an alternative, and probably complementary, explanation of the important role of moisture.

At low levels of external force, molecular and electrostatic forces are a source of attractive tendencies between individual particles. This attraction might be encountered in the mixing of dry powders or in the filling of capsule shells. Valency type molecular forces have an extremely short range and are therefore unlikely to play a major role at this stage. Van der Waals forces, however, may exert a significant effect at distances up to 100 nm, so that once an agglomerate of particles has been formed, they may serve to prevent its breakdown. Electrostatic forces generated by friction or size reduction, though generally weaker than Van der Waals forces, do have a greater range and probably produce the initial agglomerate formation in many materials. Ionic chemicals involve the additional possibility of surface polarization, which can produce marked attractive tendencies.

If the solid particles are soft, then deformation under low loads could cause more intimate contact between the particles and enhance the above bonding mechanisms. Recently, however, it has been shown that coating of a major tablet ingredient by a lower melting point component (such as a lubricant) can sometimes lead to reduced tensile strength of tablets, especially if produced at slightly elevated temperatures (10°C). It is thought that this effect is due to masking of Van der Waals forces between the particles and the formation of welded bonds by the coatings.

Some capsule filling machines operate on the “dosator principle,” that is, formation of a soft plug, which is transferred to the capsule shell. In such machines, the plugs are commonly held together by one or more of the foregoing mechanisms.

Augsburger and associates described a simple test to measure the comparative strength of different formulations in such powder plugs; the apparatus is illustrated diagrammatically in Fig. 13.42. Hiestand has reviewed in detail all the various mechanisms and their applicability to pharmaceutical powder systems, concluding that in tabletting, plastic deformation is the major mechanism leading to increased areas of intimate contact, and hence bonding, by cold welding.

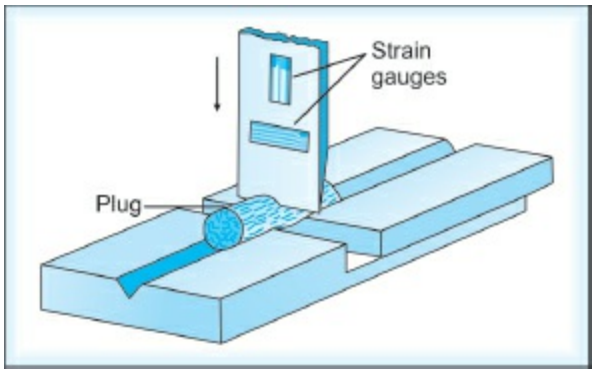



**Fig. 13.42:** Plug strengh tester

## *Role of Moisture*

At least some moisture is present in virtually all capsule and tablet formulas, and concentrations well below the 1% level can dramatically affect the behavior of these feed materials and that of the finished product. This is demonstrated clearly by data such as that given in Fig. 13.43, which shows that as little as 0.02% moisture can affect the proportion of the applied force transmitted to the lower punch, and at 0.55% moisture, the behavior is actually the reverse of that for the totally dry material. A more critical factor concerns the situation where the amount of moisture present on the powder surfaces is just sufficient to fill the remaining voids in the bed. Any further reduction in porosity, e.g. as a result of increasing compressional force, results in this water being squeezed out to the surface of the tablet. This expelled moisture may act as a lubricant at the die wall, but it could also cause material to stick to the punch faces.

Recent experiments have shown the effect of thermal dehydration on the crushing strength of tablets made from certain hydrates. The strength was found to depend on the temperature at which the dehydration was carried out. Scanning electron microscopy confirmed that dehydration had been

accompanied by a change in texture of the crystals, which led to a more porous mass. Moisture is also important in moist granulation processes, in which most of the fluids involved are aqueous in nature. Therefore, before the bonding produced in tablets is discussed in more detail, the special consolidative process that is involved in the granulation of powders by the addition of a granulating liquid is considered.

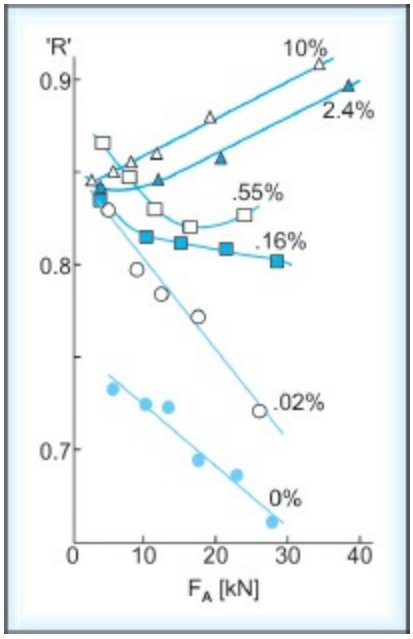



**Fig. 13.43:** Effect of moisture content on the ratio of transmitted punch force to applied punch force

## *Compression and Consolidation under High Loads*

The processes of tabletting, roll compaction, and extrusion all involve the application of massive compressive forces, which induce considerable deformation in the solid particles. With many pharmaceutical solids and perhaps most tabletting mixtures, these forces are large enough to exceed the elastic limit of the solid (or at least that of one component of the mixture). Plastic deformation and/or brittle fracture then results in the generation of new, clean surfaces, which being pressed against one another, undergo cold welding. When the compaction force reaches its maximum, a bulk solid structure of a certain overall strength will have been produced.

Irrespective of the bonding mechanism, this structure must be strong enough to withstand the new stresses induced during release of the applied

load and those generated by ejection from the die (in the case of tabletting). Ideally, relief of these stresses by elastic recovery is preferred, since at this stage, plastic deformation—and even worse, brittle fracture—may result in small failure planes, if not complete lamination, since any new surfaces tend to separate rather than consolidate.

During normal tabletting operations, consolidation is accentuated in those regions adjacent to the die wall, owing to the intense shear to which material is subjected, as it is compressed axially and is pushed along the wall surface. This consolidation results in a “skin” of material that is denser over the lateral tablet surface than in the rest of the tablet mass. This skin is in some cases visible to the naked eye. Although this thin layer of material may contribute to abrasion resistance, it may retard the escape of air during compression and the ingress of liquid during dissolution, both undesirable features. For these reasons, smaller tablet height-to-diameter ratios are preferred. This situation is advantageous from additional standpoints, which are now to be considered.

The resistance to differential movement of particles caused by their inherent cohesiveness results in the applied force not being transmitted uniformly throughout the entire mass. More specifically, in the case of a singlestation press, the force exerted by the upper punch diminishes exponentially at increasing depths below it. Thus, the relationship between upper punch force $F_A$ and lower punch force $F_L$* may be expressed in the form:

$$F_L = F_A \cdot e^{-KH/D} \quad \text{... (3)}$$

where, $K$ is an experimentally determined material-dependent constant that includes a term for the average die-wall frictional component. The values $H$ and $D$ are the height and diameter of the tablet, respectively.

The discrepancy between the two punch forces should be minimized in pharmaceutical tabletting operations, so that there is no significant difference in the amount of compression and consolidation between one region of the tablet and another. Reduction of die-wall friction effects by having smaller tablet height-to-diameter ratios and by adding a lubricant is therefore common practice. Because of their important role in the progress of the compressional sequence, frictional effects warrant further discussion.

## *Effects of Friction*

At least two major components to the frictional forces can be distinguished.

1. *Interparticulate friction:* This arises at particle/particle contacts and can be expressed in terms of a coefficient of interparticulate friction $\mu_1$; it is more significant at low applied loads. Materials that reduce this effect are referred to as *glidants*. Colloidal silica is a common example.
2. *Die-wall friction:* This results from material being pressed against the die wall and moved down it; it is expressed as $\mu_w$, the coefficient of die-wall friction. This effect becomes dominant at high applied forces when particle rearrangement has ceased and is particularly important in tabletting operations. Most tablets contain a small amount of an additive designed to reduce die-wall friction; such additives are called *lubricants*. Magnesium stearate is a common choice.

## *Force Distribution*

Most investigations of the fundamentals of tabletting have been carried out on singlestation presses (sometimes called *eccentric presses*), or even on isolated punch and die sets in conjunction with a hydraulic press. The system represented diagrammatically in Fig. 13.44 is typical of such arrangements, with force being applied to the top of a cylindric powder mass. This simple compaction system provides a convenient way to examine the process in greater detail. More specifically, the following basic relationships apply. Since there must be an axial (vertical) balance of forces:

$$F_A = F_L + F_D \quad \ldots (4)$$

where, $F_A$ is the force applied to the upper punch, $F_L$ is that proportion of it transmitted to the lower punch, and $F_D$ is a reaction at the die-wall due to friction at this surface. Because of this inherent difference between the force applied at the upper punch and that affecting material close to the lower punch, a *mean compaction force*, $F_M$, has been proposed, where:

$$F_M = \frac{F_A + F_L}{2} \quad \ldots (5)$$

A recent report confirms that $F_M$ offers a practical friction-independent measure of compaction load, which is generally more relevant than $F_A$. In single-station presses, where the applied force transmission decays

exponentially. A more appropriate *geometric mean force* ($F_G$) might be:

$$F_G = (F_A \times F_L)^{½} \quad \text{... (6)}$$

Use of these force parameters are probably more appropriate than use of $F_A$ when determining relationships between compressional force and such tablet properties as tablet strength.

## *Development of Radial Force*

As the compressional force is increased and any repacking of the tabletting mass is completed, the material may be regarded to some extent as a single solid body. Then, as with all other solids, compressive force applied in one direction (e.g. vertical) results in a decrease $\Delta H$ in the height. In the case of an unconfined solid body, this would be accompanied by an expansion in the horizontal direction of $\Delta D$. The ratio of these two dimensional changes is known as the Poisson ratio ($\lambda$) of the material, defined as:

$$\lambda = \frac{\Delta D}{\Delta H} \quad \text{... (7)}$$

The Poisson ratio is a characteristic constant for each solid and may influence the tabletting process in the following way. Under the conditions illustrated in Fig. 13.44, the material is not free to expand in the horizontal plane because it is confined in the die. Consequently, a radial die-wall force $F_R$ develops perpendicular to the die-wall surface, materials with larger Poisson ratios giving rise to higher values of $F_R$. Classic friction theory can then be applied to deduce that the axial frictional force $F_D$ is related to $F_R$ by the expression:

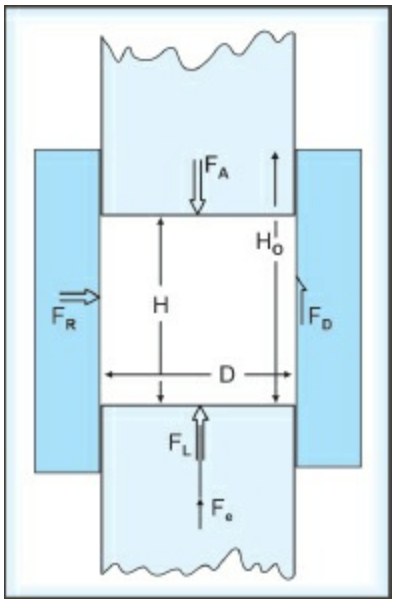



**Fig. 13.44:** Diagram of a cross-section of a typical simple punch and die assembly used for compaction studies

$$F_D = \mu_w \times F_R \ldots (8)$$

where, $\mu_w$ is the coefficient of die-wall friction. Note that $F_R$ is reduced when materials of small Poisson ratios are used, and that in such cases, axial force transmission is optimum.

The frictional effect represented by $\mu_w$ arises from the shearing of adhesions that occurs as the particles slide along the die wall. It follows that its magnitude is related to the shear strength S of the particles (or the die-wall-particle adhesions if these are weaker) and the total effective area of contact $A_e$ between the two surfaces. Therefore, force transmission is also realized when $F_D$ values are reduced to a minimum, which is achieved by ensuring adequate lubrication at the die wall (lower S) and maintaining a minimum tablet height (reducing $A_e$).

A common method of comparing degrees of lubrication has been to measure the applied and transmitted axial forces and determine the ratio $F_L/F_A$. This is called the *coefficient of lubricant efficiency*, or *R value*. The ratio approaches unity for perfect lubrication (no wall friction), and in practice, values as high as 0.98 may be realized. Values below 0.8 probably indicate a poorly lubricated system. Values of R should be considered as

relating only to the specific system from which they were obtained, because they are affected by other variables, such as compressional force and tablet *H/D* ratio.

## *Die-wall Lubrication*

Most pharmaceutical tablet formulations require the addition of a lubricant to reduce friction at the die wall. Die-wall lubricants function by interposing a film of low shear strength at the interface between the tabletting mass and the die wall, as illustrated in Fig. 13.45. Preferably, there is some chemical bonding between this "boundary" lubricant and the surface of the die wall as well as at the edge of the tablet. The best lubricants are those with low shear strength but strong cohesive tendencies in directions at right angles to the plane of shear. Table 13.11 gives the shear strength of some commonly used lubricants as measured by a punch penetration test. By utilizing materials with low shear strength as lubricants, shear failure occurs in the lubricant layers and not at the compressed powder or resultant wall interfaces (Fig. 13.45).

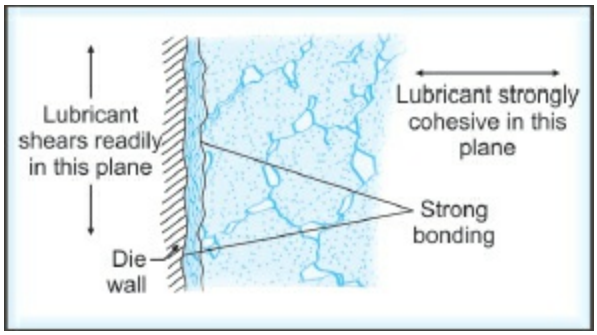



**Fig. 13.45:** Diagram illustrates the preferred characteristics of die wall lubricants

**Table 13.11:** The shear strength of some lubricants

| Material | Shear strength (MPa) | Material | Shear strength (MPa) |
|---|---|---|---|
| Stearic acid | 1.32 | Sodium stearate | 3.32 |
| Calcium stearate | 1.47 | Talc with grain | 6.20 |
| Hard paraffin | 1.86 | Talc across grain | 7.85 |
| Magnesium stearate | 1.96 | Boric acid | 7.16 |
| Potassium stearate | 3.07 | Graphite | 7.35 |

## *Ejection Forces*

Radial die-wall forces and die-wall friction also affect the ease with which the compressed tablet can be removed from the die. The force necessary to eject a finished tablet follows a distinctive pattern of three stages. The first stage involves the distinctive peak force required to initiate ejection, by breaking of tablet/die-wall adhesions. A smaller force usually follows, namely that required to push the tablet up the die wall. The final stage is marked by a declining force of ejection as the tablet emerges from the die. Variations on this pattern are sometimes found, especially when lubrication is inadequate and/or “slip-stick” conditions occur between the tablet and the die wall, owing to continuing formation and breakage of tablet/die-wall adhesions. Worn dies, which cause the bore to become barrel-shaped, give rise to a similar abnormal ejection force trace and may lead to failure of the tablet structure.

A direct connection is to be expected between die wall frictional forces and the force required to eject the tablet from the die, $F_E$. For example, well-lubricated systems (as indicated by a large R value) have been shown to lead to smaller $F_E$ values.

## FORCE-VOLUME RELATIONSHIPS

The end of the compressional process may be recognized as being the point at which all air spaces have been eliminated, i.e. $v_b = v_t$ and therefore $E = 0$. A small residual porosity is desirable, however, so there is particular interest in the relationship between applied force $F_A$ and remaining porosity E. Originally, it was suggested that decreasing porosity resulted from a two-step process: (1) the filling of large spaces by interparticulate slippage and (2) the filling of small voids by deformation or fragmentation at higher loads.

This process can be expressed mathematically:

$$\frac{E_0 - E}{E_0 \cdot (1 - E)} = K_1 e^{\frac{-K_2}{P}} + K_3 e^{\frac{-K_4}{P}} \qquad \ldots (9)$$

where, $E_0$ is the initial porosity, $E$ is the porosity at pressure $P$, and $K_1$, $K_2$, $K_3$ and $K_4$ are constants. The two terms on the right side of the equation refer to steps (1) and (2) respectively. Although equation (9) so far has only been shown to fit data from a few materials (such as alumina and magnesia), it does establish that the degree of compression achieved for a given load depends upon the initial porosity ($E_0$).

Therefore, the common practice of comparing different formulations by means of testing tablets of the same *weight* is undesirable. One variable is eliminated if experiments are carried out on tablet masses of the same true volume, and allowance should be made for varying initial values of bulk volume ($V_b$) when interpreting the results.

A more complex sequence of events during compression involves four stages, as illustrated by the data in Fig. 13.46. (Stage i) represents the initial repacking of the particles, followed by elastic deformation (stage ii) until the elastic limit is reached. Plastic deformation and/or brittle fracture then dominates (stage iii) until all voids are virtually eliminated. At this point, the onset of stage iv, compression of the solid crystal lattice, occurs.

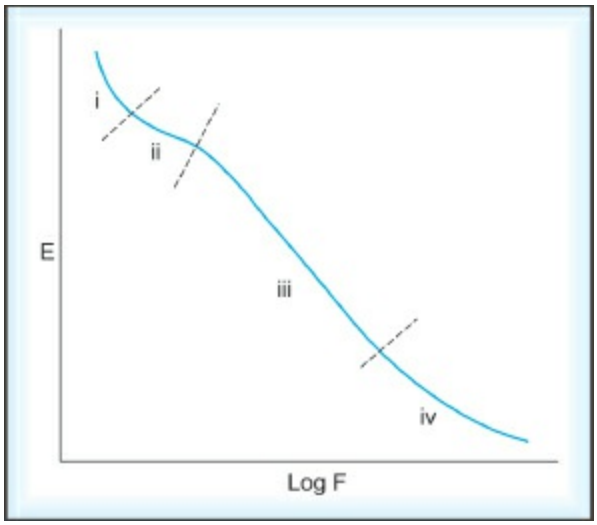



**Fig. 13.46:** Decreasing porosity with increasing compressional force for single-ended pressing

Attempts that have been made to derive equations for the first three stages, are of limited value, because in practice, the stages are not totally sequential. Owing to transmitted force variation, they may occur simultaneously in different regions of the same tablet.

In many tabletting processes, however, once appreciable force has been applied, the relationship between applied pressure (P) and some volume parameter such as porosity (E) does become linear over the range of pressure commonly used in tabletting (region iii in Fig. 13.46). For example, an equation first suggested by Shapiro has been shown to fit data obtained from several pharmaceutical materials:

$$\text{Log } E = \text{Log } E_0 - K.P. \quad \ldots (10)$$

where, $E_0$ is the porosity when the pressure is zero, and $K$ is a constant. Another equation for which there is considerable evidence, is attributed to Walker:

$$\frac{1}{1-E} = K_1 - K_2 \cdot \text{Log P} \qquad \ldots (11)$$

## *Heckel Plots*

The foregoing equations have been criticized because some of the constants apparently lack physical significance. Another equation, credited to Heckel, is free from this empiricism, however. The Heckel equation is based upon

analogous behavior to a first-order reaction, where the pores in the mass are the reactant, that is:

$$\text{Log}\frac{1}{E} = K_y P + K_r \qquad \ldots (12)$$

where, $K_y$ is a material-dependent constant inversely proportional to its yield strength S ($K_y$ = 1/3S), and $K_r$ is related to the inital repacking stage, and hence $E_0$. The above relationships may be established by simply measuring the applied compressional force $F$ and the movements of the punches during a compression cycle and translating this data into values of $P$ (applied pressure) and $E$ (porosity). For a cylindric tablet, $P$ is given by:

$$P = \frac{4F}{\pi \cdot D^2} \qquad \ldots (13)$$

where, $D$ is the tablet diameter. Similarly, values of $E$ can be calculated for any stage from:

$$E = 100 \cdot \left[1 - \frac{4w}{\rho_t \cdot \pi \cdot D^2 \cdot H}\right] \qquad \ldots (14)$$

where, $w$ is the weight of the tabletting mass, $\rho_t$ is its true density, and $H$ is the thickness of the tablet at that point (obtained from the relative punch displacement measurements). (*See* previous section, "Mass-Volume Relationships").

The particular value of Heckel plots arises from their ability to identify the predominant form of deformation in a given sample. Materials that are comparatively soft and that readily undergo plastic deformation retain different degrees of porosity, depending upon the initial packing in the die. This in turn is influenced by the size distribution, shape, etc. of the original particles. Heckel plots for such materials are shown by type '*a*' in Fig. 13.47; sodium chloride is a typical example.

Conversely, harder materials with higher yield pressure values usually undergo compression by fragmentation first, to provide a denser packing. Label '*b*' in Fig. 13.47 shows Heckel plots for different size fractions of the same material that are typical of this behavior. Lactose is one such material.

Type '*a*' Heckel plots usually exhibit a higher final slope ($K_y$) than type '*b*', which implies that the former materials have a lower yield stress. Hard,

brittle materials are, in general, more difficult to compress than soft, yielding ones because fragmentation with subsequent percolation of fragments is less efficient than void filling by plastic deformation. In fact, as the porosity approaches zero, plastic deformation may be the predominant mechanism for all materials.

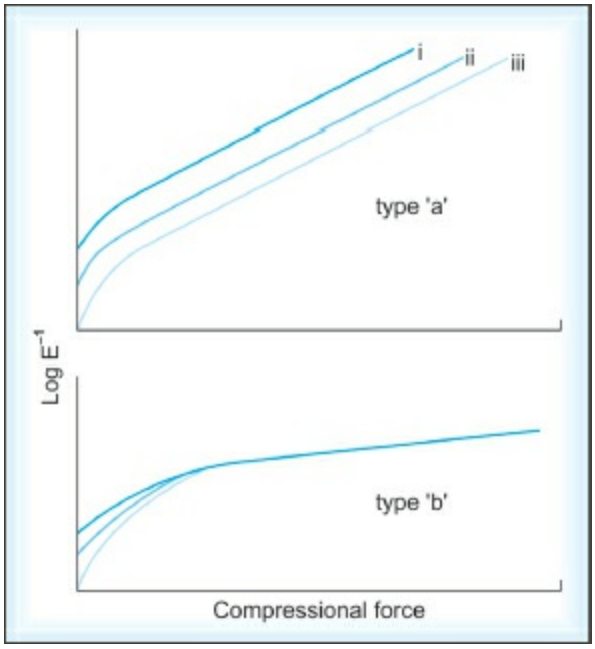



**Fig. 13.47:** Examples of Heckel plots. Curves i, ii and iii represent decreasing particle size fractions of the same material. Type ‘a’ curves are typical of plastically deforming materials, while those in which fragmentation occurs initially tend to show type ‘b’ behavior

The two regions of the Heckel plot are thought to represent the initial repacking stage and the subsequent deformation process, the point of intersection corres-ponding to the lowest force at which a coherent tablet is formed. In addition, the crushing strength of tablets can be correlated with the value of $K_y$ of the Heckel plot; larger values of $K_y$ usually indicate harder tablets. Such information can be used as a means of binder selection when designing tablet formulations. Note that Heckel plots can be influenced by the overall time of compression, the degree of lubrication, and even the size of the die, so that the effect of these variables can also be studied.

Another important factor in the use of all force-porosity relationships is that for many formulations, there is a relatively narrow optimum residual

porosity range that provides adequate mechanical strength, rapid water uptake, and hence, good disintegration characteristics. It is to the formulators' advantage to identify this optimum range and be able to predict compressing conditions needed to reach it. In addition to the predictive capability, establishing behavioral patterns for a given formulation (so-called "fingerprinting") may provide valuable diagnostic information in the event that a particular batch of the product causes problems.

Note also that the initial porosity can affect the course of the entire compressional sequence, and that in general, slow force application leads to a low porosity for a given applied load.

## Decompression

In operations such as tabletting, the compressional process is followed by a decompression stage, as the applied force is removed. This leads to a new set of stresses within the tablet as a result of elastic recovery, which is augmented by the forces necessary to eject the tablet from the die. Irrespective of the consolidation mechanism, the tablets must be mechanically strong enough to accommodate these stresses; otherwise, structural failure will occur.

For this reason, studies in which data are collected during both parts of the cycle have proved valuable. In particular, the degree and rate of relaxation within tablets, immediately after the point of maximum compression, have been shown to be characteristic for a particular system. Recording this phase of the cycle as well can provide valuable insight into the reasons behind inferior tablet quality and may suggest a remedy. For example, if the degree and rate of relaxation are high, addition of some plastically deforming component, such as polyvinyl pyrrolidone, may be advisable to reduce the risk of pronounced recovery leading to structural failure.

If the stress relaxation process involves plastic flow, it may continue after all compressional force has been removed, and the residual radial pressure will decay with time. The plastic flow can be interpreted in terms of a viscous and elastic parameter in series. This interpretation leads to a relationship of the form:

$$\text{Log } F_t = \text{Log } F_m - Kt \quad \text{... (15)}$$

where, $F_t$ is the force left in the viscoelastic region at a time $t$, and $F_m$ is the total magnitude of this force at time $t = 0$ (i.e. when decompression begins). $K$ is the *viscoelastic slope* and a measure of the degree of plastic flow. Materials with higher $K$ values undergo more plastic flow; such materials often form strong tablets at relatively low compaction forces.

Alternatively, the changing thickness of the tabletting mass due to the compactional force, and subsequently due to elastic recovery during unloading, can be used to obtain a measure of *plastoelasticity*, Specifically:

$$\gamma = \frac{H_o}{H_m} \cdot \frac{H_r - H_m}{H_o - H_m} \quad \text{... (16)}$$

where, $H_o$, $H_m$, and $H_r$ are the thickness of the tablet mass at the onset of loading, at the point of maximum applied force, and on ejection from the die, respectively. A linear relationship between $\gamma$ and log reciprocal of the tensile strength of the tablets has been demonstrated. In general, values of $\gamma$ above 9 tend to produce tablets that are laminated or capped.

## *Compaction Profiles*

Monitoring of that proportion of the applied pressure transmitted radially to the die wall has been reported by several groups of workers. For many pharmaceutical materials, such investigations lead to characteristic hysteresis curves, which have been termed *compaction profiles*. Fig. 13.48 is a typical example. Remember that the radial die-wall force arises as a result of the tabletting mass attempting to expand in the horizontal plane in response to the vertical compression. The ratio of these two dimensional changes, the *Poisson ratio*, is an important material-dependent property affecting the compressional process. The ratio is a property of solid bodies, however, and not necessarily of a porous mass of particulate solid. The anomalous results discussed in the literature may well reflect this important distinction, but certain qualitative deductions may still be possible.

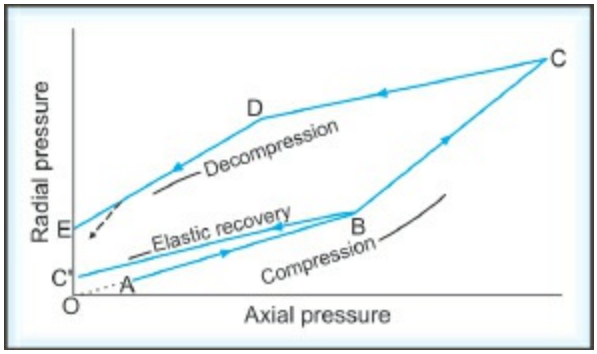



**Fig. 13.48:** Examples of compaction profiles. Dotted line O to A represents a highly variable response due to repacking, while at A, elastic deformation becomes dominant and continues until the elastic limit B is reached. From B to the point of maximum compression C, deformation is predominantly plastic, or brittle fracture is taking place. The decompression process C to D is accompanied by elastic recovery, and if a second yield point (D) is reached, by plastic deformation or brittle fracture D to E. The decompression line B to C represents the behavior of a largely elastic material

For instance, when the elastic limit of the material is high, elastic deformation may make the major contribution, and on removal of the applied load, the extent of the elastic relaxation depends upon the value of the material's modulus of elasticity (Young's modulus). If this value is low, there is considerable recovery, and unless a strong structure has been formed, there is the danger of structural failure. Maximum compressional force levels are particularly important in such cases, since most of the stored energy is released on removal of the applied load. Conversely, if the modulus of elasticity is high, there is a small dimensional change on decompression and less risk of failure. The area of the hysteresis loop (OABC') indicates the extent of departure from ideal elastic behavior, since for a perfectly elastic body, line BC' would coincide with AB.

In many tabletting operations, the applied force exceeds the elastic limit (point B), and brittle fracture and/or plastic deformation is then a major mechanism. For example, if the material readily undergoes plastic deformation with a constant yield stress as the material is sheared, then the region B to C should obey the equation:

$$P_R = P_A - 2S \quad \text{... (17)}$$

where, $S$ is the yield strength of the material. Note that the slope of this plot is unity, so that marked deviation from this value may indicate a more complex behavior. Deviation could also be due to the fact that the material is still significantly porous, which would invalidate the analogy to a solid body. Until this difference can be resolved, little is to be gained in proposing mathematical solutions for the region BC, which is often nonlinear anyway. This does not mean, however, that compaction profiles themselves cannot provide further useful information.

For example, since point C represents the situation at the maximum compressional force level, the region CD is therefore the initial relaxation response as the applied load is removed. In practice, many compaction profiles exhibit a marked change in the slope of this line during decompression, and a second yield point (point D) has been reported.

Perhaps the residual radial pressure (intercept EO), when all the compressional force has been removed, is more significant, since this pressure is an indication of the force being transmitted by the die wall to the tablet. As such, it provides a measure of possible ejection force level and

likely lubricant requirements; if pronounced, it suggests a strong tablet capable of at least withstanding such a compressive pressure. Conversely, a low value of residual radial pressure, or more significantly, a sharp change in slope (DE) is sometimes indicative of at least incipient failure of the tablet structure. In practical terms, this may mean introducing a plastically deforming component (e.g. PVP [polyvinylpyrrolidone] as binder, starch as diluent) to facilitate dissipation of these stresses, and hence a more gradual change in slope of the decompression plot, a preferred feature. In one recent study, modified compaction profiles ($P_A - P_R$) versus ($P_A + P_R$) were able to distinguish between readily consolidating and nonconsolidating materials. Specifically, two characteristic parameters (a normal stress value at zero shear and a minimum shear stress value), obtained from the unloading portion of the cycle, were shown to correlate with tensile strength and surface hardness of the compacted materials.

## Energy Involved in Compaction

Tablet machines, roll compactors, extruders, and similar types of equipment require a high input of mechanical work. The ways in which this work is converted into other forms of energy during these processes is of interest in both research and production areas. More specifically, the work requirement is a key factor in machine design, and any proportion of the applied energy stored in a product such as a tablet retains a destructive capability.

The work involved in various phases of a tablet or granule compaction operation includes (1) that necessary to overcome friction between particles, (2) that necessary to overcome friction between particles and machine parts, (3) that required to induce elastic and/or plastic deformation of the material, (4) that required to cause brittle fracture within the material and (5) that associated with the mechanical operation of various machine parts.

Normally, an appreciable amount of the energy supplied is converted to heat, which of course does not contribute toward the main objective of the process. On a theoretic basis, however, this heat does provide a means of monitoring the energy balance in the system. For example, one of the earliest experimental reports was that of Nelson and associates, who compared the energy expenditure in lubricated and unlubricated sulfathiazole granulations as shown in Table 13.12. Lubrication reduced energy expenditure by 75%, chiefly because of a lessening of the major component, namely energy utilized during ejection of the finished tablet. Note that lubrication has no apparent effect on the actual amount of energy required to compress the material, i.e. overcome resistance to relative interparticulate movement.

**Table 13.12:** Energy expended in compression of 400 mg sulfathiazole granulation

| **Compression** | **Energy expended (Joules)** | |
|---|---|---|
| **Process** | **Unlubricated** | **Lubricated** |
| Compression | 6.28 | 6.28 |
| Overcoming die wall friction | 3.35 | negligible |
| Upper punch withdrawal | 5.02 | negligible |
| Tablet ejection | 21.35 | 2.09 |

| | | |
|---|---|---|
| Totals | 36.00 | 8.37 |

These workers' estimation of the total work involved, $W_T$, was obtained by monitoring punch force and the distance D through which it acted, so that:

$$W_T = \int_{D_{F=0}}^{D_{max}} F \cdot dD \qquad \ldots (18)$$

which represents the area under the entire force-displacement plot (the area ABC in Fig. 13.49).

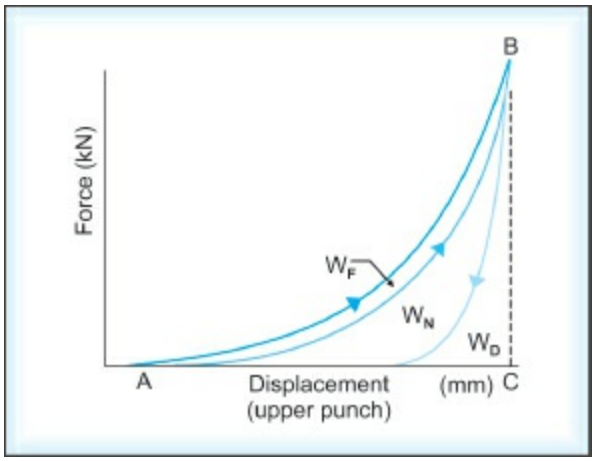



**Fig. 13.49:** Example of force-displacement (F-D) curve. (△) upper punch force; (▲) lower punch force. The area $W_F$ represents the work done in overcoming friction, while that of area $W_D$ is the elastic deformation energy stored in the tablet during compression. Thus, $W_N$ is the net mechanical energy actually used to form the tablet

This approach oversimplifies the true picture because, as can be seen from Fig. 13.49, in which both punch forces and punch displacement data have been collected throughout an entire compression-decompression press cycle, $W_T$ comprises at least three components. The region $W_F$ represents the work done in overcoming friction and therefore depends upon the properties of the tablet mass. $W_N$ is the net mechanical energy actually used to form the tablet, and $W_D$ is the elastic deformation energy stored in the tablet initially, but released during decompression. If the top punch moves too quickly during this decompression, contact with the tablet may be lost. In this eventuality, the complete work of elastic recovery of the tablet is not

transferred to the punch face, and an error is introduced into the decompression curve.

Early investigators of the technique overcame the problem by compressing the tablet a second time before ejecting it from the die. The second compression-decompression cycle provides a measure ($W_{N2}$) of the net energy required to recompress the material to the point *B* (in Fig. 13.49). $W_{N2}$ is equivalent to the amount of energy involved in the elastic recovery of the tablet.

## *Force-displacement (F-D) Curves*

Distinctive F-D curves related to the stress/strain properties of the materials involved have now been reported by several groups of workers. The technique has been shown to provide a more sensitive method for evaluating lubricant efficiency than the widely quoted R value (which is the lower to upper punch force ratio). For example, the data given in Table 13.13 show that R values are incapable of distinguishing between the result of incorporating lubricant in the granulation and the result of coating it on the die wall. The $W_N$ measurements, however, clearly indicate the lower energy expenditure if the granulation is lubricated.

**Table 13.13:** Compression of 300 mg at 440 MPa

| | Unlubricated | Die wall lubricated | Granulation lubricated |
|---|---|---|---|
| Coefficient of lubrication, R | 0.84 | 0.98 | 0.98 |
| Network for compression (Nm) | 5.6 | 4.4 | 3.4 |
| Remaining lower punch pressure (MPa) | 3.2 | 2.5 | 2.5 |

The wider utility of F-D curves is exemplified by their application to the selection of a “best” binder (from gelatin, starch, and methylcellulose) for a sulphonamide tablet, by determining the contribution of the three components $W_F$, $W_N$, and $W_D$ to the total work $W_T$. A plot of $W_N$ versus maximum compressional force produced curves such as those shown in Fig. 13.50. Since good correlation is usually found between *WN* and the crushing strength of the tablets, gelatin in this example would be chosen as the best binder. Note too the pronounced flattening of the curve for starch, so that compression of this formulation at maximum $F_A$ values above the point of

inflection would not be helpful and might even be deleterious, owing to increased elastic recovery that could lead to structural failure of the tablet.

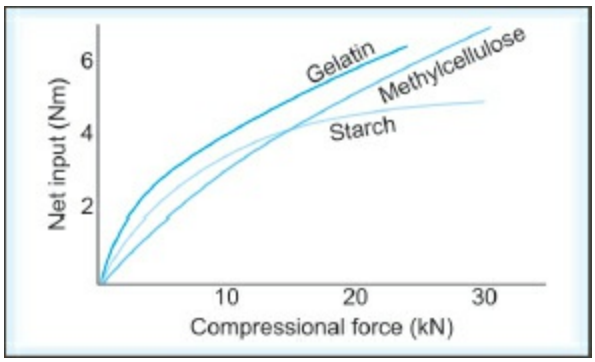



**Fig. 13.50:** Binder selection using net energy input during tabletting process

Because the rate of compression is known to affect the tabletting process, an extension of the use of F-D curve data has been suggested. This proposition takes into account the rate of loading by monitoring the power expended, rather than the work involved. In practice, the area under the F-D curve is divided by the time over which the force is applied.

## The Instrumentation of Tablet Machines

Several distinct forms of instrumentation are discussed within this section. Increasing evidence shows the value of instrumenting tablet presses to provide information on the inherent compaction characteristics of the major components in a formulation, and on the effect of additives upon them. The emphasis, therefore, is on properties of the materials in a research and development environment, which perhaps could utilize single-station machines for some of the work. On the other hand, increasing use of highspeed multi-station tablet presses, coupled with a desire to improve quality specifications for tablets, leads to forms of instrumentation intended primarily for production machines, but of interest to the quality assurance department as well.

The technology described in the first part of the chapter suggests that the mechanisms involved in the tabletting process center on utilization of the unsatisfied bonds at the solid surface. This process is enhanced by the generation of large areas of clean surface, which are then pressed together, as might occur if appreciable brittle fracture and plastic deformation were introduced into the system. It was also noted that the behavior on decompression can markedly affect the characteristics of the finished tablets, because the structure must be strong enough to accommodate the recovery- and ejection-induced stresses. Furthermore, ability to monitor ejection forces leads to valuable information on lubricant efficiency. Measurement of punch and die forces plus the relative displacement of the punches can provide raw data, which when suitably processed and interpreted, facilitate evaluation of many of these tabletting parameters.

The value of using a single-station press for developmental work on formulations that are to be manufactured on multi-station presses is strictly limited. This is less the case when material rather than process factors are most important, unless the rate of loading is critical. The sensitivity of the formulation being tested to the loading rate, however, should be determined by compression at different speeds and by monitoring of any changes in tablet properties.

Because of this rate factor, several workers have elected to instrument isolated punch and die sets and to carry out compression experiments using these sets in conjunction with a compression/tension testing machine. For

example, the assembly shown in Fig. 13.51 is capable of monitoring all of the various forces acting on the system as well as the punch displacements. Recently, a sophisticated system has been described that is capable of mimicking (in real time) the precise compression cycle of any press; thus, it has all the advantages of using a single station of tooling and still follows rotary press action. The system stores a representation of the precise compression cycle of the press in a microprocessor, which in turn controls the movements of the isolated punches, so that it reproduces the exact loading profile of the press. Such assemblies facilitate compressional studies at various compressional rates and provide a convenient means of acquiring the maximum amount of information while using the minimum amount of time and materials. They are now often referred to as "compaction simulators."

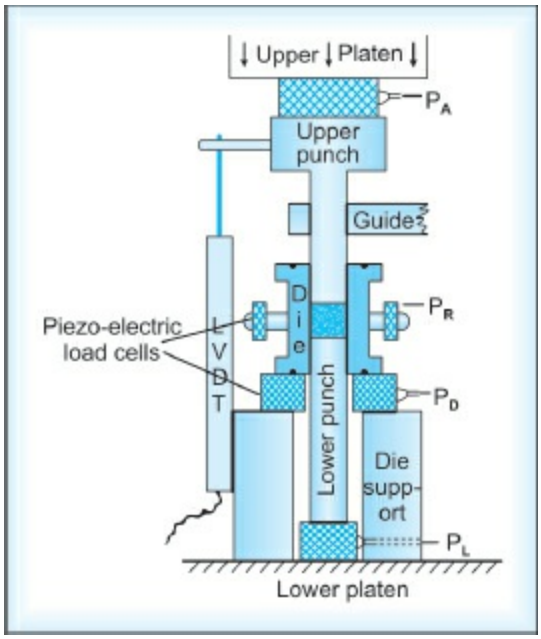


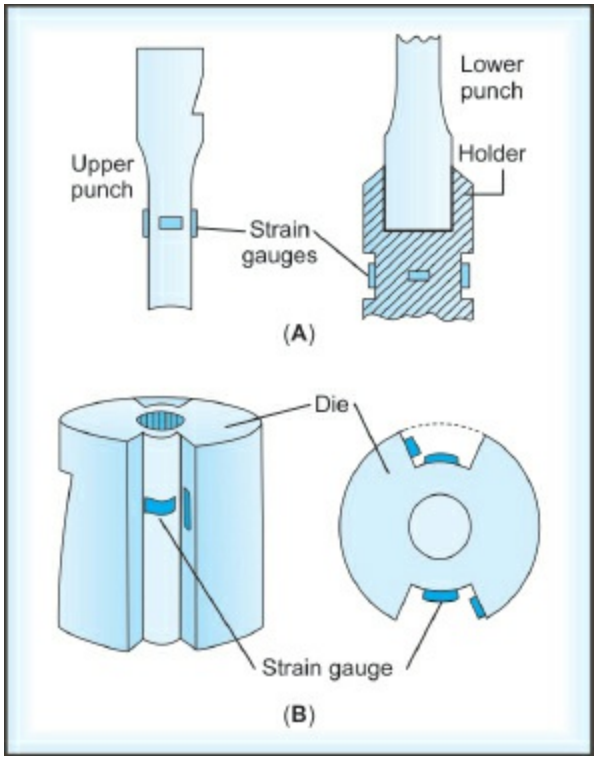

**Fig. 13.51:** Diagram of instrumented, isolated punch and die assembly

## *Single-station Presses*

Almost all reports in the literature describe strain gauge networks as a transducer for measuring the magnitude of the forces operating during the compression cycle. Resistive (metal foil) gauges are usually preferred, and ideally, they should be bonded as near to the active site as practicable, namely at the punch faces, so as to eliminate lack of correlation between signals obtained from remote regions of the machine frame and the actual

forces present in the tooling. The bonding must be over the entire area of the correctly aligned gauge as described in strain gauge manuals, at a site where the elastic change in linear dimension of the stressbearing member (due to the applied forces) can be measured.

An example of one of the commonest arrangements is shown diagrammatically in Fig. 13.52. The die-wall instrumentation requires machining of the die wall to accommodate the gauges and reduce the thickness to a point at which adequate sensitivity is achieved. The original geometry is subsequently restored with silicone rubber or similar material. The foregoing procedure may necessitate annealing and subsequent rehardening and tempering of the die. Care is needed to ensure that such treatment does not change the precise geometry of the bore.




**Figs 13.52A and B:** Strain-gauged punches (A) and instrumented die (B)

Assemblies such as those shown in Fig. 13.52 may be connected so as to conform to a Wheatstone bridge network (Fig. 13.53), which is normally energized by an AC amplifier system, the attenuated output giving a DC voltage proportional to the force being applied. Since changes in resistance are small, a full bridge network (with strain gauges in all four arms) is

preferable; one should bear in mind that such a system is then input-voltage-dependent, and therefore a stabilized supply is essential. If this arrangement is mounted as shown in Fig. 13.54 (“Poisson” configuration), compensation is provided for both temperature change and any bending of the piece.

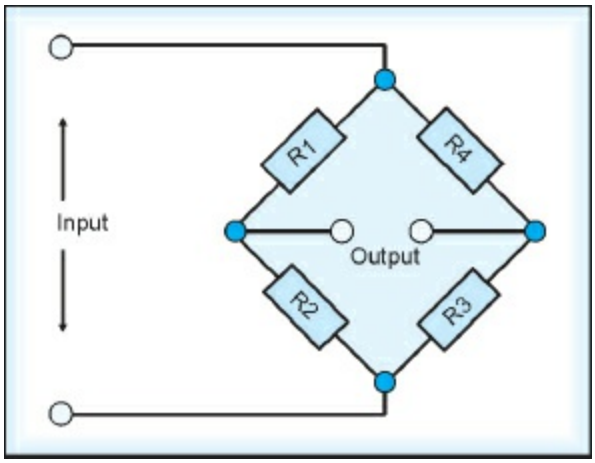



**Fig. 13.53:** Simple full-bridge Wheatstone bridge circuit, in which R1, R2, R3, and R4 represent strain gauges

Alternatively, transducers based upon the piezo-electric effect in certain crystals, notably quartz, may be used. When subjected to external forces, these materials develop an electrical charge proportional to the effect of the force. Such a transducer is connected by a high impedance cable to a charge amplifier, which converts the charge into a directly proportional DC voltage. One disadvantage of such transducers is that the charge inevitably dissipates with time, and since a difference is being measured, they are unsuitable for static force measurements.

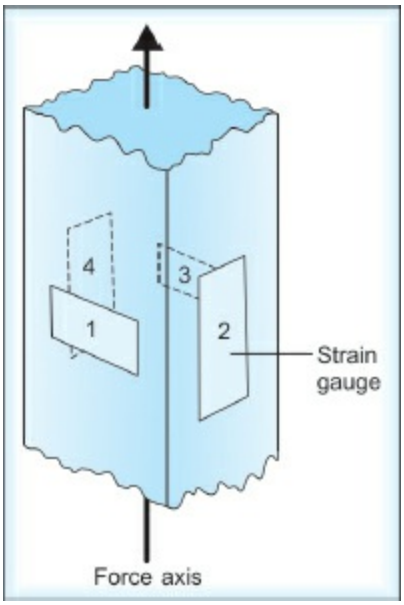



**Fig. 13.54:** Poisson arrangement. In this full-bridge strain gauge network, gauges 2 and 4 are active, while gauges 1 and 3 are temperature-compensating

Small piezo-electric load washers can easily be mounted on, or in, the upper and lower punch holders of single-station presses, as shown in Fig. 13.55. Radial die wall measurement is more difficult, but has been achieved by means of a special holder for the transducer, as illustrated in Fig. 13.56. Such systems, like their strain gauge equivalents, only monitor the radial force over a localized region of the die wall. They may give rise to misleading data, unless they are sited at the same level on the die wall as the region at which the tablet is being compressed. The advantages of piezo-electric devices include high sensitivity, robust construction, and no bonding to machine fabric, which therefore allows easy changing of tooling. Furthermore, since the signal originates as an electrical charge, the transducer may be zeroed, simply by grounding it, regardless of initial load.

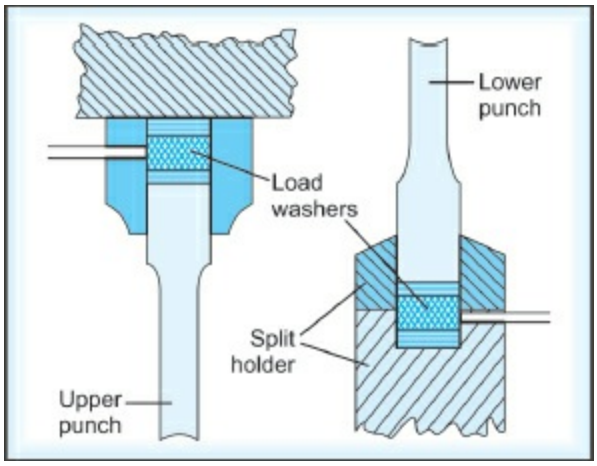



**Fig. 13.55:** Piezo-electric punch instrumentation. Preferred locations for mounting small piezo-electric force transducers are the punch holders of a singlestation press

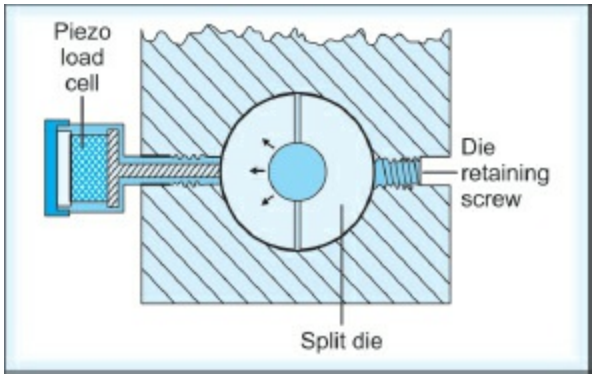



**Fig. 13.56:** Piezo-electric die wall instrumentation. An example of the use of a piezo-electric force transducer (in a special holder) in conjunction with a vertically split die to measure radial die wall forces

Although calibration data are supplied by the manufacturer of the above equipment, in situ calibration of the particular instrumentation against known loads is highly desirable. With punch assemblies, this may be achieved by use of a calibrated load cell, which can be placed on the die table (Fig. 13.57); its signal can be compared with those from the instrumented punches as they are simultaneously loaded. Alternatively, entire punch assemblies can be removed from the press and mounted in an accurate compression/tension test press. The normal procedure with die-wall instrumentation is less straightforward. Ideally, the necessary measurements are made by sealing the open ends of the die cavity, then pumping in oil at known pressures and

noting the response. A close-fitting rubber plug or rubber powder in the die cavity, however, may be regarded as a perfectly elastic material (thus radial die-wall reaction will be equal to any applied force) and can therefore be substituted for oil.

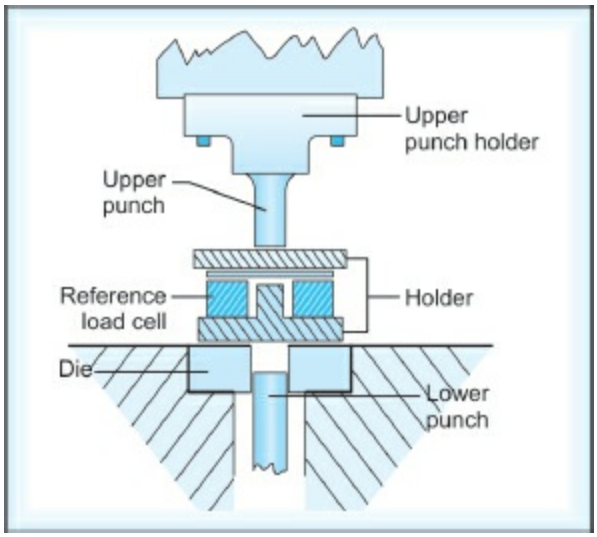



**Fig. 13.57:** Calibration using load cell. Diagram illustrates the use of a piezo-electric transducer and special holder to calibrate the upper and lower punch instrumentation on a single-station press

The preferred form of transducer for measuring punch displacements is based on the differential inductor principle, and is commonly a *linear variable differential transformer* (LVDT) as shown diagrammatically in Fig. 13.58. The movable ferrous core of the transducer is rigidly connected to the punch by a mechanical link, so that movement unbalances the “secondary” circuit, the output of which is attenuated to produce a DC voltage directly proportional to displacement.

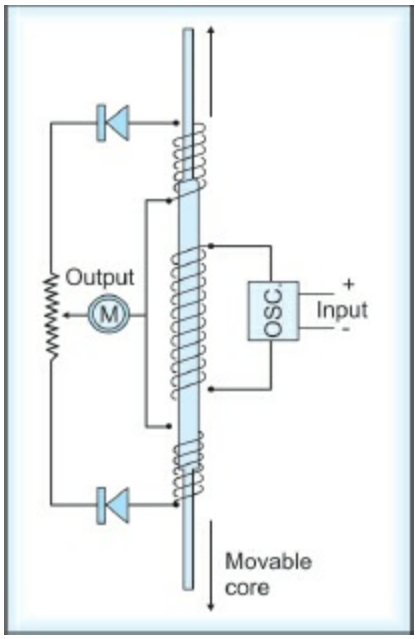



**Fig. 13.58:** Typical LVDT circuit

## *Multi-station Presses*

One of the major differences in the instrumentation of multi-station as opposed to single-station presses is the inherent difficulty in retrieving electrical signals from a revolving turret. Some early workers overcame this difficulty by employing radiotelemetry to transmit the force signal from strain-gauged upper and lower punches to external recorders. They were unsuitable for normal production conditions, since only a few stations were operative, and only low machine speeds were possible. In a system described by Ho and co-workers, however, these disadvantages have been overcome, and a significant proportion of the stations remain active, as shown in Fig. 13.59.

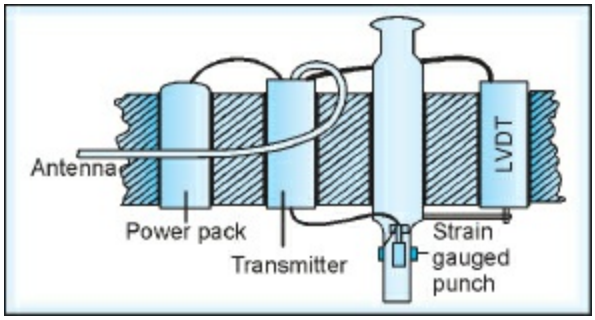



**Fig. 13.59:** Telemetry system. Diagram shows a possible arrangement for a

force and displacement measuring system, using radiotelemetry to retrieve the signals from a multi-station tablet press

Other workers have been less reluctant to utilize remote stationary sites on the machine frame, in particular the upper and lower compression roll carrier systems. For a given machine, the best location probably depends upon its actual design, but certain general points are worth noting. The response of a strain gauge is entirely a function of the change in linear dimension of the machine part to which it is bonded. Therefore, such changes must be sufficient to induce an adequate change in gauge resistance while not exceeding its elastic limit. In practice, this usually means that some part of the press has to be weakened by machining, an operation in which great care must be exercised. Cast iron components are unsuitable because of variability in their modulus of elasticity and Poisson ratio; therefore, only parts constructed of steel should be selected. Sensitivity to temperature changes may also influence site choice so as to avoid susceptible regions, e.g. near electric motors.

One of the more popular arrangements on modern high-speed presses is to attach strain gauges or to incorporate a piezo-electric load cell into one of the tie rods, as illustrated in Fig. 13.60, although strain gauges may require machining of the rod. Other sites have included the compression columns, specially modified pressure rolls, and a modified eyebolt of the lower compression roll assembly.

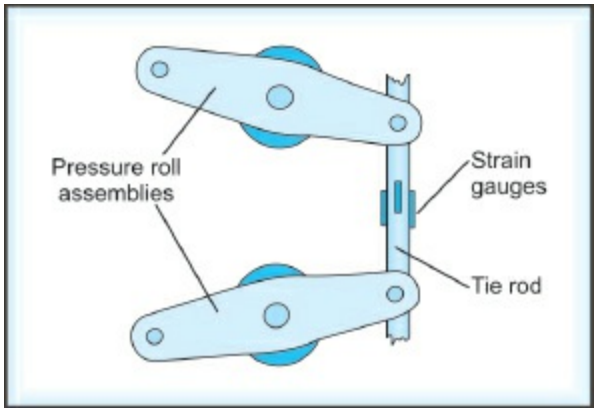



**Fig. 13.60:** Instrumented tie rod. Diagram shows the location of strain gauges on the tie rod linking the upper and lower pressure roll carriers of a multi-station tablet press

Instrumenting the normal ejection cam on a rotary tablet machine by attaching strain gauges to its bolts is of limited value, because the resulting signal is a summation of the effects of the several lower punches on it at any instant. The solution adopted by Wray employed a two-part cam so that the region responsible for tablet ejection is separate and in the form of a beam fixed at one end. This method necessitated minor modifications to other parts of the machine but did not affect normal operation. Flexure of the beam caused by the lower punches during ejection was monitored by strain gauges and was found to mimic the ejection response of an instrumented lower punch.

Alternatively, the normal cam can be cut into three sections, each clamped to the frame by a bolt, two of which are fitted with a piezoelectric load washer, for instance. The division should be such that there is only one punch on each section at a time (Fig. 13.61). The first transducer then monitors the force to initiate ejection (to break tablet die-wall adhesions), and the second monitors the force necessary to push the tablet clear of the die. This arrangement minimizes the fulcrum effect, as the punches move over the cam surface toward and then away from the actual transducer location. Certain aspects of the state of the tooling, such as sticking of the lower punches due to frictional effects, can also be detected by sensitive instrumentation of this type.

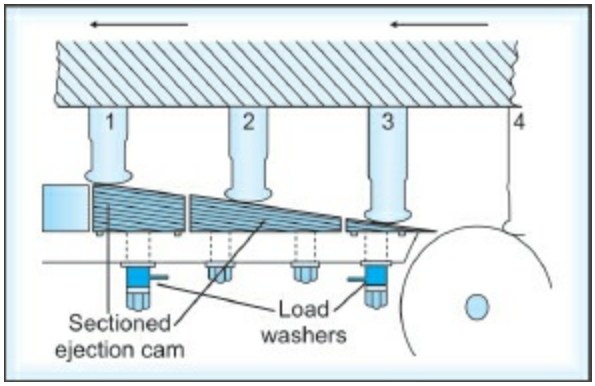



**Fig. 13.61:** Instrumented ejection cam. Diagram shows the preferred method of sectioning an ejection cam and the transducer locations for monitoring ejection forces on a multi-station tablet press

Mitrevej and Augsburger have recently described a system to measure the adhesion of tablets to the lower punch face by attaching strain gauges to a

small cantilever blade mounted on the feed frame in front of the sweep-off attachment (Fig. 13.62). They found that the force of adhesion did not necessarily reflect the ejection force or the lubricant activity of the formulation; however, the system did appear to be sensitive to batch variations in the antiadherent quality of magnesium stearate.

Regardless of which remote site is selected for instrumentation, the response should always be checked—and indeed, rechecked periodically—against signals obtained from directly instrumented tooling over the whole working range of the machine, to ensure constancy in the response relationships.

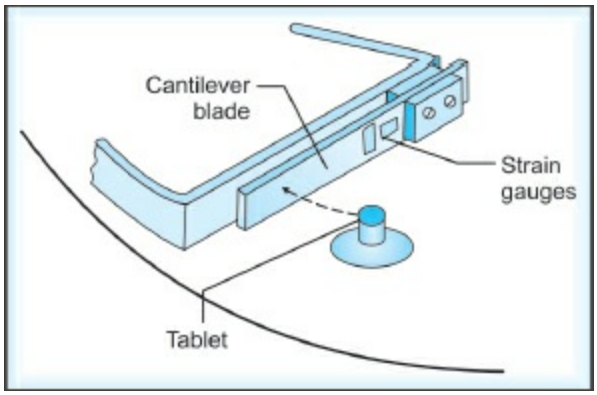



**Fig. 13.62:** Instrumentation to measure "sweep-off" force. Diagram illustrates measurement of the force of adhesion between a tablet and the lower punch of a tablet press by means of a strain-gauged cantilever blade attached to the feed frame

## *Signal Processing*

The signals from the instrumentation described in this section are usually DC voltages and can therefore be retrieved, stored, and processed by a common means. Popular practice is to display the signals on a cathode ray oscillograph (CRO) since this enables instant visualization of the instrumentation output. In the past, such displays were often photographed to provide a permanent record, but ultraviolet recording oscillographs provide better definition of traces and can facilitate a larger number of simultaneous recording channels, as illustrated in Fig. 13.63.

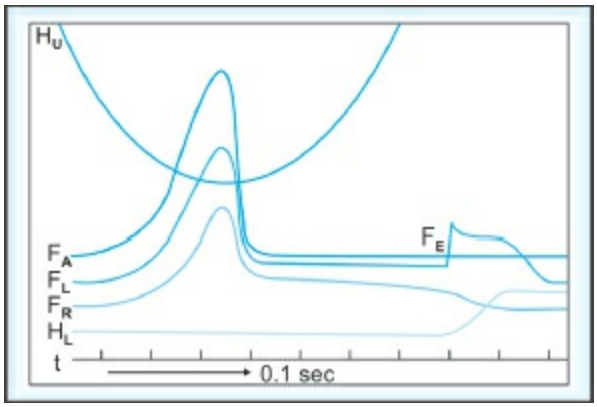



**Fig. 13.63:** Typical traces from ITM reproduction of typical traces obtained from a multichannel UV recording oscillograph connected to an single-station tablet press

Inexpensive microcomputers that are currently available can remove much of the tedium in reducing raw data from the recorders previously described; they are therefore the method of choice. The analog signals (DC voltages) can be fed by an A-to-D (analog-to-digital) convertor into memory locations in the computer. This digital data can then be recalled, manipulated, and outputted in a wide range of graphic or tabular formats. By such means, active compounds and excipients can be "fingerprinted" for their compactional characteristics. The general layout of a typical configuration is shown in Fig. 13.64.

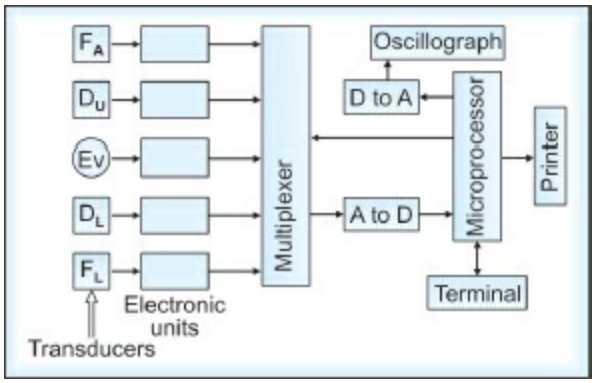



**Fig. 13.64:** ITM-microprocessor linkage. Diagram shows the major components of a typical instrumented tablet machine interfaced to a dedicated computer system

## *Role of Instrumentation in Production*

The modem tablet production department seeks innovation because of the trend toward direct compression methods, the availability of machines with increased output, and the desire to lessen tablet-to-tablet variations. In addition the design of instrumentation that monitors, or perhaps exercises some degree of control over, the tabletting process is an attractive goal because of the possibility of reducing labor involvement.

To date, the most popular approach has been the attempt to limit weight and hardness deviations, based upon the premise that compressional force is directly proportional to tablet weight, providing the following:

1. The formulation is homogeneous (i.e. has uniform density).
2. The compressional force/tablet weight function is constant.
3. The volume of the die cavities at the point of maximum compression is constant.

The third supposition is valid only when the overall length of the punches and tip geometry are constant, the die bores are uniform, and the pressure rolls are perfectly cylindric and mounted centrally.

The output from the electronic unit of most force (and displacement) measuring systems is a DC voltage. Therefore, in this context, the system produces a series of voltage pulses of short duration, each proportional to the weight of an individual tablet. These signals can be conditioned to provide a wide range of monitoring and control facilities of increased complexity. The addition of a simple event marker facilitates the identification of a particular station and of any tooling faults, which produce repetitive atypical signals. The individual compression pulses can also be used to drive counting mechanisms and to provide a reliable figure for the number of tablets made.

The foregoing devices are uncomplicated and relatively inexpensive, but usually, there is a desire to extend the signal conditioning system to monitor the process to some degree. For example, one can set upper and lower limits for acceptable tablet weight and then distinguish pulses from tablets lying outside these thresholds. When the frequency of these out-of-specification tablets exceeds some preset value, a relay can be tripped to activate an alarm and/or the machine can be automatically stopped.

In general, for a larger investment, machines can be fitted with mechanical accept/reject gates at the machine outlet, so that individual out-of-specification tablets can be diverted to a separate container. This function

requires a high level of sophistication, because the defective tablets must be “memorized” until they reach the outlet.

A second approach is to take the amplified output signals and feed them into an averaging network. This average DC voltage is compared with a reference voltage, and any difference is converted into an AC signal, which is amplified and used to drive a two-phase servomotor. The motor can be connected to the weight or pressure adjustment control of the press, so that any change in the average compressional force is reflected in an adjustment of either the weight or force control.

Regardless of which transducer systems, sites, or forces are selected, it is essential to ascertain that the response of the instrumentation is a direct function of the property needing to be monitored. Therefore, the work of Wray and his colleagues is important in that it establishes that stresses generated in certain parts of the machine frame are directly proportional to the punch forces, which in turn are related to compressional weights.

One final important aspect of instrumenting high-speed multi-station presses is the frequency response of the various components of the system. Machine outputs are now exceeding 12,000 tablets per minute, which means that the frequency of the force pulses is approximately 0.1 kHz. Since the detection of small differences in individual pulses may be necessary, all units should have flat responses well beyond this level, up to approximately 1.0 kHz.

Instrumented tablet machine technology is advancing rapidly, and its ultimate role is not yet realized. It will undoubtedly lead, however, to an even better understanding of the tabletting process, which in turn will assist in formulation development and batch quality control. In addition, the ever-increasing demand for more fully automated production will be facilitated by such machines.

---

* The various loads on a powder bed are sometimes expressed in terms of force, the preferred units being newtons (N). In other instances, the force acting over a unit cross-sectional area is used, i.e. a pressure. The unit in this case is the newton per square meter, which is called a Pascal (Pa). To facilitate comparison, expressions originally derived in other units have been converted to this (SI) system throughout this chapter.

# 14:
# Capsules

Mothes and Dublanc, two Frenchmen, are generally credited with the invention of the gelatin capsule. Their patents, granted in March and December of 1834, covered a method for producing single-piece, olive-shaped, gelatin capsules, which were closed by a drop of concentrated warm gelatin solution after filling. The two-piece telescoping capsule, invented by James Murdock of London (1848), was patented in England in 1865.

Capsules are solid dosage form in which the drug substance is enclosed within a hard or soft soluble shell generally formed from gelatin. In addition to having the advantages of elegance, ease of use, and portability, capsules have become a popular dosage form because they provide a smooth, slippery, easily swallowed, and tasteless shell for drugs; the last advantage is particularly beneficial for drugs having an unpleasant taste or odor. They are economically produced in large quantities and in a wide range of colors, and they generally provide ready availability of the contained drug, since minimal excipient and little pressure are required to compact the material.

Capsules may be classified as either hard or soft depending on the nature of the capsule shell. Soft gelatin capsules are made from a more flexible, plasticized gelatin film than hard gelatin capsule. Soft gelatin capsules are prepared by means of a plate process or, by a rotary die process, in which they are formed, filled and sealed in a single operation. Unlike, soft gelatin capsule, hard gelatin capsules are manufactured and filled in a completely separate operation. In further contrast to soft gelatin capsules, which are filled with liquids that will not solubilize the shell, hard gelatin capsules typically are filled with powders, granules or pellets. (Although recent modifications to hard gelatin capsule permit the filling of liquid or semisolid matrices, there currently are few commercial examples.) Additional attributes of both hard and soft gelatin capsules are discussed in the following section.

## CAPSULE SHELL COMPOSITION Gelatin

The capsule shell of both hard and soft capsules are made principally of gelatin blends and may contain plasticizers, water and small amounts of certified dyes, opacifying agents, flavorings, coloring agents and preservatives and medicaments to achieve desired effects.

Gelatin's chemical, physical, and physiological properties make it an ideal substance for the capsulation of pharmaceutical products. Gelatin is a heterogeneous product derived by irreversible hydrolytic extraction of treated animal collagen, and as such, it never occurs naturally. Its physical and chemical properties are mainly functions of the parent collagen, method of extraction, pH value, thermal degradation, and electrolyte content. Common sources of collagen are animal bones, hide portions, and frozen pork skin. Bone and skin gelatins are readily available in commercial quantities in most areas of the world.

Type A gelatin is derived from an acid-treated precursor and exhibits an isoelectric point in the region of pH 9, whereas type B gelatin is from an alkali-treated precursor and has its isoelectric zone in the region of pH 4.7. Although capsules may be made from either type of gelatin, the usual practice is to use a mixture of both types as dictated by availability and cost considerations. Differences in the physical properties of finished capsules as a function of the type of gelatin used are few.

Blends of bone and pork skin gelatins of relatively high gel strength are normally used for hard capsule production. The bone gelatin produces a tough, firm film, but tends to be hazy and brittle. The pork skin gelatin contributes plasticity and clarity to the blend, thereby reducing haze or cloudiness in the finished capsule. An abbreviated flowchart for the manufacture of gelatin to be used in capsules is presented in Fig. 14.1. Two recent developments have taken place in the gelatin supply area. First, "green" (fresh) bones are being used commercially as a source of Type B gelatin. Apart from additional pretreatment to remove residual tissues and fat, the processing coincides with that used for aged bones, and the gelatins obtained are indistinguishable from each other in practical use.

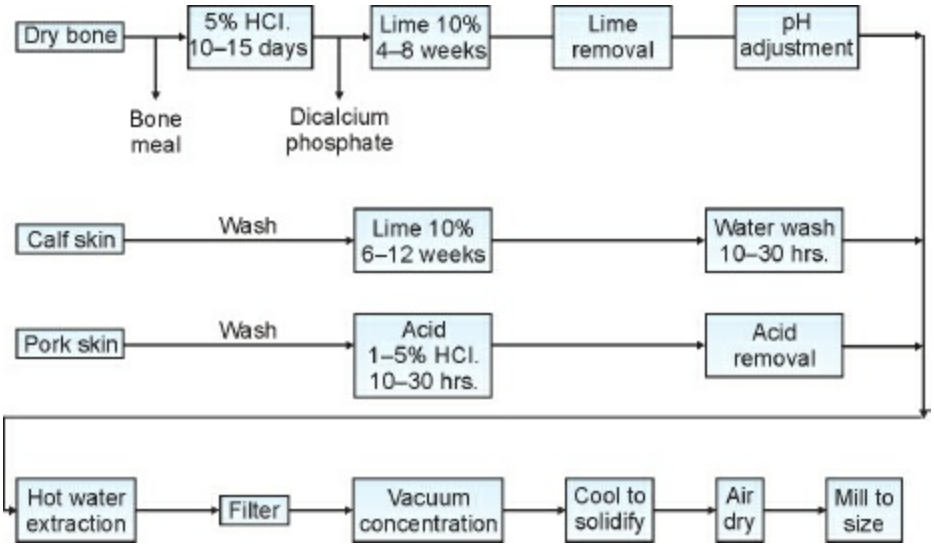



**Fig. 14.1:** The process of manufacturing gelatin used in capsules

The second development is the processing of an “acid-bone” gelatin prepared from bone by techniques essentially comparable to those for Type A gelatins. The resulting gelatin shows an altered isoelectric point (pH 5.5–6.0), and generally, intermediate physical characteristics for the film. The acid extraction technique for bones is valuable to processors of gelatin because of the decreased extraction time required. Both of the aforementioned materials are commercially available and are used in capsule production. The gelatin is USP grade with additional specifications required by the capsule manufacturer. The additional specifications concern the Bloom strength, viscosity, and iron content of the gelatins used.

The *Bloom* or *gel strength* of gelatin is a measure of the cohesive strength of the crosslinking that occurs between gelatin molecules and is proportional to the molecular weight of the gelatin. Bloom is determined by measuring the weight in grams required to move a plastic plunger of 0.5 inches in diameter, 4 mm deep into a 6⅔% gelatin gel that has been held at 10°C for 17 hours. Bloom may vary with the requirements of the individual custom manufacturer but ranges from 150 to 250 g. In general, with all other factors being equal, the higher the Bloom strength of the gelatin used, the more physically stable is the resulting capsule shell. The cost of gelatin is directly proportional to its Bloom or gel strength and thus is an important factor in the cost of soft capsules. Consequently, the higher Bloom gelatins are used only when necessary to improve the physical stability of a product or for large capsules (over 50 minims), which require greater structural strength during

manufacture.

*Viscosity* of gelatin, determined on a 6% concentration of gelatin in water at 60°C, is a measure of the molecular chain length and determines the manufacturing characteristics of the gelatin film. The desired film characteristics are usually based on standard gelatin formulations, which allow production at a set sealing temperature and definite drying conditions, and produce a firm, nontacky, nonbrittle, pharmaceutically elegant product. The viscosity for gelatin can range from 25 to 45 millipoise, but the individual manufacturer sets a narrow range, e.g. 38 ± 2 millipoise, for a particular type of gelatin, to make use of a standard formulation and thus conform to standard production conditions.

Low-viscosity (25 to 32 millipoise), high-Bloom (180 to 250 g) gelatins are used in conjunction with the capsulation of hygroscopic vehicles or solids, and standard gelatin formulas can be modified so as to require up to 50% less water for satisfactory operation on the capsulation machine. These modified formulas afford less opportunity for the hygroscopic fill materials to attract water from the shell and thereby improve the ingredient and physical stability of the product.

*Iron* is always present in the raw gelatin, and its concentration usually depends on the iron content of the large quantities of water used in its manufacture. Gelatin used in the manufacture of soft gelatin capsules should not contain more than 15 ppm of this element, because of its effect on Food, Drug, and Cosmetic (FD&C) certified dyes and its possible color reactions with organic compounds.

Although capsule shells made from gelatin predominate, recent years have seen an increased interest in nongelatin capsules made from starch, hydroxypropyl methylcellulose and pullalan. The properties and uses of these nongelatin capsules were described latter in this chapter.

## Plasticizer

The plasticizers are more common for soft gelatin capsules. The ratio (w/w) of dry plasticizer to dry gelatin determines the "hardness" of the gelatin shell, assuming that there is no effect from the capsulated material. (Some examples of glycerin/gelatin ratios are shown in Table 14.1 along with their typical usage.) The *plasticizers* used with gelatin in soft capsule manufacture are relatively few. Glycerin USP, Sorbitol USP, Pharmaceutical Grade Sorbitol Special, and combinations of these are the most prevalent.

**Table 14.1:** Typical shell "hardness" ratios and their uses

| **Hardness** | **Ratio dry glycerin/dry gelatin** | **Usage** |
|---|---|---|
| Hard | 0.4/1 | Oral, oil-based, or shell-softening products and those destined primarily for hot, humid areas. |
| Medium | 0.6/1 | Oral, tube, vaginal oil-based, water-miscible-based, or shell-hardening products and those destined primarily for temperate areas. |
| Soft | 0.8/1 | Tube, vaginal, water-miscible-based or shell-hardening products and those destined primarily for cold, dry areas. |

The ratio (w/w) of water to dry gelatin can vary from 0.7 to 1.3 (water) to 1.0 (dry gelatin) depending on the viscosity of the gelatin being used. However, for most formulations, it is approximately 1 to 1. Since only water

is lost during the capsule drying process, the percentage of plasticizer and gelatin in the shell is increased, but their ratio remains unchanged.

## Additional Components

In general, the additional *components* of the gelatin mass are limited in their use by (1) the amount required to produce the desired effect; (2) their effect on capsule manufactures and (3) economic factors. Examples of ingredients falling into the first two categories are shown in Table 14.2.

**Table 14.2:** Additional components of the gelatin mass

| Ingredient | Concentration | Purpose |
|---|---|---|
| *Category I* | | |
| Methylparaben (4 parts), Propylparaben, (1 part) | 0.2% | Preservative |
| FD&C and D&C water-soluble dyes, certified lakes, pigments, and vegetable colors, alone or in combination | q.s. | Colorants |
| Titanium dioxide | 0.2 to 1.2% | Opacifier |
| Ethyl vanillin | 0.1% | Flavoring for odor and taste |
| Essential oils | upto 2% | Flavoring for odor and taste |
| *Category II* | | |
| Sugar (sucrose) | upto 5% | To produce chewable shell and taste |
| Fumaric acid | upto 1% | Aids solubility; reduces aldehydic tanning of gelatin |

The addition of *medicaments* to the gelatin mass is not recommended usually due to economic reasons, since only 50% of the gelatin mass is incorporated into the capsules. This results in a 50% loss of the added medicament. However, certain highly active, relatively inexpensive compounds such as benzocaine (3 mg/capsule shell) in chewable cough capsules may be used successfully.

The *color* of the gelatin shell is an important aspect. This is particularly

true for soft gelatin capsules, in which the color of the capsule can be definitely affected by the color or type of material capsulated. As a general policy, the color of the capsule shell should never be lighter in hue than the capsulated material.

More specifically, darker colors are more appropriate for large-size oral products, since they will not accentuate the size. Also, before a color is chosen, mixtures should be checked in the laboratory by addition of water to ascertain if reactions take place to cause the mixture to darken, as in the case of ascorbic acid and iron salts in vitamin and mineral formulations, or as in the case of reactions between iron and compounds of a phenolic nature. Since iron is present in gelatin, dark spots may occur in the shell owing to the migration of water-soluble iron-sensitive ingredients from the fill material into the shell. As a rule, clear colors usually are employed with clear type fill materials, and opaque colors are used with suspensions, but the reverse of this rule can be chosen to achieve a particular appearance or for ingredient stability purposes. For special effects or identification purposes, two colors, both opaque or one opaque and one clear, may be chosen since the manufacturing process involves two gelatin films.

A publication by Horn and co-workers describes a gelatin disk method for the determination of the effects of agitation, temperature, dissolution medium, and shell composition on the dissolution rate of soft gelatin capsules. This information may be helpful in the formulation of gelatin capsules for various purposes.

From the foregoing discussion on the gelatin shell, one may conclude that the pharmaceutical chemist must rely heavily on the experience of the custom capsule manufacturer. However, in order to choose the proper gelatin, gelatin formula, and color, the custom manufacturer must rely on the technical and product information designed and developed by the pharmaceutical chemist. With such mutual cooperation and free exchange of information, new products or dosage forms can be efficiently developed.

# HARD CAPSULES

Hard or two-piece capsules allow for a degree of flexibility of formulation not attainable with tablets. Capsules are easier to formulate because there is no requirement that the powders be formed into a coherent compact that will stand upto handling. The majority of capsule fills are dry powder blends, which are typically simple mixtures and thus processing and filling require minimum stress. The release characteristics, rapid, controlled, or modified, of the capsule can be controlled efficiently because of the limited number of factors involved. Hard capsules can be filled with formulation that has a wide range of physical properties thus enabling the formulator to use many different types of excipients to achieve their desired effect.

From a manufacturing point of view, there perhaps is some disadvantage in the fact that the output of high speed tablet press is about five fold that of automatic capsule filling machine.

Furthermore, capsules are not usually used for the administration of extremely soluble materials such as potassium chloride, potassium bromide, or ammonium chloride since the sudden release of such compounds in the stomach could result in irritating concentrations. Capsules should not be used for highly efflorescent or deliquescent materials. Efflorescent materials may cause the capsules to soften, whereas deliquescent powders may dry the capsule shell to excessive brittleness. In some cases, this dehydration may be retarded or prevented by the use of small amounts of inert oils in the powder mixture.

## CAPSULE SHELL MANUFACTURE

The three major suppliers of empty gelatin capsules are Eli Lilly and Company, Indianapolis, IN; Capsugel, Greenwood, SC; and the R. P. Scherer Corporation, Troy, MI. Several smaller volume suppliers exist throughout the world, some of which process for their own use only.

The completely automatic machine most commonly used for capsule production consists of mechanisms for automatically dipping, spinning, drying, stripping, trimming, and joining the capsules. The stainless steel mold pin (Fig. 14.2), on which the capsule is formed, controls some of the final critical dimensions of the capsule.

(***Dipping***) One hundred and fifty pairs of these pins are dipped, as shown in Fig. 14.2, into a gelatin sol of controlled viscosity to form caps and bodies simultaneously. (Spinning) The pins are usually rotated to distribute the gelatin uniformly (1) during which time the gelatin may be set or gelled by a blast of cool air (2). (***Drying***) The pins are moved through a series of controlled air drying kilns for the gradual and precisely controlled removal of water (3 to 6). (***Stripping***) The capsules are stripped from the pins by bronze jaws and trimmed to length by stationary knives (7) while the capsule halves are being spun in chucks or collets. (***Trimming***) After being trimmed to exact length, the cap and body sections are joined and ejected from the machine. The entire cycle of the machine lasts for approximately 45 min.

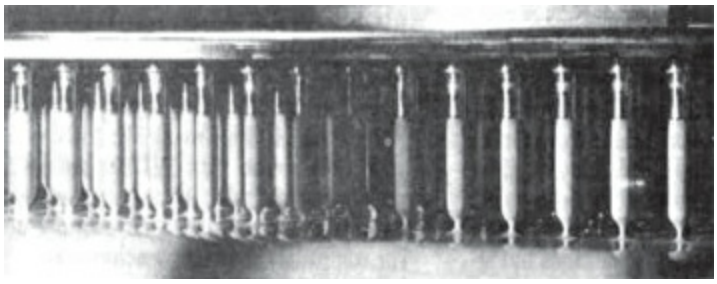


**Fig. 14.2:** Mold pin dipping

Thickness of the capsule wall is controlled by the viscosity of the gelatin solution and the speed and time of dipping. Other matters critical to the final dimensions are mold pin dimensions, precise drying, and machine control relating to cut lengths. Precise control of drying conditions is essential to the ultimate quality of the cast film. At the least, in-process controls include

periodic monitoring, and adjustment when required, of film thickness, cut lengths of both cap and body, color, and moisture content. Recent studies have been made in several areas to provide computer control of viscosity (and consequent wall thickness) during either machine operations or gelatin solution make-up.

**Inspection processes**—to remove imperfect capsules—which historically have been done visually, have recently been automated following the development and patenting of a practical electronic sorting mechanism by Eli Lilly and Company. This equipment mechanically orients the capsules and transports them past a series of optical scanners, at which time those having detectable visual imperfections are automatically rejected.

## Moisture Content

The optimum moisture content of capsule shell ranged between 12% and 15%. Below 10% moisture content, they become brittle and suffer dimensional changes, which may cause handling problems in the filling equipment. Moisture content above 16% may cause problems in the filling and in a loss of mechanical strength. At high moisture levels, the capsules absorb moisture, and may soften and become tacky. In severe cases, the capsules may absorb sufficient moisture to cause them to deform under their own weight. Exposure to either heat or moisture extremes can distort empty capsules to the extent that they cannot be handled by automatic filling equipment.

## Storage

Much consideration should be given to techniques for handling and storage of empty capsules in any production facility. This is of great importance when use rates are high, as when high-speed filling equipment is used. Empty capsules are subject to size variation as a result of moisture content variation. This can be caused by exposure to extreme variations in absolute humidity or elevated temperature. Exposure to extremely high or extremely low humidity conditions for extended periods after the containers are opened causes the capsules to either gain or loose moisture.

Handling of empty capsules is ideally carried in areas within the relative humidity range of approximately 30 to 45%, since major moisture content changes do not occur within these limits. If conditions drier than these are necessitated because of the ingredients being filled, exposures of the empty capsules prior to filling should be minimized. Strong consideration should be given to the use of air-conditioned facilities to control both temperatures and humidity when high-speed filling equipment is being operated.

## CAPSULE FILLING

The hard two-piece capsule can be filled with materials that have a wide range of physical properties. Photograph depicting filling of hard two-piece capsule is shown in Fig. 14.3, while the dosage forms that could be filled into capsules are shown in Table 14.3.

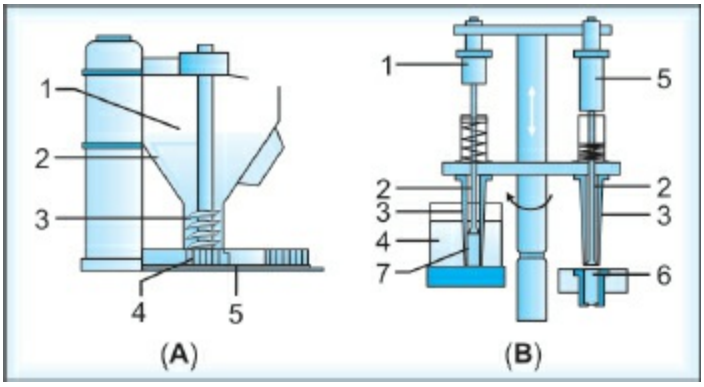

**Fig. 14.3:** Filling of hard two-piece capsules (*Courtesy* of Unichem)

**Table 14.3:** Types of dosage forms for filling into hard capsules

| Solid | Semisolid | Liquids |
|---|---|---|
| Powders | Paste | Non-aqueous |
| Granules | Thermosoftening | solutions |
| Pellets | Thixotropic mixtures | Oily liquids |
| Tablets | | |
| Capsules | | |

## Solid Filling Equipment

Filling machines are differentiated by the means by which they measure and feed the dose of material. The feeding mechanisms commonly employed in filling machines are of two types: (a) Dependent type—powder is transferred from a hopper directly to the capsule body and flow of the powder is aided either by a revolving auger (Fig. 14.4A) or by a vibrating plate. The powder inside these capsules is a loose fill. From auger machines higher fill weights are often achievable, (b) Independent type—compress the measured amount of powder to form a plug. The plug is formed inside a dosator or dosing tube with the help of a moveable piston that controls the dosing volume (Fig.14.4B). Dosator type feeding units are very versatile because the fill weight can be varied over the wide range. Filling machines are available with a range of outputs, from bench-scale to high-output industrial scale and from manual to fully automatic. Each machine type is briefly described in Table 14.4, but no attempt is made to quantify or compare weight variation figures among the various types because of obvious dependence on such factors as the conditions of the equipment, formulas, method of operation, operator competence, machine rates, and sizes of capsules.




**Figs 14.4A and B:** Feeding mechanisms of the capsule filling machine: (A) Auger feeding unit (1) powder hopper, (2) stirrer, (3) auger, (4) body holder, (5) turn table; (B) Dosator feeding unit (1) compression force, (2) piston, (3) dosing tube, (4) powder hopper, (5) plug ejection plate, (6) capsule body, (7) powder plug

**Table 14.4:** Description of capsule filling machines available from selected manufacturers

| Manufacturer unit | Model | Output (capsules/h) | Fill material | Feature: Feeding |
|---|---|---|---|---|
| *Eli Lilly & Co.* Indianapolis, IN (Fig 14.6) | Rotofil | 96,000 | Pellets | Auger type |
| Farmatic, Bologna. Italy(Fig 14.7) | 2000/15<br>2000/30<br>2000/60 | 40,000<br>80,000<br>160,000 | Powder | Dosator type |
| *Höfliger and Karg.* Waiblingen Germany (Fig 14.8) | GKF-303<br>GKF-602<br>GKF-1500<br>GKF-2500 | 18,180<br>36,120<br>90,000<br>150,000 | Powders, pellets, tablets thixotropic liquids<br>Powder, pellets | Auger type |
| Macofar, Bologna Italy (Fig 14.9) | MT-12<br>MT-13/1<br>MT-13/2 | 35,000<br>5,000<br>10,000 | Powders pellets, tablets | Dosator type |
| mG2, Bologna Italy (Fig 14.10, 14.11) | G36<br>G37N<br>G38 | 36,000<br>96,000<br>60,000 | Powder, capsule pellets | Dosator type |
| Osaka, Osaka Japan (Fig 14.12) | R-180 | 70,000 to 165,000 | Powder, granular principle | Vibratory fill |
| Perry, Green Bay WI (Fig 14.13) | ACCOFIL | 60,000 | Powder method | ACCOFIL |
| Zanasi, Bologna Italy (Figs 14.14 and 14.15) | 6E<br>40E | 6,000<br>40,000 | Powders, pellets, tablets, small capsules, paste, liquid | Dosator type |

Since the machines have essentially the same method of operation, only one description is given.

The Lilly machine is shown in Fig. 14.5. The empty capsules are fed from the storage hopper (1) and through the rectifying unit (2), into the two-piece filling ring (3A and 3B). Rectification is based on dimensional differences between the outside diameters of the cap and body portions of the capsule. The ring containing separable portions, lower portion (3A) and upper portion (3B), is rotated and a vacuum is applied on its underside. This vacuum seats the bodies into the lower portion of the ring, while the caps are retained in the upper portion. The two portion of the ring are then separated, and the upper (cap-containing) portion is set aside. The lower (body-containing) portion of the ring is placed on a variable speed turntable and is mechanically rotated under the powder hopper (4), which contains an auger for the forced delivery of the powder. After one (or more) complete rotation of the ring, the powder hopper (4) is removed, and the two portions of the ring (3A and 3B) are rejoined. The intact ring is positioned in front of the peg ring (5), and the

closing plate (6) is pivoted to a position approximately 180 degrees from that (Fig. 14.5). Pneumatic pressure is applied to the peg ring (5), which forces the capsule body into the cap, and the closing plate (6) holds the caps in position. Ejection of the filled capsules from the rings cannot occur with the plate in the closing position. For ejection of the capsules, the pressure is released, the closing plate is restored to its original position, and the capsules are expelled through the upper portion of the ring. Normal closing and ejection occur with the peg ring in a vertical position (*see* Fig. 14.5), with the filled capsules being collected through a chute (7) into a collection chamber.

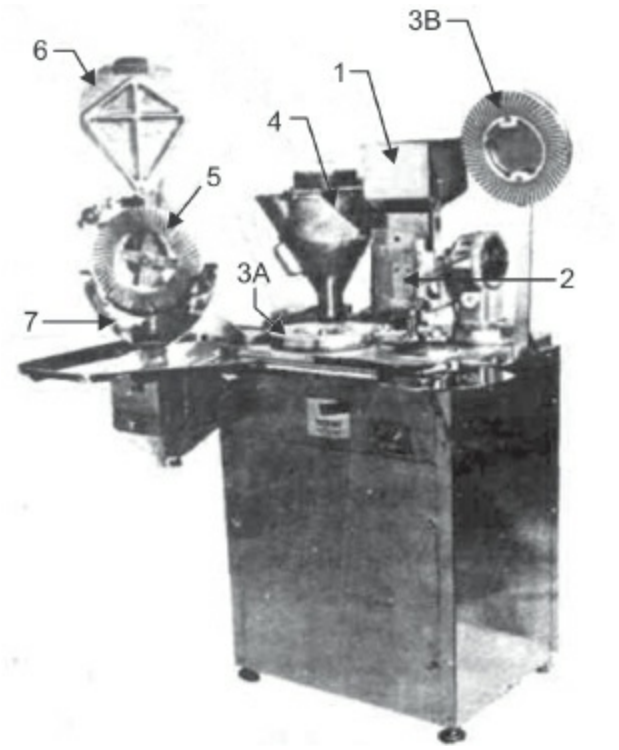



**Fig. 14.5:** Lilly capsule filling machine. *See* text for explanation of labels 1 through 7 (*Courtesy* of Eli Lilly and Co., Indianapolis, IN)

In this equipment, the powder is filled to the upper surface of the body-containing ring, and the fill is therefore primarily volumetric. Although changes in the total amount of powder can be caused by changes in the rotational speed of the turntable (which changes the amount of time for which each hole is under the auger), there is no way to produce a partially filled capsule consistently. Although slower speeds usually produce less weight variation, they also usually result in heavier total fill weights, which may not be economical because of the resultant decrease in productivity. Minimum

total fill weights (but usually maximum weight variation) are achieved with the highest turntable speed. Maximum total fill weights (but generally minimum weight variation) are achieved at the lowest rotational speed.

Some of the variables that must be properly controlled in order to achieve minimum weight variation and proper uniformity of the finished capsules are given in the following list:

1. The body-containing ring (3A) must be flat across its surface to avoid creating volumetric differences from one area of the ring to another.
2. The powder hopper (4) must be properly positioned during the Ming operation to avoid uneven powder distribution from the auger. The proper location includes consideration of both the centering of the auger over the ring holes and the parallelism between the lower surface of the hopper and the upper surface of the ring.

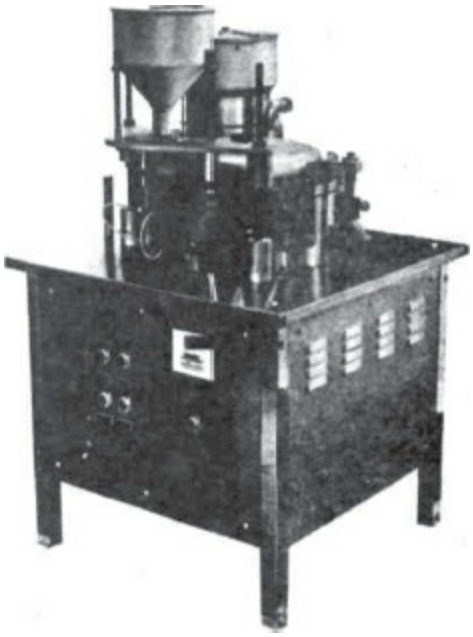


**Fig. 14.6:** Lilly ROTOFIL capsule filling machine (*Courtesy* of Eli Lilly & Co., Indianapolis, IN)

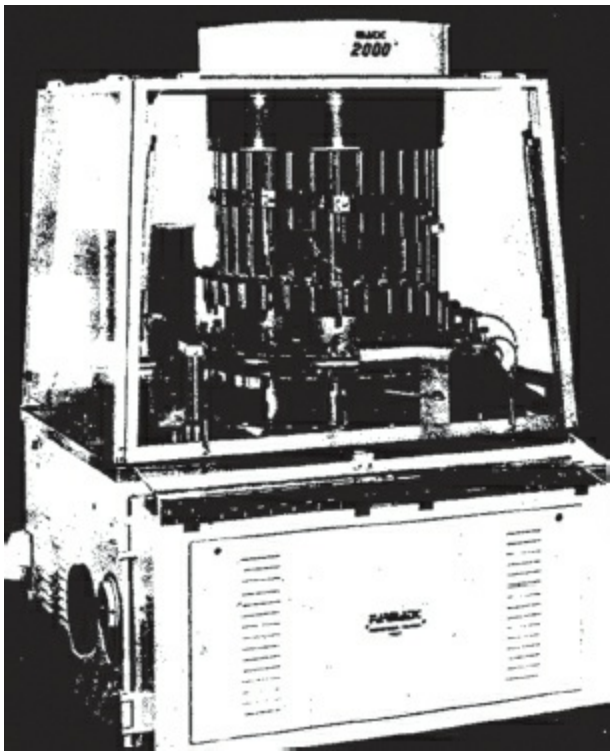



**Fig. 14.7:** Farmatic Model 2000/60 capsule filling machine (*Courtesy* of GB Gundi Bruno SpA, Bologna, Italy)

3. Extreme variations in powder level in the filling hopper (4) can cause uneven powder flow, resulting in excessive fill weight variation.
4. The individual rods in the peg ring must fit the rings being used, be of uniform length, and be perpendicular to the closing plate (6).

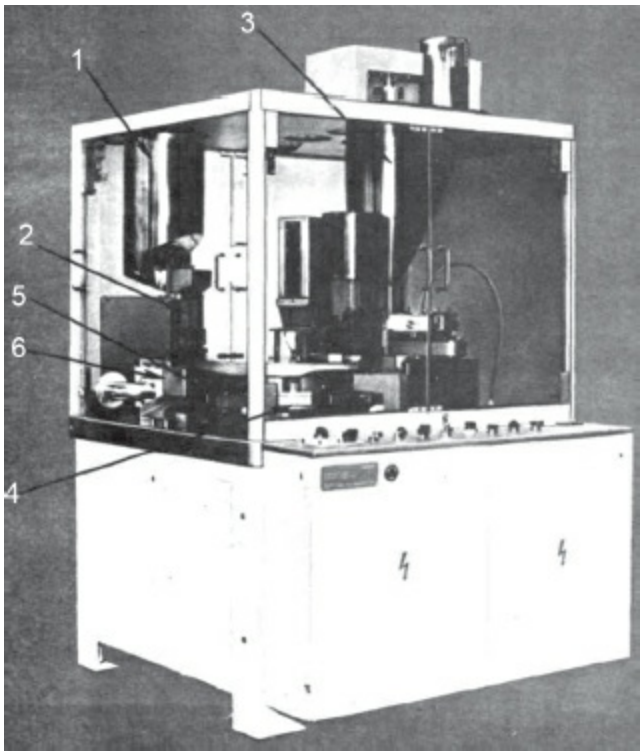



**Fig. 14.8:** H and K Model 602 capsule filling machine, 1. Empty capsule storage hopper, 2. Rectifier, 3. Bulk powder storage hopper, 4. Capsule body transport segment, 5. Closing station and 6. Filled capsules ejection station (*Courtesy* of Robert Bosch GmbH, Waiblingen, West Germany)

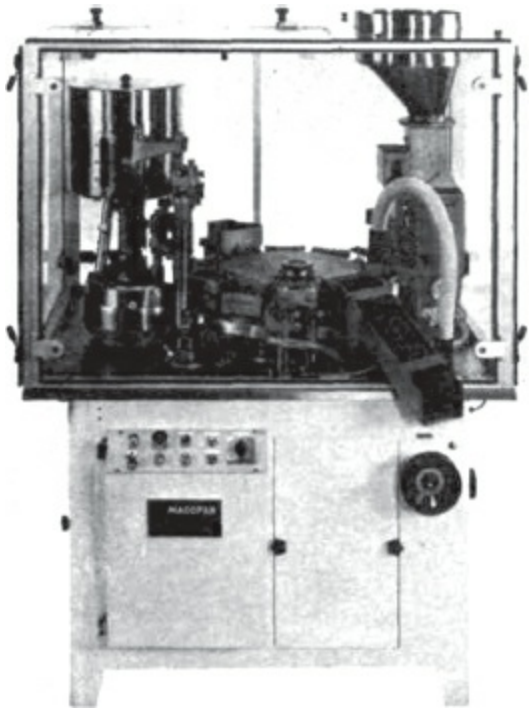


**Fig. 14.9:** Macofar model MT-12 capsule filling machine (*Courtesy* of Macofar, Bologna, Italy)

5. Flow properties of the powder being filled must be such that a constant amount of powder is available for delivery from the auger. Diluents and glidants should be selected with this phase of the operation in mind.

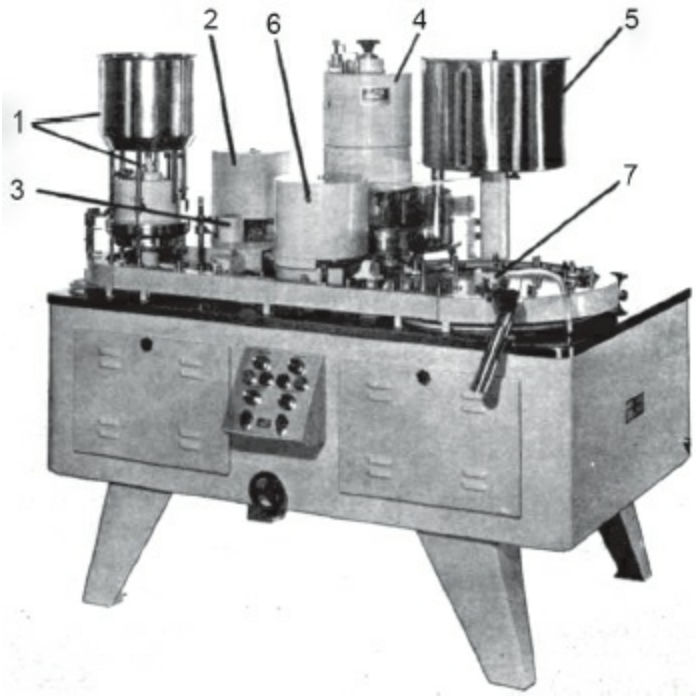



**Fig. 14.10:** MG2 model G36 capsule filling machine, 1. Empty capsule hopper and rectifier, 2. Cap holder removal station, 3. Cleaning station, 4. Powder dosing head, 5. Bulk powder hopper, 6. Cap holder replacing station, 7. Capsule closing and ejection station (*Courtesy* of mG2 macchine automatiche, Bologna, Italy)

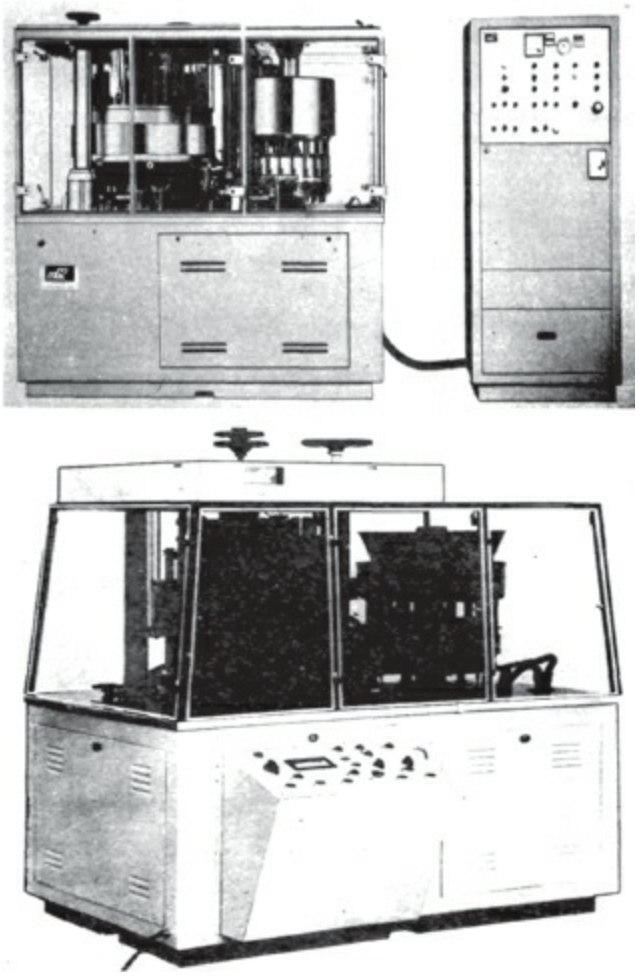


**Fig. 14.11:** mG2 Model G37N and model G38 capsule filling machine (*Courtesy* of mG2 macchine automatiche, Bologna, Italy)

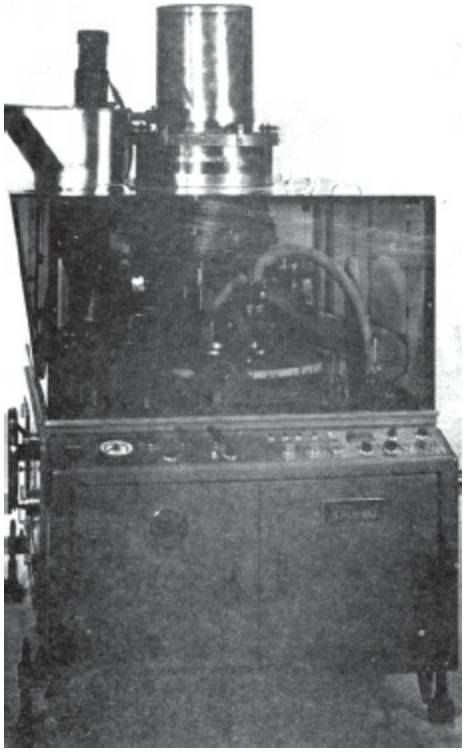


**Fig. 14.12:** Osaka model 180 capsule filling machine (*Courtesy* of Sharpies-stokes, Warminster, PA)

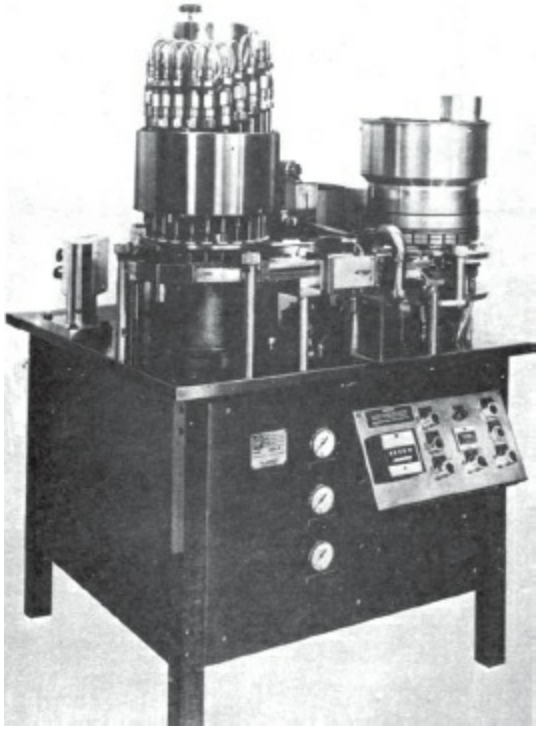


**Fig. 14.13:** Perry model CF ACCOF1L capsule filling machine (*Courtesy* of

Perry Industries, Green bay, WI)

Pelletized or granular materials may be readily filled using this equipment. Since filling is best accomplished without an auger in these cases, minimal change in fill weight can be achieved by alteration of rotational speed. It may also be desirable to perform the closing operations in a position other than the vertical position usually used for powder. In a vertical position, pellets or granules may escape from the body ring, and this may cause damage to the capsules.


**Fig. 14.14:** Zanasi model 6E capsule filling machine (*Courtesy* of Zanasi, SpA, Bologna, Italy)


**Fig. 14.15:** Zanasi model 40E capsule filling machine (*Courtesy* of Zanasi, SpA, Bologna, Italy)

## Liquid Filling

All the major machine manufacturers have made machines that can fill capsules with liquids. There are two types of liquid fills (a) non-aqueous solutions and (b) suspensions liquefied only for the filling process. Filling machines can handle materials with viscosities from 0.1 to 20 Pa. The uniformity of liquid fill is in most cases better than powder filling machine. The output of the liquid-filling machines is about 50–66% of the output of the same size powder-filling machine. The operational speed of liquid-filling machine is slower than powder filling machines, since liquid has to pass through a much smaller orifice than that for a powder. If the viscosity of the liquid is low or if the formulation is mobile at ambient temperatures, then the capsules will need to be sealed after filling.

## FORMULATION COMPONENTS

The problems encountered in handling powders during mixing and filling operations are so diverse as to preclude any but general comments. Although some problems are common to all types of filling equipment, certain machines themselves represent unique situations. Among the general problems, two major ones are listed.

1. After the powder ingredients have been homogeneously blended by any suitable technique, the flow of the resultant mixture must be adequate to ensure delivery of sufficient powder to the capsules at the time of filling. De-mixing must not occur during the powder handling in the filling equipment itself.
2. Physical incompatibilities between active ingredients, between diluents, or between active ingredients and/or diluents and the capsule shell may create problems.

The capsule seldom contains only the active ingredient(s); most capsule formulations require the use of some diluent material. Because of the wide range of materials encapsulated, no attempt can be made to outline specific criteria for the choice of suitable diluents.

The following are three major general considerations.

1. The powder mix must provide the type of flow characteristics required by the equipment. In the case of powder filling machines such as Lilly, Parke-Davis, Hofliger and Karg, Osaka and Perry, powder mix must the free flowing. In the case of slug filling machines such as Zanasi, Macofar, Farmatic and mG2, the powder must have sufficient cohesiveness to retain it slug form during delivery to the capsules. For example, while filling acetylsalicylic acid, an excipient such as a flowable cornstarch allows powder filling whereas excipients such as microcrystalline cellulose are required for filling slugs. In all cases, the powder mixture must retain its homogeneous composition without de-mixing during the machine handling operations. Lubricants, such as a metallic stearate, may be used in the former case; binders, such as mineral oil, are sometimes used in the latter. Particle sizes and powder densities of all ingredients should be matched as closely as possible to assist in the prevention of de-mixing.

2. Potential incompatibilities should be anticipated with each new mixture of materials. Reactions at elevated temperatures and humidities should be studied for effects, not only on the contained powder mixture, but also on the gelatin capsules. Studies such as these should include an evaluation in the presence of probable packaging materials. Evaluation of any test procedure should be based on sound statistical techniques.
3. The choice of excipients should be made with a view toward current Food and Drug Administration (FDA) regulations as they apply to Investigational New Drug and New Drug applications. Any applicable foreign regulations also should be considered. Some materials that may be useful as excipients are bentonite, calcium carbonate, lactose, mannitol, magnesium carbonate, magnesium oxide, silica gel, starch, talc, and tapioca powder.

If it is desirable for any reason to consider materials other than the aforementioned, first consideration should be given to materials that are given a “Generally Recognized As Safe” designation by the FDA. Obtaining approval of materials that are not in this category can be an expensive and time-consuming process, although there are occasions when it cannot be avoided.

## Diluents

The determination of amounts of diluents to be used is based on (1) the total amount of material that can possibly be put in the capsule in relation to the amount of active ingredients to be supplied by the capsule and (2) the amounts of lubricant and/or oil (generally in the order of 2% or less) that can be used. Experimentation with the actual materials is the only positive way to arrive at these figures.

## Glidants/lubricants

Materials that may be considered for improvement of flow characteristics may include the following: glycol esters, silicones, silicon dioxide, metallic stearates, stearic acid, and talc.

## Anti-dusting

Oils that may be considered for use in assisting in the control of dusting, as well as in providing additional cohesiveness to a powder mix, may include any inert, edible, FDA-approved material.

## CAPSULE FILL CAPACITY

Empty capsules are sold by sizes. The ones most commonly employed for human use range from size 0, the largest, to size 5, the smallest. Size 00 capsules may occasionally be required because of the volume of material to be filled, but this size is not used commercially in large volume. The fill capacity of a hard capsule is dependent upon the physical size of the capsule and the density of the fill material (Table 14.5). The fill weight for powders is calculated by multiplying powder tapped density by the capsule volume as provided by the capsule manufacturers. The relationship gives a reasonably accurate forecast for machine filling and holds because of the machine-dosing mechanisms. The fill volume varies with the amount of pressure employed in hand filling, or with the type of equipment utilized in machine filling. The size of the particles is also important because of the increase in voidage caused by large particles. The smaller the capsule size, the smaller should be the corresponding size of the particles to achieve uniform fill weights. The liquid-fill capacity of capsules is restricted and should not exceed 90% of the capsule body volume so as to prevent spillage of product.

**Table 14.5:** Fill capacity of hard gelatin capsule

| Capsule size | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| Weight (mg) | 96 ± 6 | 76 ± 5 | 61 ± 4 | 48 ± 3 | 38 ± 3 | 28 ± 2 |
| Fill density | | | Fill capacity (mg) | | | |
| 0.6 g/ml | 408 | 300 | 222 | 180 | 126 | 78 |
| 0.8 g/ml | 544 | 400 | 296 | 240 | 168 | 104 |
| 1.0 g/ml | 680 | 500 | 370 | 300 | 210 | 130 |
| 1.2 g/ml | 816 | 600 | 444 | 380 | 252 | 156 |

## FINISHING

Finished capsules from all filling equipment require some sort of dusting and/or polishing operation before the remaining operations of inspection, bottling, and labeling are completed. Dusting or polishing operations vary according to the type of filling equipment used, the type of powder used for filling, and the individual desires for the finished appearance of the completed capsules. The following are the methods most commonly used, based on desired output, formulation, required final appearance, and so on.

1. *Pan polishing:* Because of its unique design (primarily in the area of airflow), the Accela-Cota tablet coating pan may be used to dust and polish capsules. A polyurethane or cheese cloth liner is placed in the pan, and the liner is used to trap the removed dust as well as to impart a gloss to the capsules.
2. *Cloth dusting:* In this method, the bulk-filled capsules are rubbed with a cloth that may or may not be impregnated with an inert oil. This procedure is a hand operation, but one that can handle reasonable volumes, and that results in a positive method for removal of resistant materials. In addition, it imparts a somewhat improved gloss to the capsules.
3. *Brushing:* In this procedure, capsules are fed under rotating soft brushes, which serve to remove the dust from the capsule shell. This operation must be accompanied by application of vacuum for dust removal. Some materials are extremely difficult to remove by brushing, even to the point of impregnating the brushes and causing scratches or deformation of the capsules.

Commercial capsule sort/polish equipment are available and some of the units, in addition to the Accela-Cota pan, are as follow:

***Rotosort:*** It is a new filled capsule sorting machine sold by Eli Lilly and Company (Fig. 14.16). It is a mechanical sorting device that removes loose powder, unfilled joined capsules, filled or unfilled bodies, and loose caps. It can handle up to 150,000 capsules per hour, and can run directly off a filling machine or be used separately.

**Fig. 14.16:** ROTOSORT capsule sorting machine (*Courtesy* of Eli Lilly and Co., Indianapolis, IN)

*Erweka KEA:* The Erweka KEA dedusting and polishing machine is designed to handle the output from any capsule filling machine (Fig. 14.17). It moves the capsules between soft plastic tassels against a perforated plastic sleeve, under vacuum. Any residual powder is removed by the vacuum.

**Fig. 14.17:** Erweka KEA capsule dedusting and polishing machine (*Courtesy* of Erweka-Apparatebau, Heusenstamm, West Germany)

*PM60:* Seidenader Equipment, Totowa, NJ, offers two units that may be used separately or may be combined in the finishing of filled gelatin capsules. A belt is available that presents capsules for visual inspection, and it may include a vacuum system to automatically remove unfilled capsules. Cleaning and polishing machine PM60 (Fig. 14.18) may be used to polish finished capsules. It consists of two lamb's wool belts moving in opposite directions. The capsules are carried on the lower belt, and both belts are under suction.

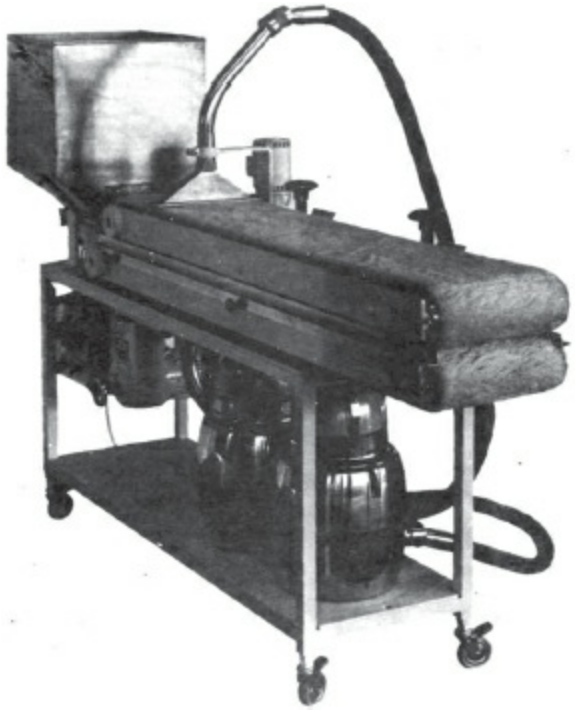


**Fig. 14.18:** Seidenader model PM60 capsule polishing machine (*Courtesy* of Seidenader Maschinenbau Miinchen, West Germany)

## IMPRINTING

Imprinting is a convenient method by which company and/or product identification information can be placed upon each capsule. The imprinting operation is best performed on the empty capsules, although filled capsules can be printed. The preference for imprinting empty capsules arises from the fact that the imprinting operation may occasionally damage some capsules. When filled capsules are imprinted, contamination, poor print quality, and actual damage to the imprinting equipment result. Various types and capacities of equipment are commercially available for this purpose in the United States. The three major suppliers of this equipment are Ackley Machine Corporation, Moorestown, NJ, RW Hartnett Company, Philadelphia, PA; and the Markem Machine Company, Keene, NH.

Hartnett offers a variety of machines with outputs as high as 500,000 capsules per hour (model B, Fig. 14.19). Also available is a unit that prints around the circumference of the capsules, as opposed to a longitudinal imprint; however, this machine operates at a slower rate. A lower-capacity unit (up to 250,000 capsules per hour) allows printing on both sides of the capsule, in different colors if desired.

Markem offers three models, which range from approximately 60,000 to 250,000 capsules per hour (Model 280A, Fig. 14.20). All three models allow for two-sided printing, but not circumferential. Ackley offers a straight-line imprinter with an output rate of about 500,000’capsules per hour, and has recently announced a new circumferential printer rated at about the same output. All imprinting machines operate on a rotogravure process, and a wide variety of colors of edible inks, both water and solvent-based, are commercially available.

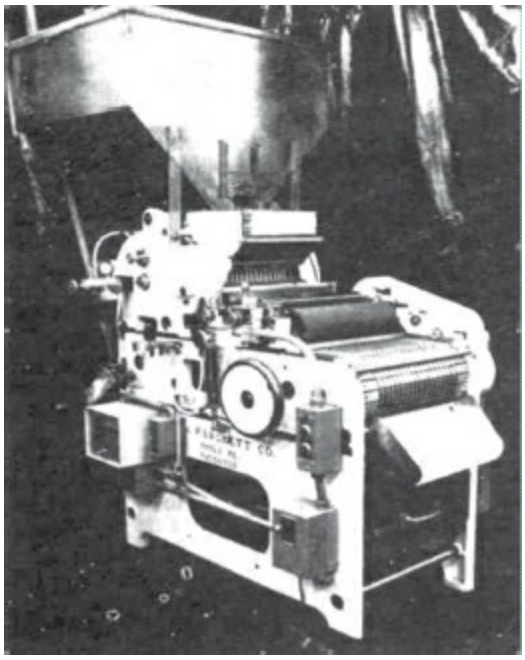


**Fig. 14.19:** Hartnett model B capsule imprinting machine (*Courtesy* of RH Hartnett Co., Philadelphia, PA)

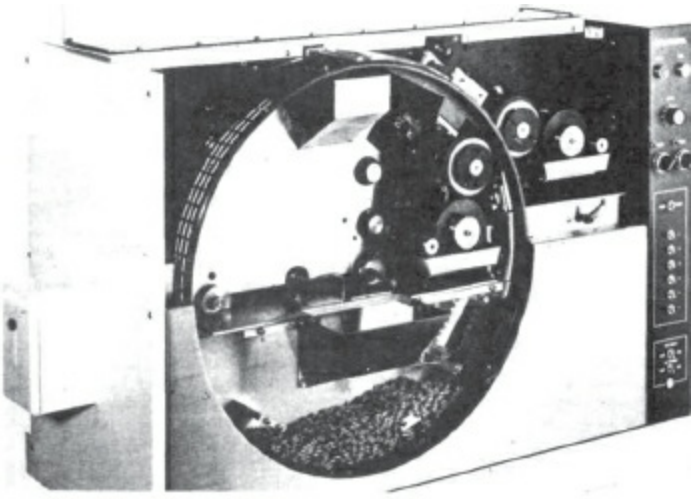


**Fig. 14.20:** Markem model 280A capsule imprinting machine (*Coursety* of Markem Co., Keene, NH)

## EVALUATION/QUALITY CONTROL TEST

Serious consideration should be given to the choice of suitable control procedures for the filling operations. In addition, it may be desirable to provide 100% weight checking after filling. The control procedures must ensure that the finished, filled capsules meet the appropriate current regulatory tests, e.g., weight variation, content uniformity, solubility, and/or disintegration. Current legal requirements should be adequately explored due to wide differences between various countries.

## Weight Variation Test

The weight variation test denned by USP XX is a sequential test, in which 20 intact capsules are individually weighed and the average weight is determined. The test requirements are met if none of the individual weights are less than 90%, or more than 110%, of the average. If the original 20 do not meet these criteria, the individual *net weights* are determined. These are averaged, and differences are determined between each individual net content and the average. The test requirements are met (1) if not more than two of the individual differences are greater than 10% of the average, or (2) if in no case any difference is greater than 25%.

If more than 2 but less than 6 net weights determined by the test deviate by more than 10% but less than 25%, the net contents are determined for an additional 40 capsules, and the average is calculated for the entire 60 capsules. Sixty deviations from the new average are calculated. The requirements are met (1) if the difference does not exceed 10% of the average in more than 6 of the 60 capsules and (2) if in no case any difference exceeds 25%.

Two new pieces of equipment, Rotoweigh and Vericap 1200, determines the weight of individual capsules, providing the automatic rejection of overfilled and underfilled capsules. These machines may be used in-line to reclaim portions of a batch as it is processed, or off-line to weigh and sort complete batch that has been shown statistically to have unacceptable weight variation.

*Rotoweigh:* The ROTOWEIGH is a highspeed capsule weighing machine sold by Eli Lilly and Company (Fig. 14.21). The capsules are gravity-fed onto vacuum pins for presentation to a unique weight detection system, which measures the reflected energy (backscatter) of a low power X-ray beam directed at each capsule. This reflected energy is proportional to the weight of the filled capsule, permitting automatic rejection of any individual capsule above or below preset weights. The machine operates at 73,000 capsules per hour, and its accuracy is more than adequate to assure compliance with the USP weight requirements.

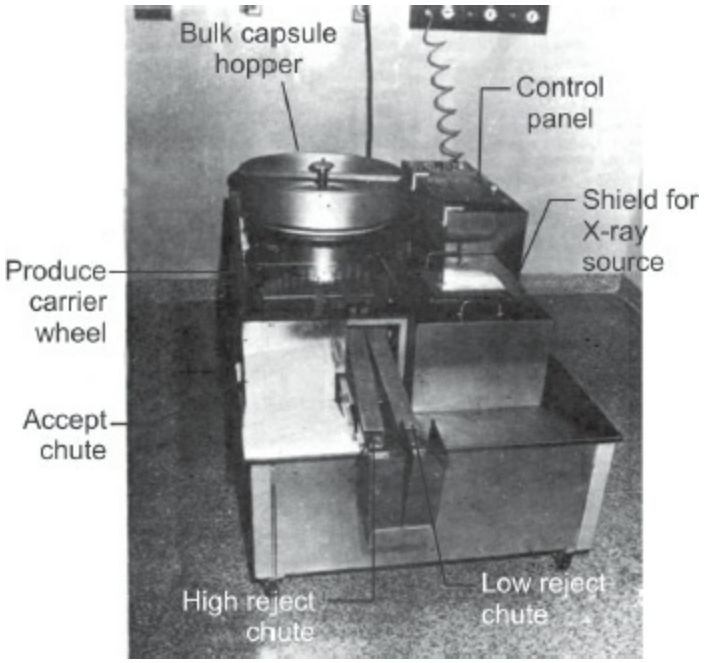



**Fig. 14.21:** ROTOWEIGH capsule weighing machine (*Courtesy* of Eli Lilly & Co., Indianapolis, IN)

*Vericap 1200:* The second unit is the Vericap 1200 machine (Fig. 14.22), which is sold by Modern Controls, Inc., Elk River, MN. It operates by detecting capacitance variation as filled capsules are propelled at high speed by compressed air between two charged plates. The measured change in dielectric constant thus produced is correlated to the weight of the capsule. Capsules that are overweight or underweight are then automatically separated from the acceptable capsules. The machine operates at a rate of 73,000 capsules per hour.

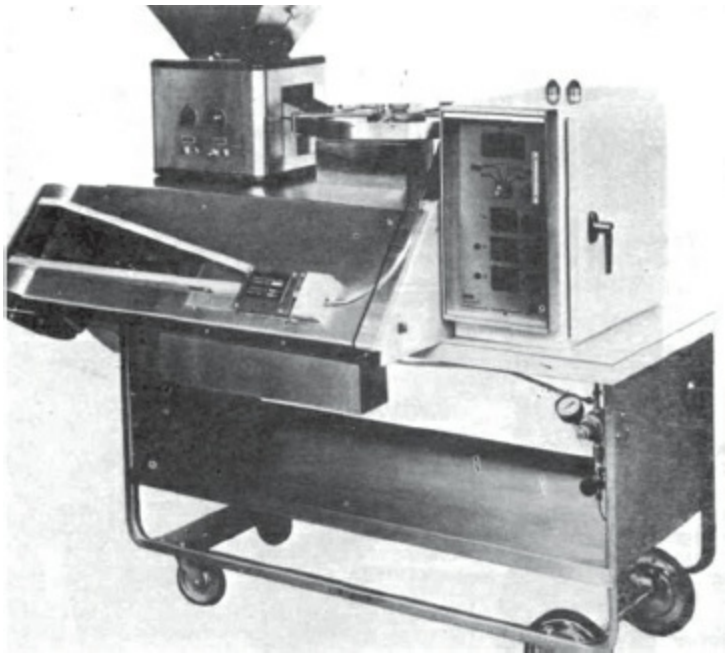


**Fig. 14.22:** Vericap 1200 capsule weighing machine (*Courtesy* of Modern Controls, Elk River, MN)

## Content Uniformity

A second test in USP XX that may apply to capsules is-that for content uniformity, which is performed as specified in individual monographs. In this case 30 capsules are selected, 10 of which are assayed by the specified procedure. The requirements are met if 9 of the 10 are within the specified potency range of 85 to 115%, and the tenth is not outside 75 to 125%

If more than 1, but less than 3, of the first 10 capsules fall outside the 85 to 115% limits, the remaining 20 are assayed. The requirements are met if all 30 capsules are within 75 to 125% of the specified potency range, and not less than 27 of the 30 are within the 85 to 115% range.

Broad generalizations about stability test programs cannot be made, since the question has to be answered according to each user's criteria. The tests should be based on adequate statistical design and should include evaluation of not only active ingredient stability, but also visual and mechanical aspects of the finished dosage form. These tests must include extended storage at various elevated temperatures with different humidity levels. Tests should include the filled capsules, both by themselves and in the presence of all contemplated packaging materials.

## In Vitro Testing

Pharmacopoeias require that hard capsules should be tested in the same apparatus as tablets even though they have very different physical properties. Filled capsules contain entrapped air, and most formulations will float on water. Devices are required to ensure that they sink as these can influence the results obtained. The way in which capsules disintegrate and dissolve is dependent upon several factors such as temperature and nature of the test media. The literature makes reference to the hard capsule effect; however, the literature shows that the rate-controlling step is the nature of the contents and not the shell.

When capsules are placed in an aqueous solution at body temperature (37°C), the capsule shell absorb water and swell. The rate of penetration of water is inversely proportional to the thickness of the wall. In gelatin capsules, water droplets can be observed on the inside surface of the shell after 30–40 sec. The shell ruptures first at the shoulders of the cap and body, which is the thinnest part. The rate of gelatin solubility is dependent upon the temperature of the solution. There is a significant decrease as the temperature falls below 30°C, and below about 26°C gelatin is completely insoluble and merely swell and distort. The rotating paddle method is the most frequently prescribed apparatus for measuring the dissolution rate of products in hard capsules. The test is used for manufacturing control purposes and for assessing product stability.

When gelatin capsules are stored under ICH-accelerated storage conditions (45°C, 75% RH), their solubility in water decreases with time. This is owing to the formation of a “pellicle” that slows down release. This effect is called cross-linking and can be caused either by interaction between gelatin and compounds containing reactive groups such as an aldehyde or by re-orientation of the gelatin molecules to a more collagen-like structure. For such stressed capsules the US Pharmacopeia recommend a two-tier dissolution test. First the dissolution of capsule is performed in the required medium and if the sample does not comply, then the test can be repeated by adding enzymes (pepsin and pancreatin) in the medium.

## SPECIAL TECHNIQUES

1. Special purpose capsules are capsules to which a special treatment has been given in an attempt to retard the solubility in some manner. This may be done in an attempt to delay absorption of the active ingredient, or to provide enteric properties. Normal solubility for gelatin capsules, either empty or filled, is not defined by the USP XX. However, the General Service Administration, in Federal Specification #U-C-115b (2/10/58), defines solubility limits for empty capsules as follows:

   (a) water resistance—fails to dissolve in water at 20 to 30°C in 15 min;

   (b) acid solubility—dissolves in less than 5 min in 0.5% aqueous HCl (w/w) at 36 to 38°C.

a. Formalin treatment has been employed to modify the solubility of gelatin capsules. Exposure to formalin vapors or treatment with aqueous formalin results in an unpredictable decrease in solubility of the gelatin film, owing to cross-linkage of the gelatin molecule initiated by the aldehyde. This result may also be noted if the product being filled contains aldehydic materials, or if aldehyde flavorants are added. Because of the nature of the reaction initiated in this manner, it is difficult to control the degree of insolubilization, or indeed, to prevent ultimate complete insolubility.

b. Various coatings have been used in an effort to provide similarly modified solubility characteristics. These coatings include salol, shellac, cellulose acetate phthalate, and certain resins that have usually been applied by usual pan coating techniques. Gelatin capsules do, however, provide a convenient way to deliver pellets or granular material when delayed or prolonged release properties have been incorporated in all or portions of the material to be filled.

2. Separation of incompatible materials (a technique used for some commercial products) is carried out by the use of a two-phase fill in the capsule. One phase consists of either soft capsule, smaller hard capsule, pill, or a suitably coated tablet that is filled into each capsule. Following this as a second phase, a powder fill is added in the usual manner. If this technique is used on commercial filling equipment, modifications must be made to the filling cycle of the machine. These changes would include, at minimum, the necessary changes in the machine operation to allow

materials to be loaded at two points during the filling cycle. Tamp type powder filling machines require the disabling of the tamp cycle.

3. Recently, there has been a revival of interest in the filling of conventional two-piece gelatin capsules with semisolids. Hard gelatin capsules were commonly used as early as the 1890s for oils, ethereal extracts, and pill masses, but the ability to fill the capsules on semiautomatic and automatic equipment is a recent development. The formulations used for filling are usually semisolids at ambient temperatures, which are melted to allow filling or they are thixotropic formulations in which the shear developed in filling allows pumping, but whose high viscosity when shear is absent prevents leakage after filling. Quantitative assessment of the gastric emptying of hard gelatin capsules filled with thixotropic liquids can be made in terms of the lag time prior to emptying, and the slope of the first order emptying curve. Results have shown that the viscosity of the fill has no significant influence on the emptying characteristics of these dosage forms. Machines for filling semisolid materials are currently available from Robert Bosch GmbH, Elanco, Harro Hofliger, and Zanasi.
4. A recent series of developments have resulted in allowing the use of, a liquid fill into two-piece hard capsules. The fills are either thermosetting or thixotropic. For liquids which are mobile, *gelatin banding* is the most successful method for the sealing of capsules to prevent the leakage. Two bands of gelatin solution are applied around the center of the filled capsule (HicapsealTM 40/100, Shionogi Qualicaps Inc.) and this is then dried using air at ambient conditions to prevent moisture loss from the gelatin shells. This band complies with the requirements of the Food and Drug Administration (FDA) “Tamper-Evident Packaging Requirements for Over-the-Counter Human Drug Products” for tamper-evident sealed capsules.

## ALTERNATIVE CAPSULES

### Hydroxypropyl Methylcellulose (HPMC) Capsules

HPMC capsules are a new type of hard two piece capsules developed as an alternative to hard gelatin capsules. HPMC capsules are odourless and flexible and their appearance corresponds to that of gelatin capsules, except that surface of HPMC capsules is matt. The physical properties of HPMC capsules compared to gelatin capsules are presented in Table 12.6. HPMC capsules have several advantages including low moisture content (2–5%), chemically inert and stable under very low moisture conditions. In addition, HPMC is a plant derived material, whereas gelatin is of swine or bovine (animal) origin. This eliminates the problem related to religious and vegetarian dietary restrictions.

HPMC capsules are manufactured by same dipping method employed for hard gelatin capsule. Shaped pins are dipped into an aqueous solution comprising 18–28% w/w HPMC having a viscosity of 2.4–5.4 × $10^{-6}$ $m^2/s$ (measured as a 2% aqueous solution at 20°C), 0.01–0.09% w/w carrageenan as a gelling agent and 0.05–0.6% w/w potassium or calcium ion as a co-gelling agent. Carrageenan and ions are added to facilitate gelling of HPMC at 48–55°C. After dipping the HPMC film is gelled, dried, trimmed and removed from the pins. The body and cap pieces are then joined. The finished HPMC capsules comprise 79.6–98.7% w/w HPMC, 0.03–0.5% w/w of carrageenan, 0.14–3.19% w/w of potassium or calcium ion and 2–5% w/w of water. HPMC capsules have been developed by Shionogi Qualicaps S.A. (HPMC), Capsugel (Vcaps™) Natural Capsules Ltd. (Cellulose capsules) and Associated capsules Ltd. (Naturecaps).

## STARCH CAPSULES

The starch capsules are made of potato starch and represents direct alternative to hard gelatin capsules. Compared to gelatin capsules, starch capsules feature several advantages: their dissolution is pH independent, they are suitable for enteric coating, they are stable due to bound moisture and capsules are tamper evident, preservative free and produced from non animal derived material. The starch capsule is odour-less, rigid and exhibits dissolution behaviour similar to the gelatin capsules. The moisture content of starch capsule ranges between 12 and 14 % w/w, with more than 30% being tightly bound. The presence of bound moisture suggests that starch capsules may provide better stability. Capsules prepared from starch are officially recognized in USPXXIII.

Recent advances in injection moulding technology have permitted the manufacture of starch capsules. During the production process, starch in the form of powder, pellet or granules is fed through a hopper onto a rotating reciprocating screw. The feed material moves along the screw towards the tip. Temperature is then increased by means of external heaters. From the feed zone to the compression zone, the feed material is gradually melted down. It is then conveyed through the metering zone, where homogenization of the melt occurs, to the end of the reciprocating screw. When sufficient melt is collected, it is injected into the mould. The polymer in the mold cools sufficiently resulting in opening of the mold and ejection of the molded parts. Temperature of between 120–180°C and pressure of about 700–2000 bars are normally seen in the operations. The West Pharmaceutical Services utilize injection moulding technique for the manufacture of starch capsules (Capill®). Recently it has been shown that enteric coated starch capsules (TARGIT®) are able to deliver formulations to the terminal ileum or colon successfully.

## CROSSLINKED DEXTRAN CAPSULES

Glutaraldehyde crosslinked dextran has also been evaluated as material for capsule shell. A reaction mixture containing dextran, $MgCl_2$, glutaraldehyde and polyethyleneglycol 400 in water was applied onto molding pins of nylon producing capsule caps and bodies. The capsule material is characterized by measuring the mechanical strength in compression and equilibrium degree of swelling. The model drug hydrocortisone was selected to study the application of crosslinked dextran as a capsule material for colon-specific drug delivery. The release was found to be about 10% during the initial 3 h in a buffer solution. Over a period of 24 h the release was about 35%. However, when the dextran capsules were challenged with a dextranase solution, simulating the arrival of the drug delivery system to the colon, the capsules quickly broke and the drug was rapidly released.

# SOFT GELATIN CAPSULES

Many pharmaceutical companies have the equipment and facilities for the development and production of tablets, liquids, and hardshell capsule products, but they usually depend upon custom manufacturers for the development and production of soft gelatin capsules. The custom manufacturers are specialists in this field, owing primarily to economic, patent, and technologic factors. Although few become directly involved in the manufacture of soft capsules, pharmaceutical chemists must be prepared to investigate this dosage form and to participate in its development, either in their own laboratories or in cooperation with the technical personnel of a custom manufacturer.

The pharmaceutical chemist should be cognizant of the inherent properties of soft gelatin capsules. Essentially, these capsules are solid dosage forms containing liquid medication and therefore offer certain advantages:

1. They permit liquid medications to become easily portable.
2. Accuracy and uniformity of dosage, capsule to capsule and lot to lot, are predominant advantages. These capsules easily pass the appropriate compendial tests and surpass other solid dosage forms in this respect, because liquid formulations can be more accurately and precisely compounded, blended, homogenized, and measured or dispensed than can dry solid formulations.
3. The pharmaceutical availability of drugs formulated for this dosage form, as measured by disintegration time, or by dissolution rate, often shows an advantage over other solid dosage formulations.
4. The physiologic availability of drugs is often improved since these capsules contain the drug in liquid form, i.e., as a liquid drug substance, drug in solution, or drug in suspension.
5. Better tamper evidence (tampering leads to puncturing and visible leakage).
6. Safer handling of highly potent or cytotoxic drug compounds.

Owing to their special properties and advantages, soft gelatin capsules are

used in a wide variety of industries, but they are used most widely in the pharmaceutical industry. Billions of capsules are made each year in various sizes and shapes (Fig. 14.23), and in a variety of colors and color combinations. Their pharmaceutical applications are:

1. As an oral dosage form of ethical or proprietary products for human or veterinary use.
2. As a suppository dosage form for rectal use, or for vaginal use. Rectal dosage forms are becoming more acceptable for pediatric and geriatric use. Vaginal use is confined to applications that require the medication to be inserted at bedtime. Because of the action of the sphincter muscle, rectal use is not similarly limited.
3. As a specialty package in tube form, for human and veterinary single dose application of topical, ophthalmic, and otic preparations, and rectal ointments.
4. In the cosmetic industry, these capsules may be used as a specialty package for breath fresheners, perfumes, bath oils, suntan oils, and various skin creams.

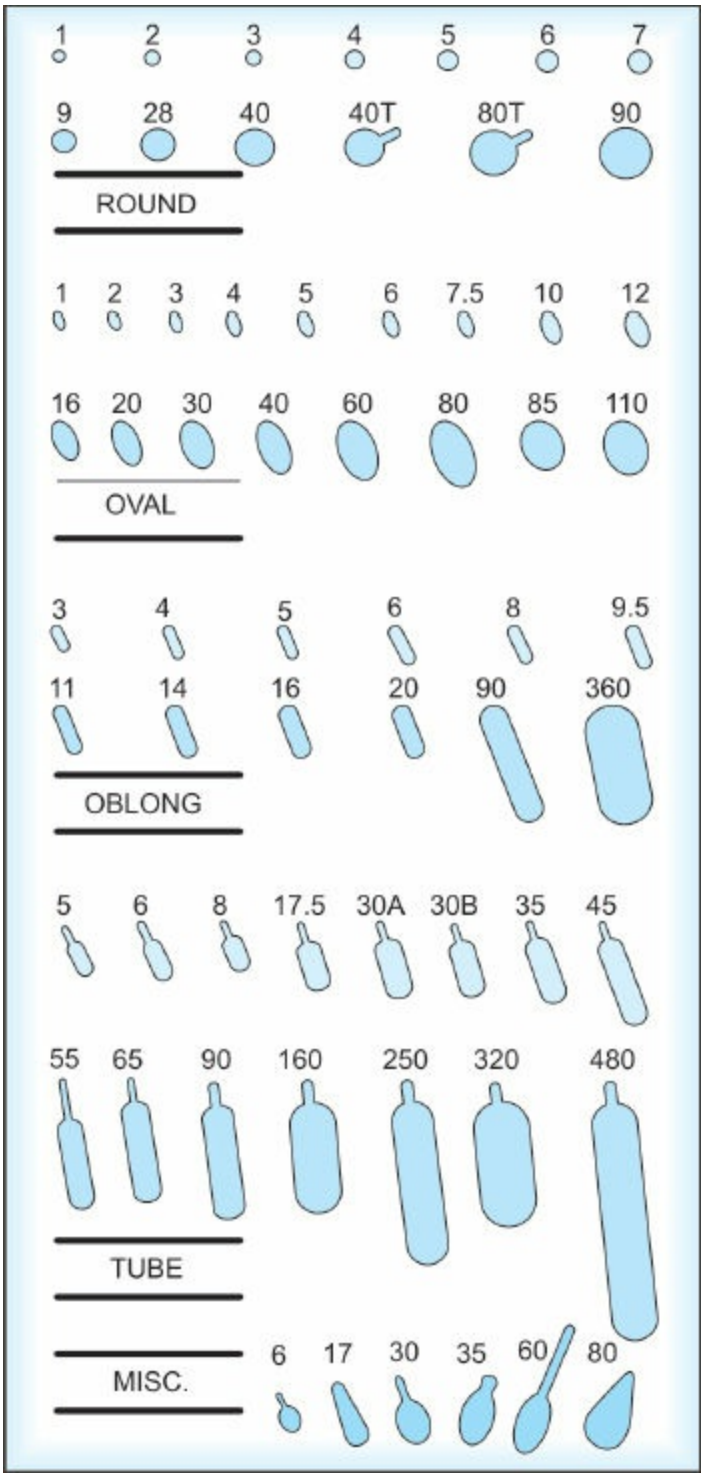



**Fig. 14.23:** Sizes and shapes of soft gelatin capsules (1 cc = 16.23 m). Numbers represent the nominal capacity in minims

# CAPSULE MANUFACTURE

## Capsule Content

Soft gelatin capsule is a hermetically sealed, one-piece capsule shell, used to dispense a variety of liquids and solids. Requirements and specifications of these materials vary, depending on the equipment of the manufacturer, but there are basic precepts that may be used as a guide for the formulation and production of commercially and therapeutically acceptable capsules, regardless of method of capsulation. The formulation of the capsule content for each product is individually developed to fulfill the specifications and end-use requirements of the product.

Except for the Accogel process, which is primarily concerned with the capsulation of dry powders, the content of a soft gelatin capsule is a liquid, or a combination of miscible liquids, a solution of a solid(s) in a liquid(s), or a suspension of a solid(s) in a liquid(s). All such materials for capsulation are formulated to produce the smallest possible capsule consistent with maximum ingredient and physical stability, therapeutic effectiveness, and production efficiency. Once the smallest capsule size is determined, personnel in the sales or marketing departments usually choose the color, shape, and ultimate size of the retail product, unless there is a technical or production reason for the development chemist to specify a particular size, shape, and color. The maximum capsule size or shape for convenient oral use in humans is the 20 minim oblong, the 16 minim oval, or the 9 minim round (*see* Fig. 14.23).

***Liquids*** are an essential part of the capsule content. Only those liquids that are both water-miscible and volatile cannot be included as major constituents of the capsule content since they can migrate into the hydrophilic gelatin shell and volatilize from its surface. Water, ethyl alcohol, and, of course, emulsions fall into this category. Similarly, gelatin plasticizers such as glycerin and propylene glycol cannot be major constituents of the capsule content, owing to their softening effect on the gelatin shell, which thereby makes the capsule more susceptible to the effects of heat and humidity. As minor constituents (up to about 5% of the capsule content), water and alcohol can be used as cosolvents to aid in the preparation of solutions for capsulation. Also, up to 10% glycerin and/or propylene glycol can be used as cosolvents with polyethylene glycol or other liquids that have a

shellhardening effect when capsulated alone.

There are a large number of liquids that do not fall into the foregoing category and thus can function as active ingredients, solvents, or vehicles for suspension-type formulations. These liquids include aromatic and aliphatic hydrocarbons, chlorinated hydrocarbons, high-molecular-weight alcohols, esters, and organic acids. Many of these are used in veterinary, cosmetic, and industrial products. For human use, however, the pharmaceutical chemist is often limited in his selection or use of a particular liquid because of government regulations, product performance specifications, ingredient incompatibilities, and liquid-solid adsorption characteristics. The most widely used liquids for human use are oily active ingredients (clofibrate), vegetable oils (soybean oil), mineral oil, nonionic surface active agents (polysorbate 80), and polyethylene glycols (400 and 600), either alone or in combination. Such active ingredient oils as fish oil may also function as a solvent, or as the suspending medium for one or more additional active ingredients, as in vitamin capsules.

All liquids, solutions, and suspensions for capsulation should be homogeneous and air-free (vide infra), and preferably should flow by gravity at room temperature, but not at a temperature exceeding 35°C at the point of capsulation, since the sealing temperature of the gelatin films is usually in the range of 37 to 40°C. In general, liquids ranging in *viscosity* from ethyl ether (0.222 cp at 25°C)[9] to heavy adhesive mixtures (exceeding 3000 cp at 25°C) may be encapsulated, but there are some exceptions since the property of viscosity alone is not the sole criterion. Liquids that exhibit the rheologic property of tack or tackiness, such as glycerin (954 cp at 25°C),[9] are exceptions, since such liquids can eventually cause the binding of slide valves and pumps in the capsule filling mechanism. Also, preparations for encapsulation should have a pH between 2.5 and 7.5, since preparations that are more acidic can cause hydrolysis and leakage of the gelatin shell, and preparations that are more alkaline can tan the gelatin and thus affect the solubility of the shell.

The capsulation of water-immiscible liquids is the simplest form of soft gelatin capsulation and usually requires little or no formulation. The minimum size capsule depends on the dosage desired, the minimum fill volume being calculated from the specific gravity of the liquid. A die size and shape may then be chosen from those shown in Fig. 14.23. The nearest die

size above the calculated fill volume may be used, or any larger die may be chosen if the active ingredient is to be diluted for some reason. For example, a 25,000-unit vitamin A capsule using vitamin A palmitate (1,000,000 units A/g) as a source for the vitamin A would have a minimum fill volume of about 0.45 minims, and thus could be diluted to any desired capsule size. On the other hand, the same potency capsule using fish oil (50,000 units A/g) as a source for the vitamin A would have a minimum fill volume of about 8.8 minims.

The minimum fill volume for water-miscible, nonvolatile liquids, such as poly-sorbate 80, is determined in the same manner. However, because of their hygroscopic nature, they cause water to migrate from the gelatin shell into the fill material. This migration is rapid and could amount to 20% of the weight of the miscible liquid. During the drying period of the capsule, most of this water returns to and passes through the gelatin shell, but up to 7.5% of the original water can remain in the fill material, depending on the hydrophilic properties of the liquid. Thus, for liquids of this type, a safety factor must be used in establishing the minimum fill volume and in choosing the die.*

Although oily liquids do not retain moisture, water does pass from the shell of the capsule into the fill material and out again during the manufacture and drying of these capsules. This is important for the formulator to remember, since such water transfer can and does have a bearing on formulations in which oily liquids are used as solvents or as vehicles for suspensions. If such suspensions contain hydrophilic solids, water may be retained up to 3% by weight of the hydrophilic material.

Combinations of miscible liquids often are used to produce desired physiological results such as increased or more rapid absorption of active ingredient (vitamin A and polysorbate 80); or to produce desired physiochemical results, such as improved flow properties (dilution or partial substitution with a thinner liquid), or improved solubility (steroid with oil and benzyl alcohol).

Except for when the Accogel process is used, solids are filled into soft gelatin capsules, in the form of either a solution or a suspension. The preparation of a suitable solution of a solid medicament should be the first goal of the pharmaceutical chemist. Usually, a solution is more easily capsulated and exhibits better uniformity, stability, and biopharmaceutical

properties than does a suspension. For oral products, the medicament must have sufficient solubility in the solvent system so that the necessary dose is contained in a maximum fill volume of 16 to 20 minims (1 to 1.25 cc).

***Solids*** that are not sufficiently soluble in liquids or in combinations of liquids are capsulated as suspensions. Most organic and inorganic solids or compounds may be capsulated. Such materials should be 80 mesh or finer in particle size, owing to certain close tolerances of the capsulation equipment and for maximum homogeneity of the suspension. Many compounds cannot be capsulated, owing to their solubility in water and thus their ability to affect the gelatin shell, unless they are minor constituents of a formula or are combined with a type of carrier (liquid or solid) that reduces their effect on the shell. Examples of such solids are strong acids (citric), strong alkalies (sodium salts of weak acids), salts of strong acids and bases (sodium chloride, choline chloride), and ammonium salts. Also, any substance that is unstable in the presence of moisture (e.g., aspirin) would not exhibit satisfactory chemical stability in soft gelatin capsules.

The capsulation of *suspensions* is the basis for the existence of a large group of products. Again, the design of suspension type formulations and the choice of the suspending medium are directed toward producing the smallest size capsule having the characteristics previously described, i.e. maximum production capacity consistent with maximum physical and ingredient stability and therapeutic efficacy.

The formulation of suspensions for capsulation follows the basic concepts of suspension technology. Formulation techniques, however, can vary depending on the drug substance, the desired flow characteristics, the physical or ingredient stability problems, or the biopharmaceutical properties desired. In most instances, these techniques must be developed through the ingenuity of the formulating chemist; however, in the formulation of suspensions for soft gelatin encapsulation, certain basic information must be developed to determine minimum capsule size.

## *Base Adsorption*

One laboratory tool for this purpose is known as the “base adsorption” of the solid(s) to be suspended. Base adsorption is expressed as the number of grams of liquid base required to produce a capsulatable mixture when mixed with one gram of solid(s). The base adsorption of a solid is influenced by

such factors as the solid's particle size and shape, its physical state (fibrous, amorphous, or crystalline), its density, its moisture content, and its oleophilic or hydrophilic nature.

In the determination of base adsorption, the solid(s) must be completely wetted by the liquid base. For glycol and nonionic type bases, the addition of a *wetting agent* is seldom required, but for vegetable oil bases, complete wetting of the solid(s) is not achieved without an additive. Soy lecithin, at a concentration of 2 to 3% by weight of the oil, serves excellently for this purpose, and being a natural product, is universally accepted for food and drug use. Increasing the concentration above 3% appears to have no added advantage.

A practical procedure for determining base adsorption and for judging the adequate fluidity of a mixture is as follows. Weigh a definite amount (40 g is convenient) of the solid into a 150 ml tared beaker. In a separate 150 ml tared beaker, place about 100 g of the liquid base. Add small increments of the base to the solid, and using a spatula, stir the base into the solid after each addition until the solid is thoroughly wetted and uniformly coated with the base. This should produce a mixture that has a soft ointment-like consistency. Continue to add liquid and stir until the mixture flows steadily from the spatula blade when held at a 45-degree angle above the mixture. The flow is even and continuous, and not in "globs." Attention should also be given to the nature of the "cut-off" quality of the mixture. As the mixture tends to stop flowing, proper cut-off is exhibited when the stream contracts rapidly upward toward the spatula blade rather than "stringing out" in intermediate flow.

At the conclusion of the foregoing test, the base adsorption is obtained by means of the following formula:

$$\text{Base adsorption} = \frac{\text{Weight of base}}{\text{Weight of solid}}$$

The base adsorption mixture is milled or homogenized, and deaerated (a desiccator under vacuum is suitable), and the specific gravity is taken. The specific gravity is the weight of mixture (W) per cubic centimeter or per 16.23 minims (V). As in the case of liquids and solutions, the specific gravity may be used to determine the die size required for a given quantity of the particular mixture.

The base adsorption is used to determine the "minim per gram" factor

(M/g) of the solid(s). The minim per gram factor is the volume in minims that is occupied by one gram (S) of the solid plus the weight of liquid base (BA) required to make a capsulatable mixture. The minim per gram factor is calculated by dividing the weight of base plus the gram of solid (BA + S) by the weight of mixture (W) per cubic centimeter or 16.23 minims (V). A convenient formula is:

$$\frac{(BA + S) \times V}{W} = M/g$$

Thus, the lower the base adsorption of the solid(s) and the higher the density of the mixture, the smaller the capsule will be. This also indicates the importance of establishing specifications for the control of those physical properties of a solid mentioned previously that can affect its base adsorption.

The BA and M/g data need not be obtained on any material that is to be capsulated alone at concentrations of 50 mg or less, since the smallest capsules can accommodate such quantities. If such material is to be used in combination, however, the data become necessary to allow for its inclusion in the formulation. The convenience of using M/g factors is particularly evident in the vitamin field, where there may be many ingredients and numerous combinations. Since the minim per gram factors are additive, they can be used for a more rapid calculation of capsule size than can be given by the preparation of the many possible mixtures in the laboratory. *See* Table 14.7 for BA and M/g data on some typical solids.

**Table 14.7:** BA and M/g factors of some typical solids

| Ingredient | Base* | BA | M/g |
|---|---|---|---|
| Acetaminophen | Veg. oil | 0.76 | 25.97 |
| Acetaminophen | PEG 400 | 0.75 | 23.07 |
| Ascorbic acid | Veg. oil | 0.60 | 20.60 |
| Ascorbic acid | Polysorbate 80 | 1.10 | 26.92 |
| $Al(OH)_3$—$MgC0_3$ (FMA 11) | Veg. oil | 1.90 | 41.30 |
| $Al(OH)_3$—$MgC0_3$ (FMA 11) | PEG 400 | 2.44 | 42.10 |
| Danthron | Veg. oil | 1.30 | 33.75 |
| Danthron | Glyceryl monooleate | 1.39 | 33.94 |
| Danthron | Polysorbate 80 | 1.38 | 31.28 |
| Danthron | PEG 400 | 1.60 | 33.62 |
| Danthron | Triacetin | 1.83 | 36.02 |
| Ephedrine $S0_4$ | Veg. oil | 1.30 | 36.80 |
| Ferrous $S0_4$ exsiccated | Veg. oil | 0.30 | 10.60 |
| Ferrous $S0_4$ exsiccated | Polysorbate 80 | 0.47 | 12.90 |
| Guaifenesin | Veg. oil | 1.28 | 34.68 |
| Lactose | Veg. oil | 0.75 | 23.87 |
| Desiccated liver | Veg. oil | 0.80 | 25.70 |
| Mephenesin | Veg. oil | 2.50 | 57.38 |
| Mephenesin | PEG 400 | 2.13 | 44.77 |
| Meprobamate | Veg. oil | 1.59 | 42.55 |
| Meprobamate | PEG 400 | 1.30 | 32.52 |
| Niacinamide | Veg. oil | 0.80 | 25.63 |
| Neomycin sulfate | Veg. oil | 0.60 | 20.66 |
| Phenobarbital | Veg. oil | 1.20 | 33.60 |
| Procaine penicillin G | Veg. oil | 0.91 | 28.63 |
| Sodium ascorbate | Veg. oil | 0.76 | 22.40 |
| Salicylamide | Veg. oil | 0.80 | 25.80 |
| Sulfathiazole | Veg. oil | 0.43 | 17.90 |
| Sulfanilamide | Veg. oil | 1.03 | 28.55 |
| Tetracycline (amphoteric) | Veg. oil | 0.61 | 21.63 |

* Vegetable oil bases contain 3% soy lecithin

The final formulation of a suspension invariably requires a *suspending agent* to prevent the settling of the solids and to maintain homogeneity prior to, during, and after capsulation. The nature and concentration of the suspending agent vary, depending on the job to be done. Also, a rather delicate balance must be achieved between the requirement for a stable suspension and the requirement for the mixture to have the proper flow characteristics. There is evidence, too, that the proper suspending agent coats the suspended solids, imparting a certain lubricity to them and thereby aiding capsulation. Also, the coating can prevent contact with possible incompatible components in the mixture. Of the examples shown in Table 14.8, the most widely used suspending agent for oily bases is wax mixture, and in nonoily

bases, the polyethylene glycols 4000 and 6000.

**Table 14.8:** Typical suspending agents

| Type | Concentration of oily base (%) | Type | Concentration of non-oily base (%) |
|---|---|---|---|
| White wax, NF | 5 | PEG 4000 and 6000 | 1–15 |
| Paraffin wax, NF | 5 | Solid nonionics | 10 |
| Animal stearates | 1–6 | Solid glycol esters | 10 |
| Wax mixture* | 10 and 30 | Acetylated monoglycerides | 5 |
| Aluminum monosterate† | 3–5 | | |
| Ethocel (100 cps)†† | 5–10 | | |

* 1 part hydrogenated soybean oil; 1 part yellow wax, +; 4 parts vegetable shortening (melting point 33 to 38°C); used at 10% on the adsorption oil and at 30% on any filler oil required ++ Used with volatile organic liquids such as butyl chloride; toluene; tetrachlorethylene; benzene.

In all instances, the suspending agent used is melted in a suitable portion of the liquid base, and the hot melt is added slowly, with stirring, into the bulk portion of the base, which has been preheated to 40°C prior to the addition of any solids. The solids are then added, one by one, with sufficient mixing between additions to ensure complete wetting. Incompatible solids are added as far apart as possible in the mixing order to prevent interaction prior to complete wetting by the base.

Additional aids to formulation involve the physical and ingredient stability of the capsules. There should be little concern with oxidation or the effects of light as a cause of ingredient instability, since the gelatin shell is an excellent oxygen barrier and may be opacified.

Most ingredient stability problems are associated with the available moisture from the gelatin shell, which when absorbed into the capsule content, can cause areas of high concentration of water-soluble solids, leading to ionization and interaction of the solids. Such problems may be alleviated or eliminated by employing a less soluble salt (procaine penicillin instead of potassium), employing coatings (gelatin-coated $B_{12}$), adjusting pH with appropriate small quantities of citric, lactic, or tartaric: acids or with less restrictive quantities of sodium ascorbate or magnesium oxide, or salting-put with appropriate small quantities of sodium chloride or sodium acetate.

Usually, the *physical stability* of a product is associated primarily with the type of gelatin and gelatin formulation used but can be aided by proper fill formulation. If a particular solid may have a deleterious action on the

gelatin shell, the form of the solid that is least water-soluble and the most oleophilic would be the form of choice for an oil-based suspension. An example would be the use of calcium salicylate rather than the sodium or magnesium salts. Also, the type of liquid base used can have an effect on physical stability. For example, the proteolytic effect of chloral hydrate on the gelatin shell is greatly reduced when a polyethylene glycol base is used in place of an oily base.

With the proper selection of materials and formulation techniques, the pharmaceutical chemist can prepare solutions or suspensions for comparisons of stability and dissolution rate with formulations of other solid dosage forms. By accurately filling two-piece gelatin capsules with such formulations, comparative absorption, urinary excretion, and metabolic studies can be made prior to the actual preparation of the soft gelatin capsule dosage form. Today the product development laboratory must evaluate all potential formulations for a new drug substance or for product improvement.

## Gelatin Processing

Except for the gelatin preparation department, the manufacturing areas of a typical plant are air-conditioned to assure the proper conditioning of the gelatin films, the proper drying of the capsules, and the consistent low moisture content of raw materials and mixtures. The temperature is usually in the range of 20 to 22°C, and the humidity is controlled to a maximum of 40% in the operating areas and a range of 20 to 30% in the drying areas.

In the *gelatin preparation* department of a typical manufacturer, the gelatin is weighed on printomatic scales and mixed with the accurately metered (printomatic) and chilled (7°C) liquid constituents in suitable equipment, such as a Pony Mixer. The resultant fluffy mass is transferred to melting tanks and melted under vacuum (29.5” Hg) at 93°C. The mixing process requires about 25 min for 270 kg of mass, and the melting procedure requires about 3 hours. A sample of the resulting fluid mass is visually compared with a color standard, and additional colorants are blended into the mass if adjustments are required. The mass is then maintained at a temperature of 57 to 60°C before and during the capsulation process.

The *materials preparation* department will nave a weigh-off and mixing area containing the necessary equipment and facilities for the preparation of the variety of mixtures that may be capsulated. Typical equipment would include printomatic scales for exacting measurements and control records; stainless-steel jacketed tanks for handling from 10-to 450-gallon batches of mix; and mixers, such as the Cowles, for the initial blending of solids with the liquid base. After the initial blending is completed, the mixture is put through a *milling* or *homogenizing* process, using equipment such as the homoloid mill, stone mill, hopper mill, or the Urschel Comitrol. The purpose of the milling operation is not to reduce particle size, but to break up agglomerates of solids and to make certain that all solids are “wet” with the liquid carrier, so as to achieve a smooth and homogenous mixture.

Following the milling operation, all mixtures are subjected to *deaeration*, and particularly so if the capsulation machine is equipped with a positive displacement pump. Deaeration is necessary to achieve uniform capsule fill weights; it also protects against loss of potency through oxidation prior to and during capsulation. When small amounts of volatile ingredients are included in a formulation, they are carefully added and blended into the bulk mixture

after deaeration. Most liquids and suspensions may be deaerated by means of equipment designed to expose thin layers of the material continuously to a vacuum (29.5” Hg) and at the same time transfer the material from the mixing tank to the container that will be used at the capsulation machine. Suspensions or liquid mixtures containing volatile liquids or liquid surface active agents as chief constituents of the formula may be deaerated by subjection to temperatures up to 60°C for the period required to achieve the results desired. After deaeration, the mixture is ready to be capsulated.

## MANUFACTURING EQUIPMENT

Soft gelatin capsules have been available since the middle of the nineteenth century. Originally, they were made one at a time; leather molds—and later, iron molds—were used for shaping the capsule. The capsules were filled by medicine dropper and sealed by hand with a “glob” of molten gelatin. Since those early days, many methods of capsulation have been proposed and patented, but this discussion is confined to equipment of commercial significance in present use. As technology advanced, the individual iron molds gave way to multiple molding units, and these eventually led to sets of plates containing the pockets.

## Plate Process

The plate process is the oldest commercial method of manufacture, but today this equipment can no longer be purchased, and consequently, only a few companies still use this process. The plate process—a batch process—involves pressing two sheets of wet gelatin together between two molds, provided with depressions. One of the gelatin sheet is placed over mold and application of vacuum produce depressions in the gelatin sheet into which active fill was then placed. A second gelatin sheet was laid over the first and both were pressed together with fill material sandwiched in between. The pressure of the plate sealed the top and bottom sheets of gelatin together. Softgels were produced by cutting followed by drying. The plate process that requires two or three operators for each machine has given way to the more modern continuous processes, which require considerably less manpower for operation.

## Rotary Die Process

The continuous processes became a commercial reality in 1933, when the late R. P. Scherer invented the rotary die process. Prior to this invention, soft gelatin capsules were not looked on favorably by the pharmaceutical industry, owing to the relatively large amount of the capsulated material (15 to 20%) lost during manufacture, and to the variation in the net content of the capsule (possibly 20 to 40%). The rotary die process reduced manufacturing losses to a negligible figure and content variation to less than ±3%.

A schematic drawing of the rotary die process is presented, however, to acquaint the pharmaceutical chemist with the fundamental aspects of capsulation (Fig. 14.24).

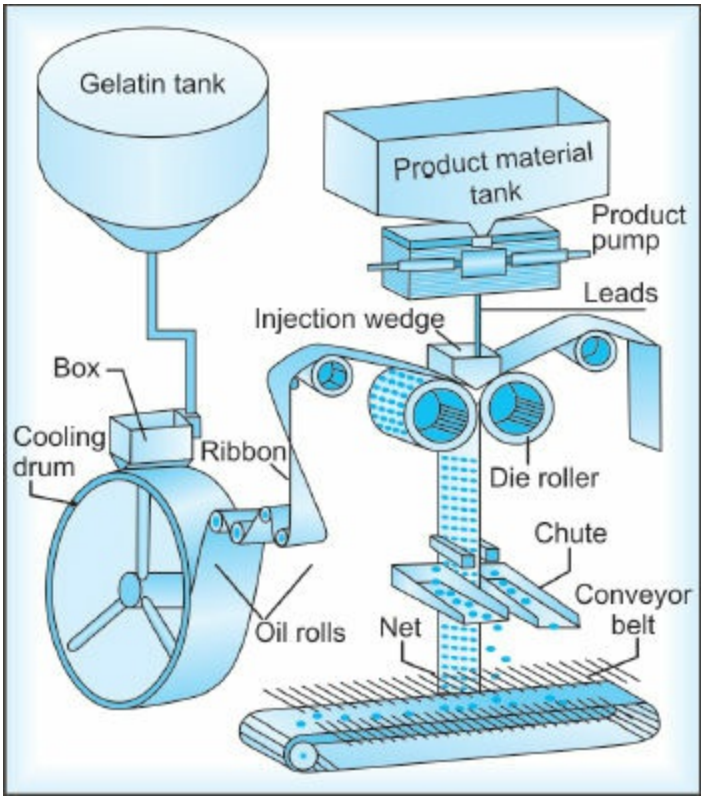



**Fig. 14.24:** Schematic drawing of rotary die process

The gelatin mass is fed by gravity to a metering device (spreader box), which controls the flow of the mass onto air-cooled (13 to 14°C) rotating drums. Gelatin ribbons of controlled (± 10%) thickness are formed. The wet shell thickness may vary from 0.022 to 0.045 inch, but for most capsules, it is

between 0.025 and 0.032 inch. Thicker shells are used on products requiring greater structural strength. Product cost is directly proportional to shell thickness. The ribbons are fed through a mineral oil lubricating bath, over guide rolls, and then down between the wedge and the die rolls.

The material to be capsulated flows by gravity into a positive displacement pump. The pump accurately meters the material through the leads and wedge and into the gelatin ribbons between the die rolls. The bottom of the wedge contains small orifices lined up with the die pockets of the die rolls. The capsule is about half sealed when the pressure of the pumped material forces the gelatin into the die pockets, where the capsules are simultaneously filled, shaped, hermetically sealed, and cut from the gelatin ribbon. The sealing of the Capsule is achieved by mechanical pressure on the die rolls and the heating (37 to 40°C) of the ribbons by the wedge.

Immediately after manufacture, the capsules are automatically conveyed through a naphtha wash unit to remove the mineral oil lubricant. The washed capsules may be automatically subjected to a preliminary infrared drying step, which removes 60 to 70% of the water which is to be lost, or may be manually spread directly on trays. Capsules from the infrared dryer are also spread on trays, and all capsules are allowed to come to equilibrium with forced air conditions of 20 to 30% relative humidity at 21 to 24°C. Capsules at equilibrium with 20 to 30% RH at 21 to 24°C are considered “dry,” and the shell of such a capsule contains 6 to 10% water, depending on the gelatin formula used.

## Reciprocating Die Process

The early success of the rotary die process led others to develop continuous methods of soft gelatin capsule manufacture. One such method, known as the *reciprocating die process*, was announced in 1949 and was developed by the Norton Company, Worchester, MA.

## Accogel Technology

Another continuous process, also announced in 1949, was developed by the Lederle Laboratories Division of the American Cyanamid Company and has been used solely in the manufacture of that company's products. This equipment, known as the *Accogel machine*, is unique in that it is the only equipment that accurately fills powdered dry solids into soft gelatin capsules.

A discussion of the comparative advantages and disadvantages of the foregoing four processes—plate, rotary die, reciprocating die, and Accogel machine—is beyond the scope of this chapter and would have little instructive value, since the pharmaceutical chemist seldom has the opportunity to choose between the four types of equipment. However, one must consider, that for maximum production efficiency, the continuous processes demand almost 24 hours per day, 5 (preferably 7) days per week, of continuous operation. Thus, medicament formulations must be so designed as to maintain their desired physical characteristics during this period of operation as well as during periods of weekend shutdowns. The production capacity of each of these machines is determined by (1) the size, which determines the number of the pockets on the standard-sized die plate, rotary die, or reciprocating die, (2) the speed of the machine (of the operators for the plate process) and (3) the physical characteristics of the material to be capsulated. Formulations are designed to achieve maximum production capacity consistent with maximum physical and ingredient stability and therapeutic efficacy

## QUALITY CONTROL

During manufacture, capsule samples are taken periodically for seal thickness and fill weight checks. The *seal thickness* is measured under a microscope, and changes in ribbon thickness, heat, or the pressure are made if necessary. Acceptable seal thickness is one half to two thirds of the ribbon thickness. *Fill weight checks* are made by weighing the whole fresh capsule, slitting it open, and expressing the contents. The shell is then washed in a suitable solvent (petroleum ether), and the empty shell is reweighed. If necessary, adjustments in the pump stroke can be made to obtain the proper fill weight.

The *moisture content* of the shell is determined by the toluene distillation method, collecting the distillate over a period of one hour. Additional water may be removed from "dry" capsules by further heating, e.g., at 40°C, but such a manufacturing step has not been found to be practical or necessary.

After drying, the capsules are transferred to the inspection department and held until released by the quality control department. The *inspection and quality control* steps in the processing of capsules are much the same as with other dosage forms and must conform to good manufacturing practices. Control tests specifically applicable to the quality of soft gelatin capsules may involve seal thickness determinations, total or shell moisture tests, capsule fragility or rupture tests, and the determination of freezing and high temperature effects.

Also, capsules may be sent after drying to a finishing department for heat branding or ink printing for purposes of identification. Final physical control processing and packaging may be accomplished by the following inline continuous operations.

1. A capsule *diameter sorter* passes the next unit capsule within ±0.020 inch of the theoretic diameter of the particular capsule being tested. Overfills, underfills, and "foreign" capsules are discarded. The unit is fed from a hopper, and the capsules are passed through a final naphtha washing unit just prior to the sorter. The unit employs a syntron vibrator, which is a series of divergent wire lanes, and can be used for capsule diameters ranging from 0.200 to 0.500 inch.
2. A capsule *color sorter* is the next unit in line. The capsules are fed into automatically from the diameter sorter by a pneumatic conveyor. In this

unit, any capsule whose color does not conform to the reference color standard for that particular product is discarded, while satisfactory capsules pass immediately to an electronic counting and packaging unit.

3. The electronic *counting unit* can count as many as 8,000 capsules per minute (depending upon size) directly into the bulk shipping carton. A printout of the content of each carton and a printout of the number of cartons are automatically produced and made a part of the production record. Following this step, the cartons are labeled, sealed, and palletized and are then ready for shipment.

## PHYSICAL STABILITY AND PACKAGING

Unprotected soft gelatin capsules (i.e. capsules that can breathe) rapidly reach equilibrium with the atmospheric conditions under which they are stored. This inherent characteristic warrants a brief discussion of the effects of temperature and humidity on these products, and points to the necessity of proper storage and packaging conditions and to the necessity of choosing an appropriate retail package. The variety of materials capsulated, which may have an effect on the gelatin shell, together with the many gelatin formulations that can be used, makes it imperative that physical standards are established for each product.

*General statements* relative to the effects of temperature and humidity on soft gelatin capsules must be confined to a control capsule that contains mineral oil, with a gelatin shell having a dry glycerin to dry gelatin ratio of about 0.5 to 1 and a water to dry gelatin ratio of 1 to 1, and which is dried to equilibrium with 20 to 30% RH at 21 to 24°C. The physical stability of soft gelatin capsules is associated primarily with the pick-up or loss of water by the capsule shell. If these are prevented by proper packaging, the above control capsule should have satisfactory physical stability at temperatures ranging from just above freezing to as high as 60°C.

a. For the unprotected control capsule, low humidities (<20% RH), low temperatures (<2°C) and high temperatures (>38°C) or combinations of these conditions have only transient effects. The capsule returns to normal when returned to optimum storage conditions. The transient effects are primarily brittleness and greater susceptibility to shock, requiring greater care in handling or a return to proper storage conditions prior to further handling.

b. As the humidity is increased, within a reasonable temperature range, the shell of the unprotected control capsule should pick up moisture in proportion to its glycerin and gelatin content in accordance with the curves shown in Fig. 14.25. The total moisture content of the capsule shell, at equilibrium with any given relative humidity within a reasonable temperature range, should closely approximate the sum of the moisture content of the glycerin and the gelatin when held separately at the stated conditions. For example, the shell of the described control capsule contains 400 mg of dry gelatin and 200 mg of dry glycerin per gram. At

equilibrium with 30% RH at room temperature (21 to 24°C), the curves show that the gelatin should retain about 12% (48 mg) of water, and the glycerin 7% (14 mg) of water. Thus, the "dry" shell would contain about 9.4% water (62 mg $H_20$/662 mg of shell).

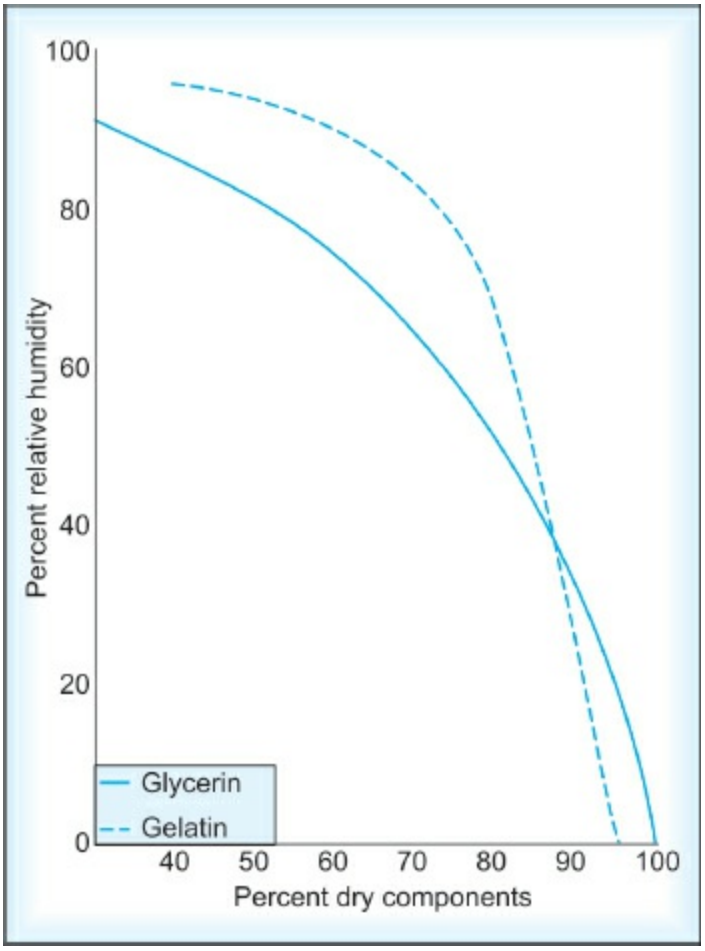



**Fig. 14.25:** Equilibrium content by weight of dry glycerin (20 to 100°C) and dry gelatin (25°C) at various relative humidities

If the conditions are changed to 60% RH (21 to 24°C), the moisture content should be approximately 17.4%. In actual practice, however, such calculated moistures are considered maximum, since moisture assays (toluene distillation method) of the shells of oil-filled capsules give results somewhat lower than the theoretical values. The deviation most likely is due to an interaction between the plasticizer and gelatin, partially satisfying their respective water-binding capacities and thereby causing a lower moisture content than would be theoretically expected.

Nevertheless, the curves serve to illustrate the hygroscopic nature of capsule shells, the relative effect of changes in the glycerin to gelatin ratio on the hygroscopicity of the shell, and the potential effects of humidity on the chemical and physical stability of the product.

c. High humidities (>60% RH at 21 to 24°C) produce more lasting effects on the capsule shell, since as moisture is absorbed, the capsules become softer, tackier, and bloated. The capsules do not leak unless the increased moisture has allowed a deleterious ingredient in the capsule content to attack the gelatin. On return to optimum storage conditions, the capsules are dull in appearance and most likely inseparably stuck together. An increase in temperature (>24°C), together with humidity (>45%), results in more rapid and pronounced effects and may even cause the unprotected capsules to melt and fuse together. Capsules containing water-soluble or miscible liquid bases may be affected to a greater extent than oil-based capsules, owing to the residual moisture in the capsule content and to the dynamic relationship existing between capsule shell and capsule fill during the drying process.

The capsule manufacturer routinely conducts *accelerated physical stability tests* on all new capsule products as an integral part of the product development program. The following tests have proved adequate for determining the effect of the capsule content on the gelatin shell. The tests are strictly relevant to the integrity of the gelatin shell and should not be construed as stability tests for the active ingredients in the capsule content. The results of such tests are used as a guide for the reformulation of the capsule content or the capsule shell, or for the selection of the proper retail package. The test conditions are: (1) 80% RH at room temperature in an open container; (2) 40°C in an open container; (3) 40°C in a closed container (glass bottle with tight screw-cap). The capsules at these stations are observed periodically for 2 weeks. Both gross and subtle effects of the storage conditions on the capsule shell are noted and recorded. The control capsule should not be affected except at the 80% RH station, where the capsule would react as described under the effects of high humidity.

In the case of a newly developed product, the gross effects such as disintegration, leakers, unusual brittleness or softening, apparent color fading, or discoloration are obvious. The more subtle changes may be the loss of a volatile ingredient as detected by slight capsule indentation, or the slight

darkening or widening of the capsule seams, or slight changes in color hue. Capsules often show a “soft spot” at the site at which they lie next to the tray or against another capsule. This spot is the result of slower drying and is of no consequence in the control capsule, since such areas firm up and are not flaws in the capsule shell. On the other hand, if such areas do not become firm, usually because of action by the capsule content, then physical stability problems pan be anticipated during the shelf-life of the product. Such defects must be corrected before the product can be considered for production. Correction of such defects depends upon identifying their cause. Most defects can be corrected by appropriate changes in gelatin or fill material formulations, but in some cases, different colorants, machine speeds, and machine dies may have to be used. The experience and mature judgment of the custom manufacturer is invaluable in the solution of such problems.

Chemists conducting the physical stability tests in their own laboratories should keep two important points in mind (1) prior to testing, the capsules should be equilibrated to known atmospheric conditions, preferably 20 to 30% RH at 21 to 24°C, (2) evaluation of the results of the previously described heat tests should be made only after the capsules have returned to equilibrium with room temperature.

After the capsules have passed the shell integrity tests, additional physical studies should be conducted using the various types of retail packages being considered for the product. These latter tests should be designed to determine the shelf-life of the product and may conform to most of the standard testing procedures employed by a company for its other solid dosage forms. Exceptions may involve those tests conducted at temperatures exceeding 45°C for time periods exceeding a month.

The soft gelatin capsule manufacturer takes great care in the production of capsules to meet the specifications of the product set forth by the customer. When *bulk shipments* of capsules are made by the manufacturer, they are temporarily protected from normal changes in humidity by a suitable moisture barrier such as a 0.003 inch polyethylene bag within a standard fiber board carton. Since such packaging is not a permanent moisture barrier, the capsules should be retail packaged as soon as possible after receipt. If immediate packaging is not practical, the bulk capsules in their original unopened cartons should be stored in an air-conditioned area in which the humidity does not exceed 45% RH at 21 to 24°C. The retail packaging of

these capsules should be done under similar conditions, for the maximum physical and chemical stability of the product.

Soft gelatin capsules may be *retail packaged* using any modern packaging equipment, including the electronic type. Capsules may be packaged in glass or plastic containers or may be strip-packaged. Suppliers of rigid plastics and plastic films can be of immeasurable service in suggesting the proper types of packaging for testing. Since strip packaging is usually done by the specialists in this field, their advice should be solicited, and test strips should be made and tested for adequacy.

# BIOPHARMACEUTICAL ASPECTS

## Dissolution Rate

In the dissolution rate studies of twenty drugs, presented previously, which included a wide variety of chemical types and pharmacologic classes, the authors showed that in the majority of cases, the drugs were more rapidly and completely available from the soft gelatin capsule than from the commercial tablets or capsules. For these studies, the NF XII (second supplement, 1967) rotating-bottle method was used.

A rationale for using a rotating-bottle method for dissolution studies on soft gelatin capsules is expressed, and examples are given by Withey and Mainville. Their dissolution studies on thirteen brands of commercial chloramphenicol capsules, using their modified USP apparatus, showed the soft gelatin capsule brand to release only 22.3 to 24.8% of its chloramphenicol content in 30 min, while upon changing to a rotating-bottle method, the 30-min recoveries were 100% from the soft gelatin capsule brand, 82% from brand B2, and 70% from brand D.

The difference could be attributed to greater agitation by the bottle method and less opportunity for the capsule to adhere to the sides or bottom of the apparatus. The effect of several variables on capsule dissolution is discussed by Horn and associates, who indicate that the degree of agitation, the pH of the dissolution medium, and the presence or absence of pepsin in the medium are important to the dissolution of soft gelatin capsules.

## Bioavailability

Nelson, in his review, points out that the availability of a drug for absorption, from various types of oral formulations, usually decreases in the following order: solution, suspension, soft gelatin capsule, powder-filled capsule, compressed tablet, coated tablet. A study by Wagner and co-workers confirms Nelson's observation. Their study involved the effect of dosage form on the serum levels of indoxole (solubility in water 0.1 μg/ml) and showed the serum level decreased in the following order: emulsion, soft gelatin capsule (drug in polysorbate 80 solution), aqueous suspension, powder-filled capsule.

Maconachie, in his review article on soft gelatin capsules, gives some specific examples of how this dosage form can improve drug absorption. His examples involve acetaminophen, chlormethiazole, and temazepam. The 4-hour urinary recovery of acetaminophen from three soft gelatin rectal suppository formulations (oil base and water-miscible type base) was found to be five to eight times greater than from the traditional fatty type suppository formulation. The switch from a tablet to a soft gelatin capsule form not only improved the stability of chlormethiazole by protecting the drug from oxidation, but increased its bioavailability as evidenced by comparative blood levels and by earlier onset of a minor side effect (nose tingling). Major side effects of the tablet dosage form were also eliminated or ameliorated. The capsule formulation allowed the use of the liquid drug substance (chlormethiazole base) rather than the solid derivatives used in the tablet formulations. A temazepam soft gelatin formulation, when compared with hard gelatin capsule formulations of temazepam, nitrazepam, amobarbital sodium, and a placebo, gave superior bioavailability as indicated by "onset of sleep." Furthermore, this was achieved at a lower dosage (20 mg per soft capsules versus 30 mg per hard capsule).

In an article on soft gelatin capsules, Ebert discusses and reports on the bioavailability and content uniformity of digoxin solutions in soft gelatin capsules. The capsulated solutions were 0.05 mg, 0.10 mg and 0.20 mg of digoxin dissolved in a base consisting of polyethylene glycol 400, USP (89.4% w/w); alcohol, USP (6.5% w/w); propylene glycol, USP (3.4% w/w); and purified water, USP (0.6% w/w). These capsules were tested by various investigators for bioavailability in comparison with brand name tablets,

digoxin solution, and digitoxin elixir. In all studies, the bioavailability of the soft gelatin capsule formulation was found to be superior to the commercially available tablets. The most surprising finding of all these studies, according to Ebert, was that the capsulated solution exhibited a more rapid and complete absorption than did the same solution not encapsulated. The commercially available tablets contain 20% more drug than the capsulated solutions. Thus, the capsule dosage form allows for a significant reduction in dose for this relatively toxic drug.

Another comparative bioavailability study of digoxin soft gelatin capsules and tablets was reported by Astorri and co-workers. They found that in heart patients using digoxin, the absorption of digoxin from the capsulated solution was 36% higher than from the tablet, while in healthy volunteers absorption from the capsule was 20% higher than from the tablets.

The bioavailability of theophylline from soft gelatin capsules in comparison to a commercially available liquid aminophylline preparation and to a nonalcoholic aminophylline solution was studied by Ebert and by Lesko and coworkers. Both studies found that the two dosage forms were bioequivalent as measured by the area under the plasma-leveltime curves. This is an example of the capsule providing a convenient portable dosage form for a liquid medication. The capsule also effectively masked the bitter taste of theophylline.

Papaverine hydrochloride bioavailability from soft gelatin capsules was studied by Arnold et al. These authors found that the peak blood level and area under the blood-level-time curve from soft gelatin capsules were equal to those obtained from an elixir and superior to those from a sustained release hard-shell capsule formulation. Healthy volunteers were used in this study. Lee et al. found not only high blood levels of papaverine 120 min after a 150 mg dose in soft gelatin capsules, but a higher degree of vasodilation after four doses (150 mg × 4) of the soft capsule dosage form when compared with equivalent doses from a sustained-release dosage form. Patients with severe arteriosclerosis obliterans were used in the study. Both studies conclude that the soft gelatin capsule dosage form shows significant advantage over the sustained release tablet form of the drug.

The bioavailability of diazepam (structurally similar to that of temazepam, mentioned previously) was studied by Yamahira et al. They compared diazepam, capsulated in soft gelatin using a medium-chain

triglyceride base, with a tablet dosage form. They reported that when these dosage forms were repeatedly orally administered to an individual subject, the capsule dosage form showed a tendency toward faster drug absorption and superior reproducibility of the plasma-level-time curve than the tablet dosage form. This suggests that the capsule dosage forms have a more uniform drug absorption rate than tablets. The authors suggests that diazepam, though it a weak base, was emptied from the stomach while mostly retained in the lipid, and this was affected by the movement of the triglyceride in the gastrointestinal tract.

Orally administered drugs, particularly if used chronically, can be irritating to the stomach. The dosage form of such drugs can affect gastric tolerance, as indicated by the studies of Caldwell, et al. These authors compared the degree of irritation or ulcerogenic potential of soft gelatin capsule formulations of dexamethasone with a tablet formulation of the drug. Several liquid formulations and tablet formulations were administered to rats, and both ulcerogenic potential and bioavailability were determined. The authors concluded that the liquid or capsule formulations had a reduced ulcerogenic potential when compared to the tablet formulation, and that this effect is apparently not a reflection of reduced bioavailability.

The pharmaceutical chemist should certainly consider the bioavailability potential of soft gelatin formulations. The biopharmaceutical characteristics of such formulations can be altered or adjusted more easily than those of other solid dosage forms. Through the selection and use of liquids and combinations of liquids that range from water-immiscible through emulsifiable to completely water-miscible, and by altering the type or quantity of thickening or suspending agents, capsule formulations allow the formulating chemist more flexibility in the design of a dosage form to fit the biopharmaceutical specifications of a particular therapeutic agent.

---

* For example, a capsule to contain 500 mg of polysorbate 80 would have a calculated $\left(\frac{0.5\text{g} \times 16.23\,\text{minim}}{1.08\text{g}}\right)$ fill volume of about 7.5 minims. Assuming, however, that there is 5% residual water in the dry capsule, the final fill volume would be about 8 minims $\left(\frac{0.525\text{g} \times 16.23\,\text{minim}}{1.08\text{g}}\right)$

# 15:
# Microencapsulation

Microencapsulation is a rapidly expanding technology. As a process, it is a means of applying relatively thin coatings to small particles of solids or droplets of liquids and dispersions. For the purpose of this chapter, microencapsulation is arbitrarily differentiated from macrocoating techniques, in that the former involves the coating of particles ranging dimensionally from several tenths of a micron to 5000 microns in size. As the technology has developed, it has become apparent that the concept offers the industrial pharmacist a new working tool. Microencapsulation provides the means of converting liquids to solids, of altering colloidal and surface properties, of providing environmental protection, and of controlling the release characteristics or availability of coated materials. Several of these properties can be attained by macropackaging techniques; however, the uniqueness of microencapsulation is the smallness of the coated particles and their subsequent use and adaptation to a wide variety of dosage forms and product applications, which heretofore might not have been technically feasible. Because of the smallness of the particles, drug moieties can be widely distributed throughout the gastrointestinal tract, thus potentially improving drug sorption.

This new technology does not exclude problem areas; for instance, no single microencapsulation process is adaptable to all core material candidates or product applications. Difficulties, such as incomplete or discontinuous coating, inadequate stability or shelf-life of sensitive pharmaceuticals, nonreproducible and unstable release characteristics of coated products, and economic limitations are often encountered in the attempt to apply a particular microencapsulation method to a specific task. Many times, successful adaptation is, in part, a result of the technical ingenuity of the investigators.

Microencapsulation is receiving considerable attention fundamentally, developmentally, and commercially. In view of this interest, it is the purpose of this chapter to present a description of the more prominent

microencapsulation methods and some of their capabilities and limitations. The microencapsulation methods to be discussed are air suspension, coacervation-phase separation, spray drying and congealing, and polymerization techniques. A survey of the ever-expanding patent and published literature reveals that not all microencapsulation techniques are included within the methods cited in this chapter; however, the methods described represent the currently established, most highly developed and widely used commercial processes, although some may not be applicable to pharmaceuticals at this time.

## Applications

The applications of microencapsulation might well include sustained-release or prolonged-action medications, taste-masked chewable tablets, powders and suspensions, single-layer tablets containing chemically incompatible ingredients, and new formulation concepts for creams, ointments, aerosols, dressings, plasters, suppositories, and injectables. Pharmaceutically related areas, such as hygiene, diagnostic aids, and medical equipment design, also are amenable to microencapsulation applications.

Three important areas of current microencapsulation application are the stabilization of core materials, the control of the release or availability of core materials, and separation of chemically reactive ingredients within a tablet or powder mixture.

The following examples illustrate the concept of improved stabilization. Microencapsulation of certain vitamins to retard degradative losses has been practiced for many years. The potency retention properties of a microencapsulated vitamin A palmitate oil are illustrated in Fig. 15.1. The conversion of volatile liquids to dry, free-flowing powders with subsequent retention of the liquid core material during extended storage is another example of stabilization. Figure 15.2 depicts typical microcapsule stabilities of an anthelmintic (carbon tetrachloride), methyl salicylate, and a flavor. An example of stability enhancement, accomplished by microencapsulation of incompatible admixed constituents, is given in Fig. 15.3. Compared in the graph is the formation of the aspirin hydrolysis product, salicylic acid, occurring with a mixture of aspirin and chlorpheniramine maleate.

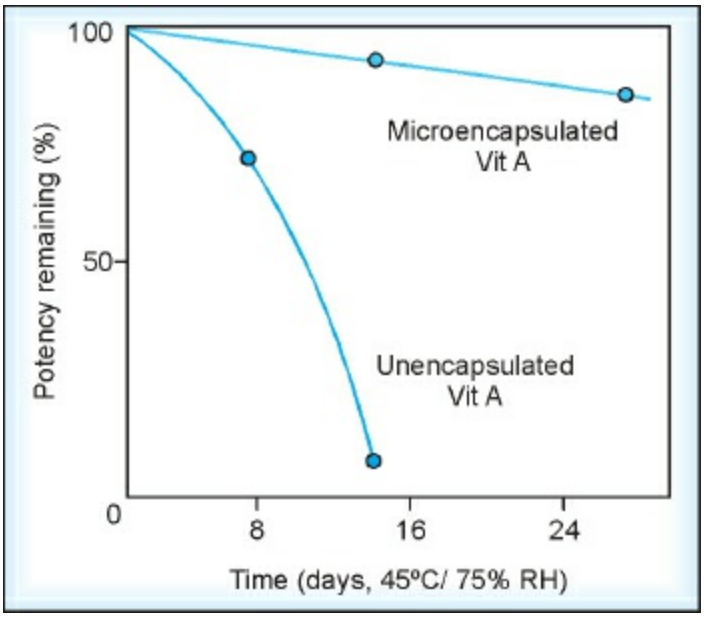



**Fig. 15.1:** Stability of a microencapsulated vitamin A palmitate com oil prepared by phase-separation/coacervation technique, compared with an unencapsulated control

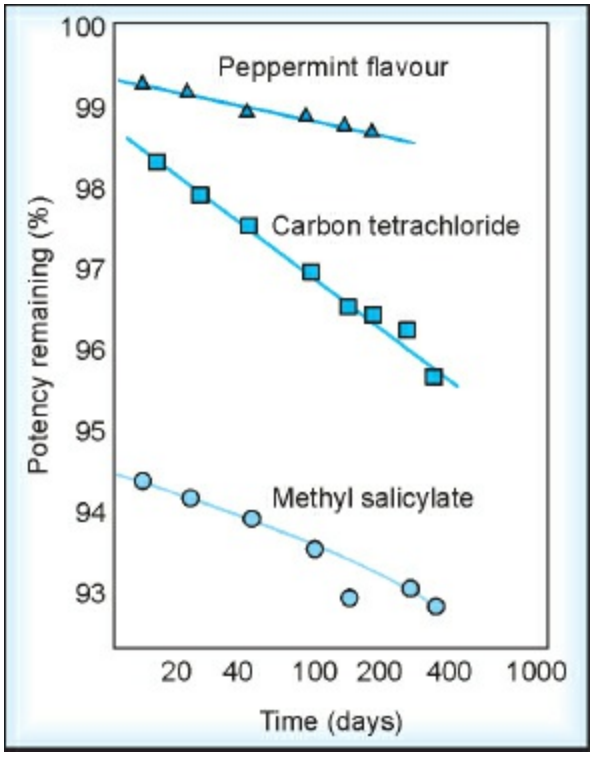



**Fig. 15.2:** Stability of microencapsulated volatile liquids, prepared by phase-separation/coacervation techniques

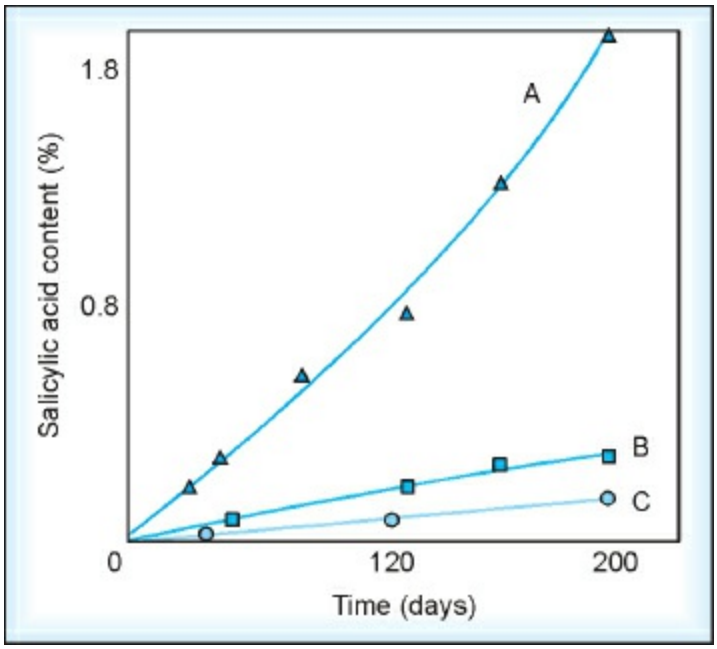



**Fig. 15.3:** Stability, enhancement of incompatible aspirin mixture by microencapsulation: (A) Aspirin hydrolysis of chlorpheniramine maleate aspirin mixture; (B) Aspirin hydrolysis of microencapsulated mixture; (C) Hydrolysis of aspirin control

The release properties of microencapsulated materials require detailed consideration, as the coated material must be released in a predictable and reproducible manner. A wide variety of mechanisms is available to release encapsulated core materials. Disruption of the coating can occur by pressure, shear, or abrasion forces, any of which affords a release mechanism. Other mechanisms involve permeability changes brought about enzymatically. Also, release can be achieved from inert coatings by diffusion or leaching of a permeant fluid. The rate of release is a function of the permeability of the coating to the extraction fluid; the permselectivity, if any, of the coating to core material solute; the dissolution rate of the core material; the coating thickness; and the concentration gradient existing across the coating membrane.

Prolonged-action or sustained-release formulations are obvious examples of controlled release from microencapsulated products, and Figs 15.4 and 15.5 demonstrate the versatility of two diverse microencapsulation processes and coating materials. The in vitro release patterns achieved by applying varied amounts of an ethylcellulose coating to small aspirin crystals using

coacervation phase-separation encapsulation techniques are shown in Fig. 15.4. Release of aspirin is accomplished in this case by a leaching or diffusion mechanism from inert, pH-insensitive ethylcellulose coating.

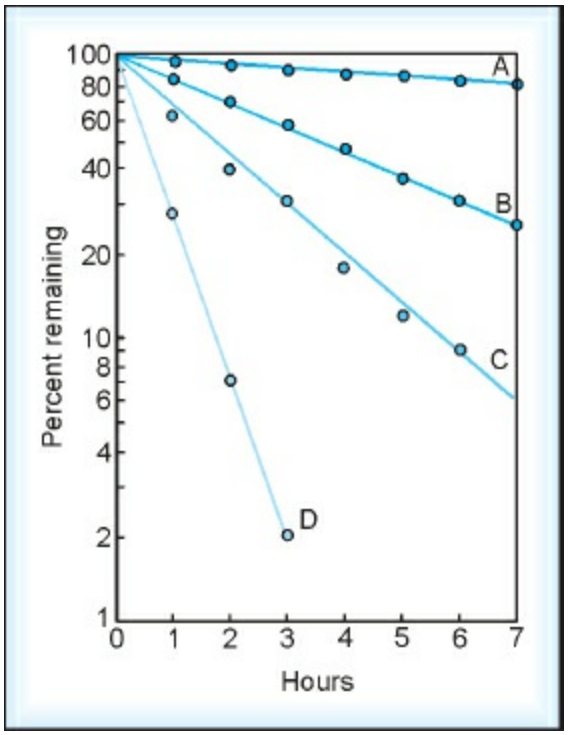



**Fig. 15.4:** In vitro release patterns of crystalline aspirin coated with various amounts of ethylcellulose using phase-separation/coacervation techniques: (A) 52% coating; (B) 29%-coating; (C) 16% coating; D 13% coating

Depicted in Fig. 15.5 are the in vitro release responses of amphetamine sulfate pellets that have been microencapsulated with varying amounts of a wax-fat coating applied by a pan coating process. Figure 15.5 illustrates the release rate effects of varied amounts of coating as well as the enteric nature of the coatings when subjected to the simulated gastrointestinal extraction conditions. The nonlinearity of the release curves plotted semilogarithmically suggests that the release of the amphetamine is accomplished initially by a leaching action through a gastric fluid-resistant coating followed by release from a dissolvable or disintegratable coating accomplished by action of the simulated intestinal fluid. Also shown in Fig. 15.5 is the resultant release pattern (segmented curve) for an appropriate sustained release blend of two of the coated materials and a non-coated fraction of the amphetamine.

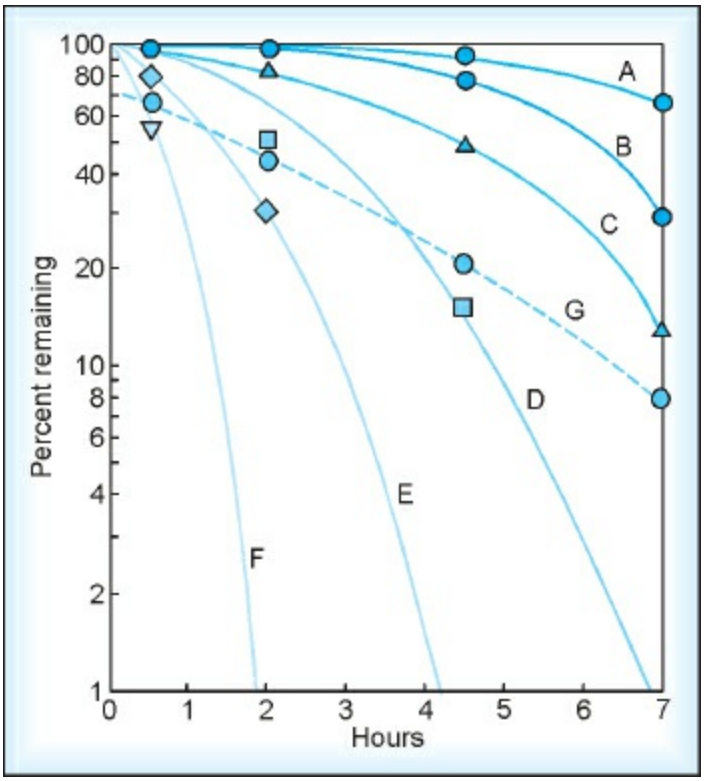



**Fig. 15.5:** In vitro release patterns of amphetamine sulfate pellets pan coated with various amounts of a fat-wax coating: (A) 17% coating; (B) 15% coating; (C) 13% coating; (D) 11% coating; (E) 9% coating; (F) 796 coating; (G) Selected blend of uncoated pellets and coated pellets

The in vitro release properties of the microencapsulated product forms already described—a rapidly disintegrating tablet containing coated particles, a microencapsulated powder, and a blend of coated pellets—all illustrate, in vitro, apparent first-order release kinetics, although diverse mechanisms are involved. Consequently, the rate of release of the drugs from these examples is proportional to the amount of drug remaining – $da/dt = k_r a$, or after integration, $a = a_0 e^{-krt}$ or In $(a/a_0) = -k_r t$, where $a_0$ is the total dose in the preparations, $a$ is the fraction of the total dose remaining in the coated preparation at time $t$ (if $a_0$ is unity), and $k_r$, is the apparent first-order release constant. For sustained-release preparations formulated to release a fraction ($f_i$) of the dose immediately, and another fraction ($f_r$) exponentially as equated previously, the amount of drug release ($a_r$) by any time thereafter is therefore:

$$a_r = a_0 f_i + a_0 f_r (1 - e^{-krt})$$

Many sustained-release products release all or part of their drug content exponentially according to the first-order rate equation. Improved gastrotolerability of drugs can be obtained by microencapsulation while good bioavailability is maintained. To prove the tolerability of the microencapsulated product, one study of gastrointestinal bleeding was done with Cr-labeled red blood cells in 20 subjects treated for 3 months with microencapsulated KCl in amount of 3 g daily. The daily blood loss in the feces of 20 subjects after adminis-tration of microencapsulated KCl was practically nil, mean values ranging from 0.432 to 0.596 ml in any 24 hour period, as opposed to 0.1 to 2.2 ml blood loss with raw KCl. In a study by McMahon et al. healthy male volunteers who had had no previous evidence of gastrointestinal disease were examined by endoscopy after ingestion of microencapsulated KCl and wax matrix KCl tablets. The results of the study indicated that microcapsules had minimal endoscopically discernible lesions when compared with wax matrix tablets.

## FORMULATION CONSIDERATIONS

The process of microencapsulation involves a basic understanding of the general properties of microcapsules, such as the nature of the core and coating materials, the stability and release characteristics of the coated materials, and the microencapsulation methods. One should also note that the method employed in the manufacture of microcapsules may well result in products of varied composition, quality, and utility.

## Core Material

The *core material*, defined as the specific material to be coated, can be liquid or solid in nature. The liquid core can include dispersed and/or dissolved material, whereas the solid core can be a mixture of active constituents, stabilizers, diluents, excipients, and release-rate retardants or accelerators. The ability to vary the core material composition provides definite flexibility and utilization of this characteristic often allows effectual design and development of the desired microcapsule properties.

It is not possible to discuss, or even list, all of the potential core materials and product applications that are or may be amenable to microencapsulation. However, to aid in illustrating the diversity of the materials and their applications, some of these products are listed in Table 15.1.

**Table 15.1:** Properties of some microencapsulated core materials

| Core material | Characteristic property | Purpose of encapsulation | Final product form |
|---|---|---|---|
| Acetaminophen | Slightly water-soluble solid | Taste-masking | Tablet |
| Activated charcoal | Adsorbent | Selective sorption | Dry powder |
| Aspirin | Slightly water-soluble solid | Taste-masking; sustained release; reduced gastric irritation; separation of incompatibles | Tablet or capsule |
| Islet of Langerhans | Viable cells | Sustained normalization of diabetic condition | Injectable |
| Isosorbide dinitrate | Water-soluble solid | Sustained release | Capsule |
| Liquid crystals | Liquid | Conversion of liquid to solid; stabilization | Flexible film for thermal mapping of anatomy |
| Menthol/methyl salicylate camphor mixture | Volatile solution | Reduction of volatility; sustained release | Lotion |
| Progesterone | Slightly water-soluble solid | Sustained release | Varied |
| Potassium chloride | Highly water-soluble solid | Reduced gastric irritation | Capsule |
| Urease | Water-soluble enzyme | Permselectivity of enzyme, substrate, and reaction products | Dispersion |
| Vitamin A palmitate | Nonvolatile liquid | Stabilization to oxidation | Dry powder |

## Coating Materials

The selection of a specific coating material from a lengthy list of candidate materials presents the following questions to be considered by the research pharmacist.

1. What are the specific dosage or product requirements—stabilization, reduced volatility, release characteristics, environmental conditions, etc?
2. What coating material will satisfy the product objectives and requirements?
3. What microencapsulation method is best suited to accomplish the coated product objectives?

The selection of the appropriate coating material dictates, to a major degree, the resultant physical and chemical properties of the microcapsules, and consequently, this selection must be given due consideration. The coating material should be capable of forming a film that is cohesive with the core material; be chemically compatible and nonreactive with core material; and provide the desired coating properties, such as strength, flexibility, impermeability, optical properties, and stability. The coating materials used in microencapsulation methods are amenable, to some extent, to in situ modification. For example, colorants may be added to achieve product elegance or masking, or coatings may be plasticized or chemically altered through cross-linking, for instance, to achieve controlled dissolution or permeability. A partial listing of typical coating materials commonly used in the various microencapsulation methods is suggested in Table 15.2.

**Table 15.2:** Representative coating materials and applicable microencapsulation process

| Coating materials | Processes | | | | | |
|---|---|---|---|---|---|---|
| | Multiorifice—centrifugal | Phase Separation—coacervation | Pan coating | Spray drying and congealing | Air suspension | Solvent evaporation |
| **Water-soluble resins** | | | | | | |
| Gelatin | X | X | X | X | X | X |
| Gum arabic | | X | X | X | X | X |
| Starch | | X | X | X | X | |
| Polyvinylpyrrolidone | X | X | X | X | X | |
| Carboxymethylcellulose | | X | X | X | X | |
| Hydroxyethylcellulose | | X | X | X | X | X |
| Methylcellulose | | X | X | X | X | |
| Arabinogalactan | | X | X | X | X | |
| Polyvinyl alcohol | X | X | X | X | X | X |
| Polyacrylic acid | | X | X | X | X | X |
| **Water-insoluble resins** | | | | | | |
| Ethylcellulose | | X | X | X | X | X |
| Polyethylene | X | | | | X | X |
| Polymethacrylate | | X | X | X | X | X |
| Polyamide (nylon) | | | | | X | X |
| Poly (ethylene-vinyl acetate) | X | X | X | X | | X |
| Cellulose nitrate | X | X | X | X | | X |
| Silicones | | | X | X | | |
| Poly (lactide-co-glycolide) | | X | X | | | X |
| **Waxes and lipids** | | | | | | |
| Paraffin | X | X | X | X | X | |
| Carnauba | | | X | X | X | |
| Spermaceti | | X | X | X | X | |
| Beeswax | | | X | X | X | |
| Stearic acid | | | X | X | | |
| Stearyl alcohol | | | X | X | X | |
| Glyceryl stearates | | | X | X | X | |
| **Enteric resins** | | | | | | |
| Shellac | | X | X | X | X | |
| Cellulose acetate phthalates | | X | X | X | X | X |
| Zein | | X | | | X | |

It is not within the scope of this discussion to describe the physical and chemical properties of coatings per se. However, it is pointed out, that typical coating properties such as cohe-siveness, permeability, moisture sorption, solubility, stability, and clarity must be considered in the selection of the proper microcapsule coating material. The selection of a given coating often can be aided by the review of existing literature and by the study of free or cast films, although practical use of free-film information often is impeded

for the following reasons:

1. Cast or free films prepared by the usual casting techniques yield films that are considerably thicker than those produced by the microencapsulation of small particles; hence, the results obtained from the cast films may not be extrapolatable to the thin microcapsule coatings.
2. The particular microencapsulation method employed for the deposition of a given coating produces specific and inherent properties that are difficult to simulate with existing film-casting methods.
3. The coating substrate or core material may have a decisive effect on coating material properties.

Hence, the selection of a particular coating material involves consideration of both classic free-film data and applied results.

As previously stated, the uniqueness of microcapsules in their properties and use involves their characteristic smallness. Consequently, the protective coatings that are applied are quite thin. Although the active content of many microencapsulated products can be varied from a few percent to over 99%, the effective coating thickness that can be realized, regardless of the method of application employed, varies from tenths of a micron to a few hundred microns, depending on the coating-to-core ratio and the particle size (surface area) of the core material. Figure 15.6 illustrates the theoretic film thickness that can be applied to small spherical particles. The thinness of microencapsulation coatings, although not necessarily limiting, must be of prime consideration. Just as the smallness of microcapsules allows unique properties and formulations to be accomplished, the thinness of the resultant coatings also can present unique problems. For example, most polymers exhibit microscopic discontinuities and some degree of ordered or random crystallinity. The total thickness of the coatings achieved with microencapsulation techniques is microscopic in size, and therefore, what might be a minor non-homogeneity occurring on the surface of a 5-mil coating can penetrate the entire thickness of a microencapsulation coating.

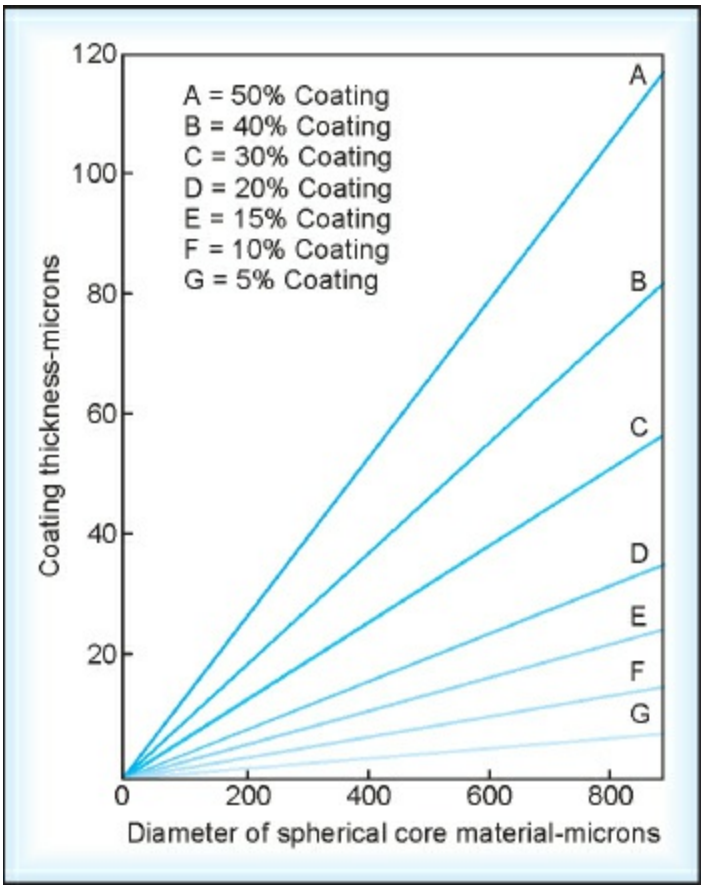



**Fig. 15.6:** Theoretic coating thickness for spherical core material having various amounts of coating

## METHODOLOGY

Microencapsulation methods that have been or are being adapted to pharmaceutical use include air suspension, coacervation-phase separation, spray drying and congealing, pan coating, and solvent evaporation techniques. Methods not currently applicable to pharmaceutical preparations are vacuum deposition and polymerization techniques. The physical nature of the core materials and the particle size ranges applicable to each process are given in Table 15.3. A representative photomicrograph of a colloidal particle formed through modified solvent evaporation method is shown in Fig. 15.7.

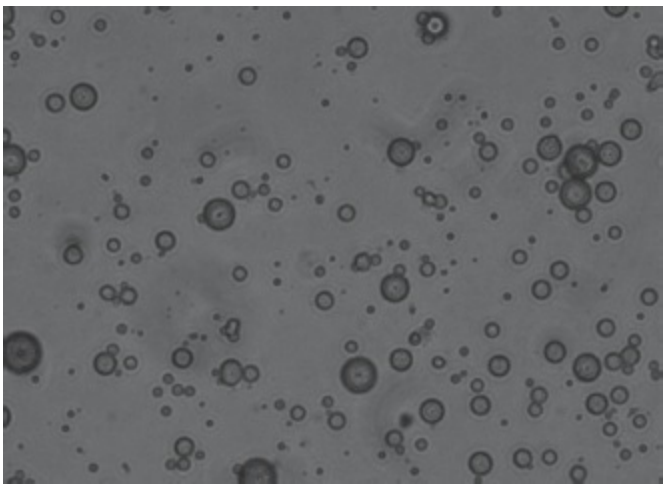


**Fig. 15.7:** Photomicrograph of a colloidal particle formed through modified solvent evaporation method

**Table 15.3:** Microencapsulation processes and their applicabilities

| Microencapsulation process | Applicable core material | Approximate particle size (μm) |
|---|---|---|
| Air suspension | Solids | 35–5000* |
| Coacervation-phase separation | Solids and liquids | 2–5000* |
| Multiorifice centrifugal | Solids and liquids | 1–5000* |
| Pan coating | Solids | 600–5000* |
| Solvent evaporation | Solids and liquids | 5–5000* |

| | | |
|---|---|---|
| Spray drying and congealing | Solids and liquids | 600 |

* The 5000 μm size is not a particle limitation. The methods are also applicable to macrocoating, i.e. particles greater than 5000 μm to size

## Air Suspension

Microencapsulation by air suspension techniques is generally ascribed to the inventions of Professor Dale E. Wurster during his tenure at the University of Wisconsin. Basically, the Wurster process consists of the dispersing of solid, particulate core materials in a supporting air stream and the spray-coating of the air-suspended particles. Figure 15.8 depicts a type of the Wurster air suspension encapsulation unit. Within the coating chamber, particles are suspended on an upward moving air stream as indicated in the drawing. The design of the chamber and its operating parameters effect a recirculating flow of the particles through the coating zone portion of the chamber, where a coating material, usually a polymer solution, is spray-applied to the moving particles. During each pass through the coating zone, the core material receives an increment of coating material. The cyclic process is repeated, perhaps several hundred times during processing, depending on the purpose of microencapsulation, the coating thickness desired, or whether the core material particles are thoroughly encapsulated. The supporting air stream also serves to dry the product while it is being encapsulated. Drying rates are directly related to the volume temperature of the supporting air stream.

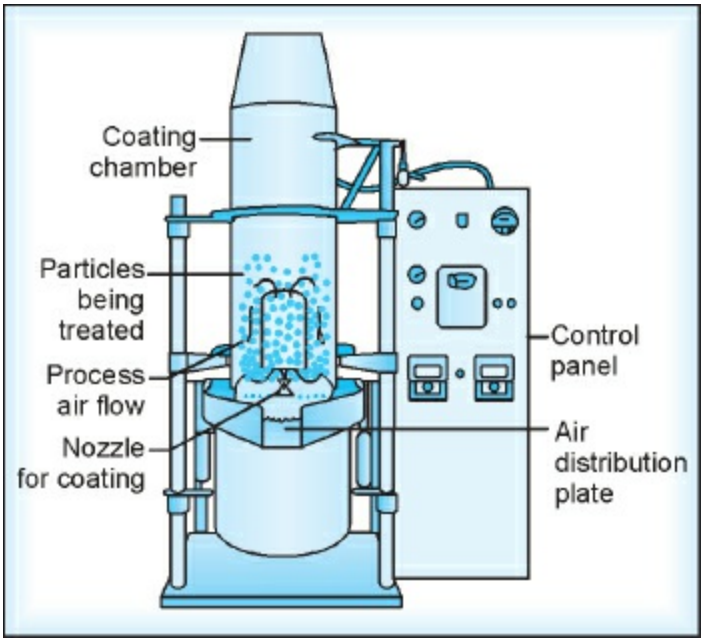



**Fig. 15.8:** Schematic drawings of Wurster air suspension apparatus

Processing variables that receive consideration for efficient, effective

encapsulation by air suspension techniques include the following:

1. Density, surface area, melting point, solubility, friability, volatility, crystallinity, and flowability of the core material.
2. Coating material concentration (or melting point if not a solution).
3. Coating material application rate.
4. Volume of air required to support and fluidize the core material.
5. Amount of coating material required.
6. Inlet and outlet operating temperatures.

The air suspension process offers a wide variety of coating material candidates for microencapsulation. The process has the capability of applying coatings in the form of solvent solutions, aqueous solutions, emulsions, dispersions, or hot melts in equipment ranging in capacities from one pound to 990 pounds. A partial listing of the coating materials is listed in Table 15.2. The coating material selection appears to be limited only in that the coating must form a cohesive bond with the core material. The process generally is considered to be applicable only to the encapsulation of solid core materials as indicated in Table 15.3. Indirectly, however, liquids can be encapsulated by the process at relatively low active levels by coating solid sorbents that have been pretreated with liquid sorbates. In regard to particle size, the air suspension technique is applicable to both *microencapsulation* and *macroencapsulation* coating processes. The practical particle size range for microencapsulation, however, is considered to be in excess of 74 microns. Under idealized conditions, particles as small as 37 microns can be effectively encapsulated as single entities. Core materials comprised of micron or submicron particles can be effectively encapsulated by air suspension techniques, but agglomeration of the particles to some larger size is normally achieved.

## Coacervation-phase Separation

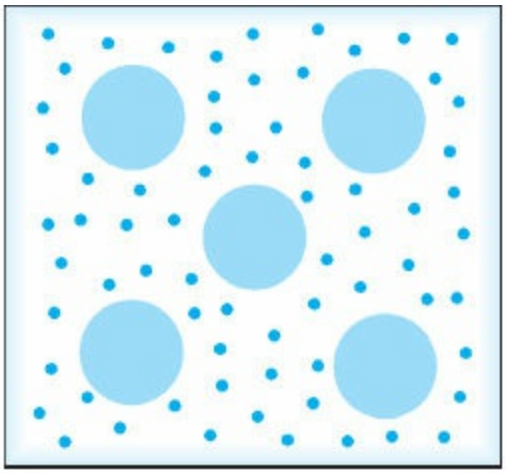
Microencapsulation by coacervation-phase separation is generally attributed to The National Cash Register (NCR) Corporation and the patents of B. K. Green et al. The general outline of the processes consists of three steps carried out under continuous agitation (1) formation of three immiscible chemical phases, (2) deposition of the coating and (3) rigidization of the coating (Fig. 15.9).

**Step 1** of the process is the formation of three immiscible chemical phases: a liquid manufacturing vehicle phase, a core material phase, and a coating material phase. To form the three phases, the core material is dispersed in a solution of the coating polymer, the solvent for the polymer being the liquid manufacturing vehicle phase. The coating material phase (an immiscible polymer in a liquid state) is formed by utilizing one of the methods of phase separation-coacervation, that is, by changing the temperature of the polymer solution; or by adding a salt, nonsolvent, or incompatible polymer to the polymer solution; or by inducing a polymer-polymer interaction. These general modes of effecting liquid-liquid phase separation are discussed in more detail later in this chapter.

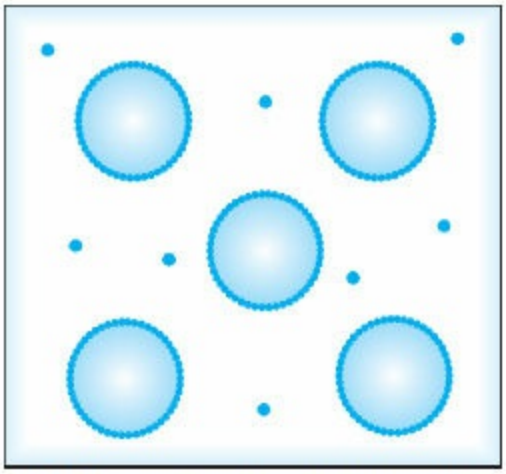
**Step 2** of the process consists of depositing the liquid polymer coatings upon the core material. This is accomplished by controlled, physical mixing of the coating material (while liquid) and the core material in the manufacturing vehicle. Deposition of the liquid polymer coating around the core material occurs if the polymer is adsorbed at the interface formed between the core material and the liquid vehicle phase, and this adsorption phenomenon is a prerequisite to effective coating. The continued deposition of the coating material is promoted by a reduction in the total free interfacial energy of the system, brought about by the decrease of the coating material surface area during coalescence of the liquid polymer droplets.

**Step 3** of the process involves rigidizing the coating, usually by thermal, cross-linking, or desolvation techniques, to form a selfsustaining microcapsule.

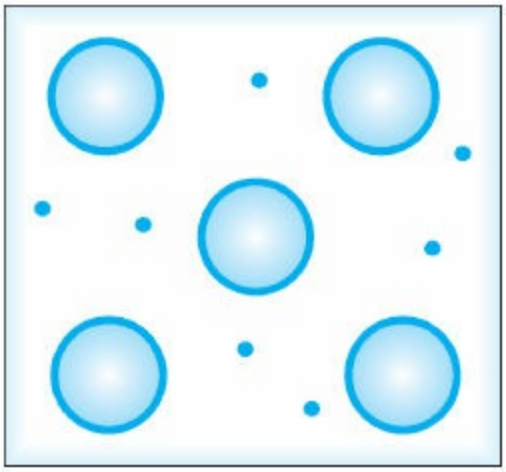
Because of the latitude in the material systems associated with the utilization of this general scheme of accomplishing microencapsulation, a representative example to illustrate each method is given here.



Step 1: Core and liquid coating in manufacturing vehicle



Step 2: Deposition of liquid coating material



Step 3: Completed capsules in manufacturing vehicle

**Fig. 15.9:** Schematic representation showing coating formation during phase-

separation/coacervation process. White- 1st phase, grey- 2nd phase, black- 3rd phase

## *Temperature Change*

Figure 15.10 illustrates a general temperature-composition phase diagram for a binary system comprised of a polymer and a solvent. A system having an overall composition, represented as point X on the abscissa, exists as a single-phase, homogeneous solution at all points above the phase-boundary or binodal curve, FEG. As the temperature of the system is decreased from point A along the arrowed line *AEB*, the phase boundary is crossed at point E, and the two-phase region is entered. Phase separation of the dissolved polymer occurs in the form of immiscible liquid droplets, and if a core material is present in the system, under proper polymer concentration, temperature, and agitation conditions, the liquid polymer droplets coalesce around the dispersed core material particles, thus forming the embryonic microcapsules. The phase-boundary curve indicates that with decreasing temperature, one phase becomes polymer-poor (the microencapsulation vehicle phase) and the second phase (the coating material phase) becomes polymer-rich. At point B, for instance, the segmented tie-line suggests that vehicle phase is essentially pure solvent, point C, whereas the coexisting phase, point D, is a concentrated polymer-solvent mixture. In practice, the loss of solvent by the polymer-rich phase can constitute gelatin of polymer, and hence rigidization or solidification of the microcapsule polymeric coating.

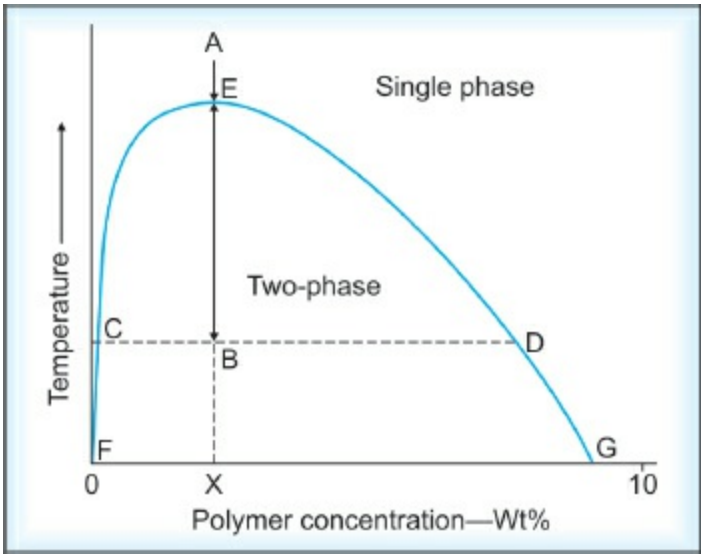


**Fig. 15.10:** General phase diagram-coacervation induced thermally

The following example illustrates a microencapsulation procedure that utilizes the phase-separation/coacervation principle. Ethylcellulose, a water-insoluble polymer, is applied to a water-soluble core material, N-acetyl p-aminophenol, by utilizing the temperature-solubility characteristics of the polymer in the hydrocarbon solvent cyclohexane. The etherified cellulosic, containing a relatively high ethoxyl content (high degree of substitution), is insoluble in cyclohexane at room temperature, but is soluble at elevated temperatures. Consequently, a working example involves dispersing ethylcellulose in cyclohexane to yield a polymer concentration of 2% by weight. The mixture is heated to the boiling point to form a homogeneous polymer solution. The core material, finely divided crystalline N-acetyl p-aminophenol, is dispersed in the solution with stirring at a coating-to-core material (dry) ratio of 1:2. Allowing the mixture to cool, with continued stirring, effects phase-separation/coacervation of the ethyl-cellulose and microencapsulation of the core material. Allowing the mixture to cool further to room temperature accomplishes gelation and solidification of the coating. The microencapsulated product can then be collected from the cyclohexane by filtration, decantation, or centrifugation techniques.

## *Incompatible Polymer Addition*

Liquid phase separation of a polymeric coating material and microencapsulation can be accomplished by utilizing the incompatibility of dissimilar polymers existing in a common solvent. Microencapsulation using this phenomenon is best described by considering the process in conjunction with the general phase diagram shown in Fig. 15.11.

The diagram illustrates a ternary system consisting of a solvent, and two polymers, X and Y. If an immiscible core material is dispersed in a solution of polymer Y (point A in Fig. 15.11) and polymer X is added to the system, denoted by the arrowed line, the phase boundary will be crossed at point *E*. As the two-phase region is penetrated with the further addition of polymer X, liquid polymer, immiscible droplets form and coalesce to form embryonic microcapsules. The coating of the microcapsules existing at point B, for example, consists of a concentrated solution of polymer Y dispersed in a solution comprised principally of polymer X, as indicated by the segmented tie-line and the phase-boundary intercepts C and *D*. The polymer that is more

strongly adsorbed at the core material-solvent interface, in this case polymer Y, becomes the coating material.

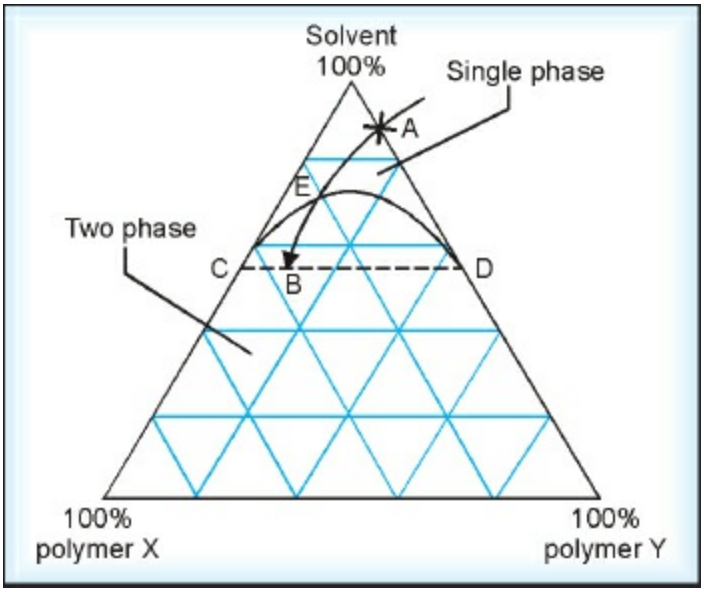



**Fig. 15.11:** General phase diagram of phase-separation/coacervation induced by the addition of an incompatible polymer

In practice, solidification of the coating material is accomplished by further penetration into the two-phase region, chemical cross-linking, or washing the embryonic microcapsules with a liquid that is a nonsolvent for the coating, polymer Y, and that is a solvent for polymer X.

The microencapsulation of methylene blue hydrochloride with ethylcellulose by this mode of phase separation (incompatible polymer addition) is described as follows. Ethylcellulose is dissolved in toluene to yield a polymer concentration of 2% by weight. Crystalline methylene blue hydrochloride, being essentially insoluble in toluene, is dispersed, with stirring, in the polymer solution at a ratio, for instance, of 4 parts methylene blue hydrochloride to 1 part ethylcellulose. Phase-separation/coacervation is accomplished by slowly adding liquid polybutadiene in sufficient quantity to yield a ratio of 25 parts polybutadiene to 1 part ethylcellulose. The polybutadiene, being quite soluble in toluene and incompatible with ethylcellulose, effects the demixing of the ethylcellulose from the polybutadiene toluene solution, and subsequent microencapsulation of the dispersed core material results. The ethylcellulose coating is solidified by adding a nonsolvent for the coating polymer such as hexane. Also, the

polybutadiene, being soluble in hexane, is washed from the mixture by decantation and by additional hexane wash cycles. The resultant product, crystalline methylene blue hydrochloride coated with ethylcellulose, is collected by standard filtration and drying techniques.

## *Nonsolvent Addition*

A liquid that is a nonsolvent for a given polymer can be added to a solution of the polymer to induce phase separation, as indicated by the general phase diagram given in Fig. 15.12. The resulting immiscible, liquid polymer can be utilized to effect microencapsulation of an immiscible core material as illustrated in the following example. A 5%, weight to volume, methyl ethyl ketone solution of cellulose acetate butyrate is prepared, and in it, micronized methylscopolamine hydrobromide is dispersed with stirring. A core-material to coating-material ratio (methylscopolamine hydrobromide to cellulose acetate butyrate) of about 2:1 is used. The resulting mixture is heated to 55°C, and isopropyl ether, a nonsolvent for the coating polymer, is added slowly to effect phase-separation/coacervation and microencapsulation of the suspended core material. The system is slowly cooled to room temperature, and the microencapsulatesd particles are separated by centrifugation, washed with isopropyl ether, and dried in vacuum.

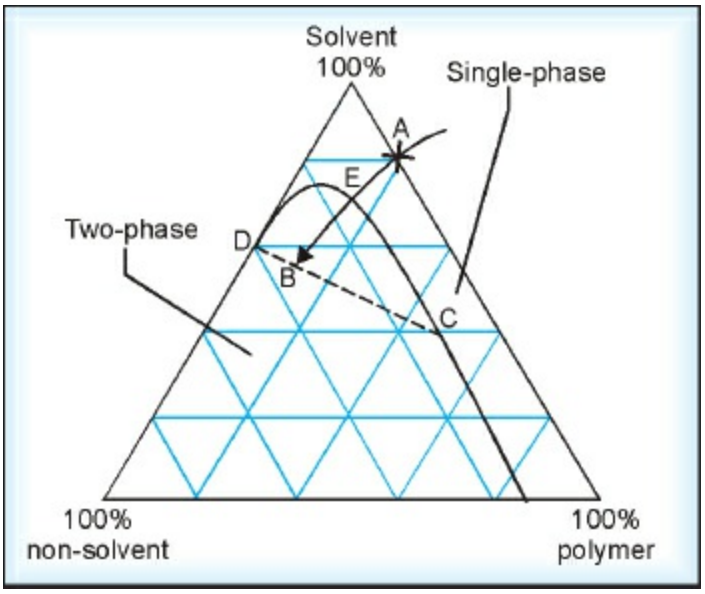



**Fig. 15.12:** General phase diagram for phase-separation/coacervation induced by the addition of a nonsolvent

## *Salt Addition*

Soluble inorganic salts can be added to aqueous solutions of certain water-soluble polymers to cause phase separation (Fig. 15.13). The following example of an oil-soluble vitamin microencapsulation induced by adding sodium sulfate to a gelatin solution illustrates the concept. An oil-soluble vitamin is dissolved in corn oil and is emulsified to the desired drop size in a 10% solution of high-quality pigskin gelatin having an isoelectric point at about pH 8.9. Twenty parts oil to 100 parts water, by weight, are used for the preparation of the oil/water emulsion. The emulsification process is conducted at 50°C, well above the gelation temperature of the gelatin. With the temperature of the emulsion maintained at 50°C, phase-separation/coacervation is induced by slowly adding a 20% solution of sodium sulfate. The salt solution is added in a ratio of 10 parts emulsion to 4 parts salt solution. The addition of the salt solution to the continuously stirred emulsion effects the microencapsulation of the oil droplets with a uniform coating of gelatin. The resultant protein coating is rigidized by transferring the mixture into a sodium sulfate solution that is 7% by weight and is maintained at 19°C, with continued agitation. The gelatin salt solution comprises a volume approximately ten times that of the microencapsulation mixture volume. The microencapsulated product is collected by filtration, washed with water, chilled below the gelation temperature of the gelatin (to remove the salt), and voided of water by standard drying techniques such as spray drying.

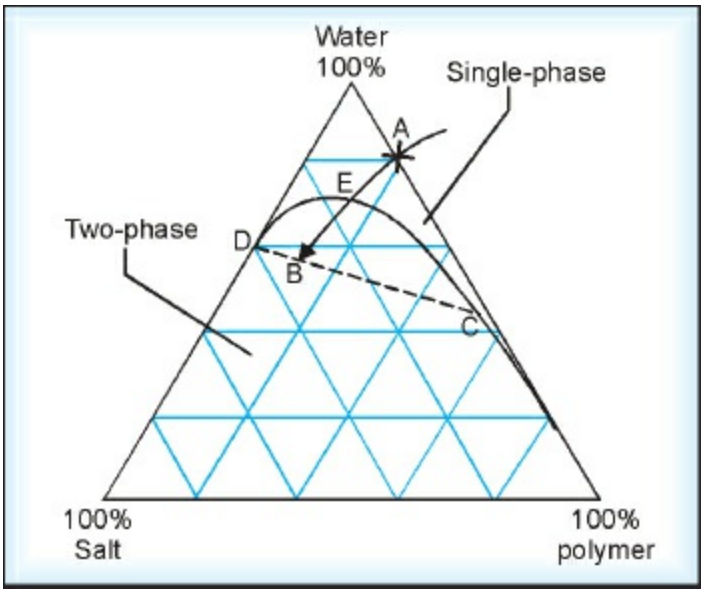


**Fig. 15.13:** General phase diagram for phase-separation/coacervation induced by salt addition

## *Polymer-polymer Interaction*

The interaction of oppositely charged polyelectrolytes can result in the formation of a complex having such reduced solubility that phase separation occurs. Figure 15.14 illustrates the phase diagram for a ternary system comprised of two dissimilarly charged polyelectrolytes and the solvent, water. In the dilute solution region, interaction of the oppositely charged polyelectrolytes occurs, inducing phase separation within the phaseboundary curve *ABA*. The segmented tie-lines indicate that a system, having an overall coacervation by polymer interaction composition within the two-phase region (point C for example), consists of two phases, one being polymer poor, point A, and one containing the hydrated, liquid complex, $Pa^+$ and $Pa^-$, point B.

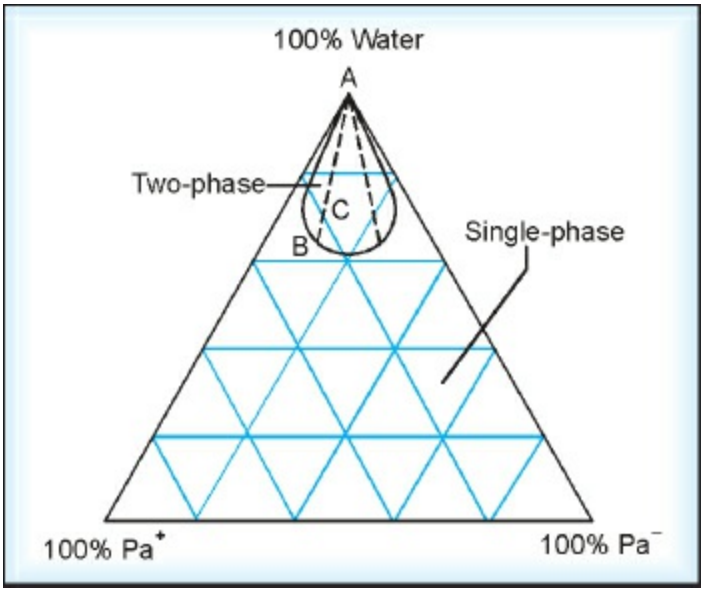



**Fig. 15.14:** Phase diagram for phase-separation/

As in the case of the previously described phase-separation/coacervation phenomenon, microencapsulation can be accomplished by polymer-polymer interaction. Gelatin and gum arabic are typical polyelectrolytes that can be caused to interact. Gelatin, at pH conditions below its isoelectric point, possesses a net positive charge, whereas the acidic gum arabic is negatively charged. Under the proper temperature, pH, and concentrations, the two

polymers can interact through their opposite electrical charges, forming a complex that exhibits phase separation/coacervation. The water-immiscible liquid, methyl salicylate, is an example of this process. Aqueous solutions of gum arabic and pigskin gelatin (isoelectric point 8.9) are prepared, each being 2% by weight in concentration. The homogeneous polymer solutions are mixed together in equal amounts, diluted to about twice their volume with water, adjusted to pH 4.5, and warmed to 40 to 45°C. The oppositely charged macromolecules interact at these conditions and undergo phase-separation/coacervation. While maintaining the warm temperature conditions, the liquid core material, methyl salicylate, is added at a weight ratio of, for instance, 25 parts methyl salicylate to 1 part gelatin-gum arabic (dry). The core material is emulsified by stirring to yield the desired drop size. The mixture is then slowly cooled to 25°C, with continued stirring, over a period of about one hour. During the cooling cycle, phase-separation/coacervation is further enhanced, resulting in the microencapsulation of the core material with the gelatin-gum arabic complex. The coating is rigidized for drying purposes by cooling the mixture to about 10°C. The materials used in coacervation-phase separation techniques are summarized in Table 15.4.

**Table 15.4:** Summary of examples for phase-separation/coacervation techniques

| Phase-separation/ coacervation method | Liquid manufacturing vehicle | Coating meterial | Core material |
|---|---|---|---|
| Temperature change (high to room temperature) | Cyclohexane | Ethylcellulose | N-acetyl p-amino-phenol |
| Incompatible polymer addition (liquid polybutadiene) | Toluene | Ethylcellulose | Methylene blue |
| Nonsolvent addition (isopropyl ether) | Methyl ethyl ketone | Cellulose acetate butyrate | Methylscopolamine Hydrobromide |
| Salt addition (sodium sulfate solution) | Water | Pig skin gelatin | Oil soluble vitamin |
| Polymer-polymer interaction (+ve charged gelatin and –ve charged gum arabic solution) | Water | Gelatin-gum arabic complex | Methyl salicylate |

Owing to the fact that core materials are microencapsulated while being dispersed in some liquid manufacturing vehicle, subsequent drying operations may be required. Typical drying methods such as spray, freeze, fluid bed,

solvent, and tray drying techniques are amenable to the microencapsulated products. The phase-separation/coacervation processes are conducted as batch operations using common production equipment in the manner illustrated in Fig. 15.15. A wide variety of liquids, solids, or suspensions can be microencapsulated in various sizes (*see* Table 15.1) having a variety of coatings (*see* Table 15.2). Microcapsules are being manufactured in vessels up to 2000 gallons in capacity, at a multimillion pound per annum rate.

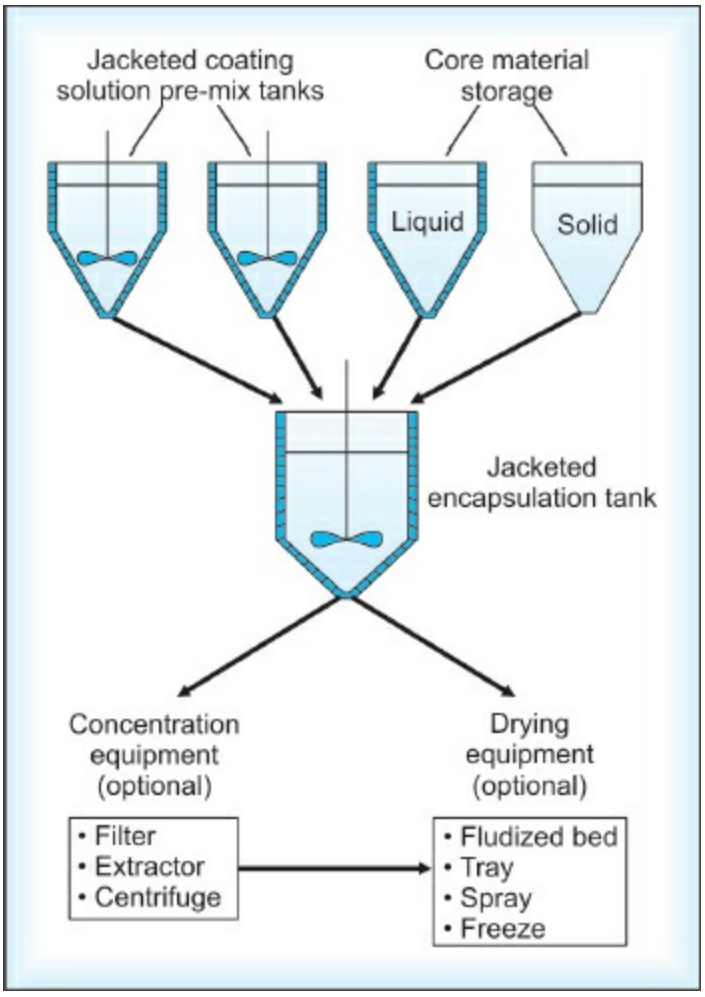



**Fig. 15.15:** Flow diagram of a typical phase-separation/coacervation process

## Multiorifice-centrifugal Process

The Southwest Research Institute has developed a mechanical process for producing microcapsules that utilizes centrifugal forces to hurl, a core material particle through an enveloping microencapsulation membrane, thereby effecting mechanical microencapsulation. The apparatus, illustrated cross-sectionally in Fig. 15.16, depicts a rotating cylinder, 1, a major and essential portion of the device. Located within the cylinder are three circumferential grooves, 2, 3, and 4. Countersunk in the intermediate groove, 3, are a plurality of orifices spaces closely and circumferentially around the cylinder. The upper and lower grooves, also located circumferentially around the cylinder, carry the coating material in molten or solution form, via tubes, 5, to the respective grooves. The ridges of the coating material grooves, 2 and 4, serve as a weir over which the coating material overflows when the volume of the upper and lower grooves is exceeded by the volume of material pumped into the system. The coating material, 6, under centrifugal force imparted by the cylinder rotation, flows outward along the side of the immediate groove into the countersunk portion and forms a film across the orifice. A counter rotating disc, 7, mounted within the cylinder, atomizes or disperses the core material fed through the centrally located inlet, 8. The rotating disc flings the particulate core material (liquid droplets or solid particles) toward the orifices. The core material arrives at the orifices and encounters the coating material membrane. The impact and centrifugal force, generated by the rotating cylinder, hurls the core material through the enveloping coating membrane, 9, which is immediately regenerated by the continually overflowing coating material.

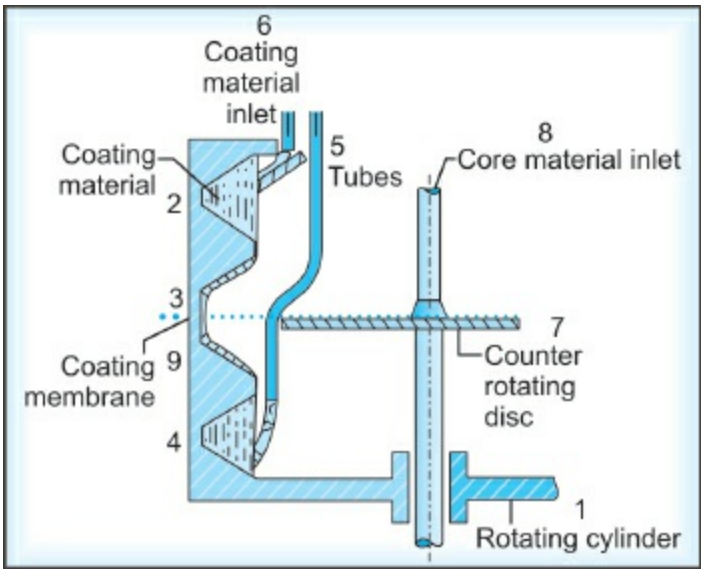



**Fig. 15.16:** Sectional diagram of multiorifice-centrifugal microencapsulation apparatus

The embryonic microcapsules, upon leaving the orifices, are hardened, congealed, or voided of coating solution by a variety of means. For example, the microcapsules can be flung into a heated, countercurrent air stream to harden or congeal coatings containing residual solvent. Also, the microcapsules can be forced into a rotating hardening or congealing bath. The coating material, if a melt, can be hurled into a cool liquid (nonsolvent for the coating material) decreasing the temperature below the melting point of the coating. Also, the hardening bath can contain a coating nonsolvent that is capable of extracting the coating solution solvent. The rotating hardening bath not only provides a coating desolvation or congealing function, but serves as a means of removing the microcapsules from their impact points, thus reducing agglomeration tendencies. It also provides a means of accumulating the coated product. The hardening liquid, after removing the microencapsulated product to where it can be collected, can be recycled to the hardening bath for subsequent reuse.

Processing variables include the rotational speed of the cylinder, the flow rate of the core and coating materials, the concentration and viscosity of the coating material, and the viscosity and surface tension of the core material. The multiorifice-centrifugal process is capable of microencapsulating liquids and solids (if the solids are dispersed in a liquid) of varied size ranges, with diverse coating materials (*see* Tables 15.2 and 15.3). The encapsulated

product can be supplied as a slurry in the hardening media or as a dry powder. Production rates of 50 to 75 pounds per hour have been achieved with the process.

## Pan Coating

The macroencapsulation of relatively large particles by pan methods has become widespread in the pharmaceutical industry, and the topic is covered in depth in Chapter 13 of this book. With respect to microencapsulation, solid particles greater than 600 microns in size are generally considered essential for effective coating, and the process has been extensively employed for the preparation of controlled-release beads. Medicaments are usually coated onto various spherical substrates such as nonpareil sugar seeds, and then coated with protective layers of various polymers.

In practice, the coating is applied as a solution, or as an atomized spray, to the desired solid core material in the coating pan. Usually, to remove the coating solvent, warm air is passed over the coated materials as the coatings are being applied in the coating pans. In some cases, final solvent removal is accomplished in a drying oven.

Blythe describes a method of preparing sustained-release pellets in which nonpareil seeds are coated initially with dextroamphetamine sulfate, and then with a release-rate retarding wax-fat coating. Nonpareil seeds (sugar pellets), 15.5 kg and 12 to 40 mesh in size, are placed in a rotating, 36-inch coating pan. USP syrup, 240 ml, is slowly poured onto the pellets to wet them evenly. An 80/20 mixture (750 g) of dextroamphetamine and calcium dihydrate are sprinkled onto the wetted nonpareil seeds. The pellets were dried with warm air. This coating operation is repeated three times. The fifth coating, talc, is accomplished by wetting the product with 240 ml of syrup followed by dusting 600 g of talc on the seeds. The pellets are rolled until dry, and the excess talc is removed by vacuum. The product is then screened through a 12-mesh screen and 20.0 kg of coated product is collected. One quarter of the batch is set aside, and the remainder is coated with a wax-fat coating solution consisting of 6300 g of glyceryl monostearate, 700 g of white beeswax, and 2100 ml of carbon tetrachloride, maintained at 70°C. The wax-fat solution, 425 ml is applied to the rotating pellets and subsequently dried with air. The coating operation is repeated until a 10% coating weight is achieved, whereupon one-third of the batch is removed.

The remaining product is again coated with the wax-fat solution as described previously. Subsequently, one half of the material is removed and the remainder is again coated to yield an additional coating of about 10% by

weight. The four groups of pellets are then thoroughly mixed to yield the sustained-release form of the sympathomimetic.

## Spray Drying and Spray Congealing

Spray-drying and spray-congealing methods have been used for many years as microencapsulation techniques. Because of certain similarities of the two processes, they are discussed together.

Spray-drying and spray-congealing processes are similar in that both involve dispersing the core material in a liquefied coating substance and spraying or introducing the core-coating mixture into some environmental condition, whereby relatively rapid solidification (and formation) of the coating is affected. The principal difference between the two methods, for the purpose of this discussion, is the means by which coating solidification is accomplished. Coating solidification in the case of spray drying is effected by rapid evaporation of a solvent in which the coating material is dissolved. Coating solidification in spray congealing methods, however, is accomplished by thermally congealing a molten coating material or by solidifying a dissolved coating by introducing the coating-core material mixture into a nonsolvent. Removal of the nonsolvent or solvent from the coated product is then accomplished by sorption, extraction, or evaporation techniques.

In practice, microencapsulation by *spray drying* is conducted by dispersing a core material in a coating solution, in which the coating substance is dissolved and in which the core material is insoluble, and then by atomizing the mixture into an air stream. The air, usually heated, supplies the latent heat of vaporization required to remove the solvent from the coating material, thus forming the microencapsulated product. The equipment components of a standard spray dryer include an air heater, atomizer, main spray chamber, blower or fan, cyclone and product collector, as described in detail in Chapter 4.

Process control variables include feed material properties such as viscosity, uniformity, and concentration of core and coating material, feed rate, method of atomization, and the drying rate, which is normally controlled by the inlet and outlet temperatures and the air stream solvent concentration. The process produces microcapsules approaching a spherical structure in the size range of 5 to 600 microns (Table 15.3). Characteristically, spray drying yields products of low bulk density, owing to the porous nature of the coated particles. Low active contents are normally required to provide the necessary

protection desired. For instance, the adequate retention of volatile, liquid core materials is difficult to achieve without maintaining low active content levels, perhaps below 20%.

Many coating materials (*see* Table 15.2) can be applied to liquid and solid core materials by spray drying coating solutions containing the dispersed core material. The process is commonly employed in the microencapsulation of liquid flavors yielding dry, free-flowing powders for use in foods and pharmaceuticals.

Microencapsulation by *spray congealing* can be accomplished with spray drying equipment when the protective coating is applied as a melt. General process variables and conditions are quite similar to those already described, except that the core material is dispersed in a coating material melt rather than a coating solution. Coating solidification (and microencapsulation) is accomplished by spraying the hot mixture into a cool air stream. Waxes, fatty acids and alcohols, polymers and sugars, which are solids at room temperature but meltable at reasonable temperatures, are applicable to spray-congealing techniques. Typically, the particle size of spray-congealed products can be accurately controlled when spray drying equipment is used, and has been found to be a function of the feed rate, the atomizing wheel velocity, dispersion of feed material viscosity, and other variables.

## Solvent Evaporation

This technique has been used by companies including The NCR Company, Gavaert PhotoProduction NV, and Fuji Photo Film Co., Ltd. to produce microcapsules. The processes are carried out in a liquid manufacturing vehicle. The microcapsule coating is dissolved in a volatile solvent, which is immiscible with the liquid manufacturing vehicle phase. A core material to be microencapsulated is dissolved or dispersed in the coating polymer solution. With agitation, the core coating material mixture is dispersed in the liquid manufacturing vehicle phase to obtain the appropriate size microcapsule. The mixture is then heated (if necessary) to evaporate the solvent for the polymer. In the case in which the core material is dispersed in the polymer solution, polymer shrinks around the core. In the case in which the core material is dissolved in the coating polymer solution, a matrix-type microcapsule is formed. Once all the solvent for the polymer is evaporated, the liquid vehicle temperature is reduced to ambient temperature (if required) with continued agitation. At this stage, the microcapsules can be used in suspension form, coated on to substrates or isolated as powders.

Process variables would include, but not be limited to, methods of forming dispersions, evaporation rate of the solvent for the coating polymer, temperature cycles, and agitation rates. Important factors that must be considered when preparing microcapsules by solvent evaporation techniques include choice of vehicle phase and solvent for the polymer coating, as these choices greatly influence microcapsule properties as well as the choice of solvent recovery techniques.

The solvent evaporation technique to produce microcapsules is applicable to a wide variety of liquid and solid core materials. The core materials may be either water-soluble or water-insoluble materials. A variety of filmforming polymers can be used as coatings as exemplified in Table 15.2.

## Polymerization

A relatively new microencapsulation method utilizes polymerization techniques to form protective microcapsule coatings in situ. The methods involve the reaction of monomeric units located at the interface existing between a core material substance and a continuous phase in which the core material is dispersed. The continuous or core material supporting phase is usually a liquid or gas, and therefore the polymerization reaction occurs at a liquid-liquid, liquid-gas, solid-liquid, or solid-gas interface.

The polymerization method most applicable, perhaps, to pharmaceutical or medical use is that developed by Chang in his research at McGill University. Chang has been able to accomplish permselective membrane properties for microcapsules having coatings of nylon formed by interfacial polymerization, or collodion formed by phase-separation/coacervation techniques. The membranes, typically about 200 Å thick, have an equivalent aqueous pore radius of about 16 Å. The microcapsules have been shown to be permselective in that protein and enzyme core materials, for instance, do not transfer out of the microcapsule, whereas smaller molecules such as enzyme substrates and resultant reaction products can permeate the membrane.

Chang's interfacial polymerization method for forming polyamide (nylon) membranes involves the reaction occurring at the liquid-liquid interface existing between an aqueous solution of an aliphatic diamine and a water-immiscible organic solution of a dicarboxylic acid halide. The polymerization reaction depends on the fact that acid halides, such as sebacoyl chloride, are nearly water-insoluble, and diamines, such as hexanediamine, have an appreciable partition coefficient toward the water-immiscible organic phase. Hence, the hexanediamine diffuses to the organic sebacoyl chloride phase, and the polycondensation reaction occurs forming the polyamide. Because the chemical reaction rate exceeds the diffusion rate of the diamine into the nonaqueous phase, the polyamide is deposited almost entirely at the interface existing between the two solutions.

Using this phenomenon, Chang prepares microcapsules containing protein solutions by incorporating the protein in the aqueous diamine phase. Chang has demonstrated the permselectivity of microcapsules containing the enzyme, urease, by their ability to convert blood urea to ammonia, the enzyme remaining within the microcapsules when incorporated within an

extracorporeal shunt system. Numerous companies including National Lead Corporation, Union Carbide Corporation, Pennwalt Corporation, Eurand America Inc., Appleton Papers Inc., and Moore Business Forms, Inc. are utilizing polymerization techniques to accomplish microencapsulation.

## EQUIPMENT AND PROCESSING

The equipment required to conduct microencapsulation varies from complex machines designed specifically for microencapsulation to rather simple processing equipment common to many laboratories. The variation of microencapsulation equipment is evidenced by the descriptions included in the above “Methodology” section.

Microcapsules as bulk materials, in either dry powder or dispersed form, can be processed into final product applications using common equipment such as V-blenders, tablet machines, granulators, homogenizers, kneaders, hard-gelatin capsule filling machines, or coating equipment if deposition onto a substrate is desired. The specific processing equipment employed depends on the final product form desired and on the microcapsule properties. All processing and formulation operations must be conducted with continual caution to avoid possible adverse effects as a result of rupture, attrition, or dissolution of the thin microcapsule coating.

# 16:
# Sustained Release Oral Dosage Forms

With many drugs, the basic goal of therapy is to achieve a steady-state blood or tissue level that is therapeutically effective and nontoxic for an extended period of time. The design of proper dosage regimens is an important element in accomplishing this goal. A basic objective in dosage form design is to optimize the delivery of medication so as to achieve a measure of control of the therapeutic effect in the face of uncertain fluctuations in the in vivo environment in which drug release takes place. This is usually accomplished by maximizing drug availability, i.e. by attempting to attain a maximum rate and extent of drug absorption; however, control of drug action through formulation also implies controlling bioavailability to reduce drug absorption rates.

The United States Pharmacopoeia defines the extended-release (ER) dosage form as the one that allows at least a 2-fold reduction in dosing frequency or significant increase in patient compliance or therapeutic performance when compared with that presented as a conventional dosage form (a solution or a prompt drug-releasing dosage form). The terms “controlled release (CR)”, “prolonged release”, “sustained or slow release (SR)” and “long-acting (LA)” have been used synonymously with “extended release”. The commercial branded products in this category are often designated by suffixes such as CR, CD (controlled delivery), ER, LA, PD (programmed or prolonged delivery), SA (slow-acting), SR, TD (timed delivery), TR (timed release), XL and XR (extended release).

## Sustained Release Concept

Sustained release, sustained action, prolonged action, controlled release, timed release, depot, and repository dosage forms are terms used to identify drug delivery systems that are designed to achieve a prolonged therapeutic effect by continuously releasing medication over an extended period of time after administration of a single dose. In the case of injectable dosage forms, this period may vary from days to months. In the case of orally administered forms, however, this period is measured in hours and critically depends on the residence time of the dosage form in the gastrointestinal (GI) tract. The term "*controlled release*" has become associated with those systems from which therapeutic agents may be automatically delivered at predefined rates over a long period of time. Products of this type have been formulated for oral, injectable, and topical use, and include inserts for placement in body cavities as well.

The pharmaceutical industry provides a variety of dosage forms and dosage levels of particular drugs, thus enabling the physician to control the onset and duration of drug therapy by altering the dose and/or mode of administration. In some instances, control of drug therapy can be achieved by taking advantage of beneficial drug interactions that affect drug disposition and elimination, e.g. the action of probenicid, which inhibits the excretion of penicillin, thus prolonging its blood level. Mixtures of drugs might be utilized to potentiate, synergize, or antagonize given drug actions. Alternately, drug mixtures might be formulated in which the rate and/or extent of drug absorption is modified. Sustained release dosage form design embodies this approach to the control of drug action, i.e. through a process of either drug modification or dosage form modification, the absorption process, and subsequently drug action, can be controlled.

### *Advantages*

Physicians can achieve following desirable therapeutic advantages by prescribing sustained release forms:

1. Since the frequency of drug administration is reduced, patient compliance can be improved, and drug administration can be made more convenient as well.

2. The blood level oscillation characteristic of multiple dosing of

conventional dosage forms is reduced, because a more even blood level is maintained.

3. A less obvious advantage, implicit in the design of sustained release forms, is that the total amount of drug administered can be reduced, thus maximizing availability with a minimum dose.
4. In addition, better control of drug absorption can be attained, since the high blood level peaks that may be observed after administration of a dose of a high-availability drug can be reduced by formulation in an extended action form.
5. The safety margin of high-potency drugs can be increased, and the incidence of both local and systemic adverse side effects can be reduced in sensitive patients.
6. Overall, administration of sustained release forms enables increased reliability of therapy.

## *Disadvantages*

In evaluating drugs as candidates for sustained release formulation, the disadvantages of such formulations that must be considered include the following:

1. Administration of sustained release medication does not permit the prompt termination of therapy. Immediate changes in drug need during therapy, such as might be encountered if significant adverse effects are noted, cannot be accommodated.
2. The physician has less flexibility in adjusting dosage regimens. This is fixed by the dosage form design.
3. Sustained release forms are designed for the normal population, i.e. on the basis of average drug biologic half-lives. Consequently, disease states that alter drug disposition, significant patient variation, and so forth are not accommodated.
4. Economic factors must also be assessed, since more costly processes and equipment are involved in manufacturing many sustained release forms.

Not all drugs are suitable candidates for formulation as prolonged action medication. Table 16.1 lists specific drug characteristics that would preclude formulation in peroral sustained release forms. Drugs with long biologic half-

lives (e.g. digoxin – 34 hours) are inherently long-acting and thus are viewed as questionable candidates for sustained release formulation. For some drugs in this group, however, a properly designed sustained release formulation may be advantageous. Because single doses capable of producing equally prolonged effects often yield significant concentration peaks immediately after each dosing interval, control of drug release may be indicated if toxicity or local gastric irritation is a hazard. Drugs with narrow requirements for absorption (e.g. drugs dependent on position in the GI tract for optimum absorption) are also poor candidates for oral sustained release formulation, since absorption must occur throughout the length of the gut. Very insoluble drugs whose availability is controlled by dissolution (e.g. griseofulvin) may not benefit from formulation in sustained release forms since the amount of drug available for absorption is limited by the poor solubility of the compound.

**Table 16.1:** Characteristics of drugs unsuitable for peroral sustained release forms

| **Characteristics** | **Drugs** |
|---|---|
| Not effectively absorbed in the lower intestine | Riboflavin, ferrous salts |
| Absorbed and excreted rapidly; short biologic half-lives (< 1 hr) | Penicillin G, furosemide |
| Long biologic half-lives (>12hr) | Diazepam, phenytoin |
| Large doses required (>1 g) | Sulfonamides |
| Cumulative action and undesirable side effects; drugs with low therapeutic indices | Phenobarbital, digitoxin |
| Precise dosage titrated to individual is required | Anticoagulants, cardiac glycosides |
| No clear advantage for sustained release formulation | Griseofulvin |

Before proceeding with the design of a sustained release form of an appropriate drug, the formulator should have an understanding of the

pharmacokinetics of the candidate, should be assured that pharmacologic effect can be correlated with drug blood levels, and should be knowledgeable about the therapeutic dosage range, including the minimum effective and maximum safe doses.

# THEORY

## Design and Fabrication

To establish a procedure for designing sustained release dosage forms, it is useful to examine the properties of drug blood-leveltime profiles characteristic of multiple dosing therapy of immediate release forms. Figure 16.1 shows typical profiles observed after administration of equal doses of a drug using different dosage schedules: every 8 hours (curve A), every 3 hours (curve B), and every 2 hours (curve C). As the dosage interval is shortened, the number of doses required to attain a steady-state drug level increases, the amplitude of the drug level oscillations diminishes, and the steady state average blood level is increased. As a first approximation, the optimum dosage interval can be taken to be equal to the biologic half-life, in this case, 3 hours. Curve D represents a profile in which the first or loading dose is made twice that of all subsequent doses administered, i.e. the maintenance doses. This dosing regimen allows the relation between the loading ($D$) and maintenance ($D_m$) doses to be determined as follows:

$$D_i = D_m\,(1 - \exp^{-0.693\tau/t_{1/2}})$$

where, τ is the dosing interval and $t_{1/2}$ is the biologic half-life. If $t_{1/2} = \tau$, $D_i = 2D_m$. Selection of the proper dose and dosage interval is a prerequisite to obtaining a drug level pattern that will remain in the therapeutic range.

Elimination of drug level oscillations can be achieved by administration of drug through constant-rate intravenous infusion. Curve E in Fig. 16.1 represents an example whereby the infusion rate was chosen to achieve the same average drug level as a 3 hour dosage interval for the specific case illustrated. The objective in formulating a sustained release dosage form is to be able to provide a similar blood level pattern for up to 12 hours after oral administration of the drug. To design an efficacious sustained release dosage form, one must have a thorough knowledge of the pharmacokinetics of the drug chosen for this formulation. Figure 16.2 shows a general pharmacokinetic

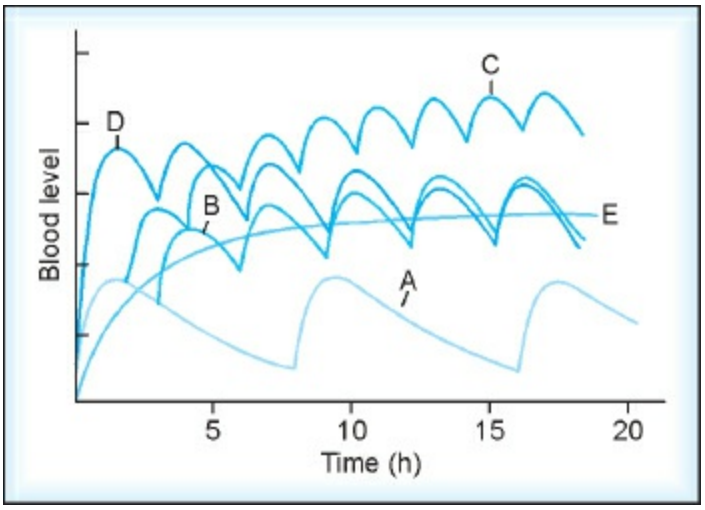



**Fig. 16.1:** Multiple patterns of dosage that characterize nonsustained peroral administration of a drug with a biologic half-life of 3 hr and a half-life for absorption of 20 min. Dosage intervals are: A, 8 hr; B, 3 hr; C, 2 hr; and D, 3 hr (loading dose is twice the maintenance dose). E, Constant rate intravenous infusion model of an ideal sustained release dosage form. For the purposes of this discussion, measurements of drug blood level are assumed to correlate with therapeutic effect and drug kinetics are assumed to be adequately approximated by a one-body-compartment model. That is, drug distribution is sufficiently rapid so that a steady state is immediately attained between the central and peripheral compartments, i.e. the blood-tissue transfer rate constants, $k_{12}$ and $k_{21}$, are large.

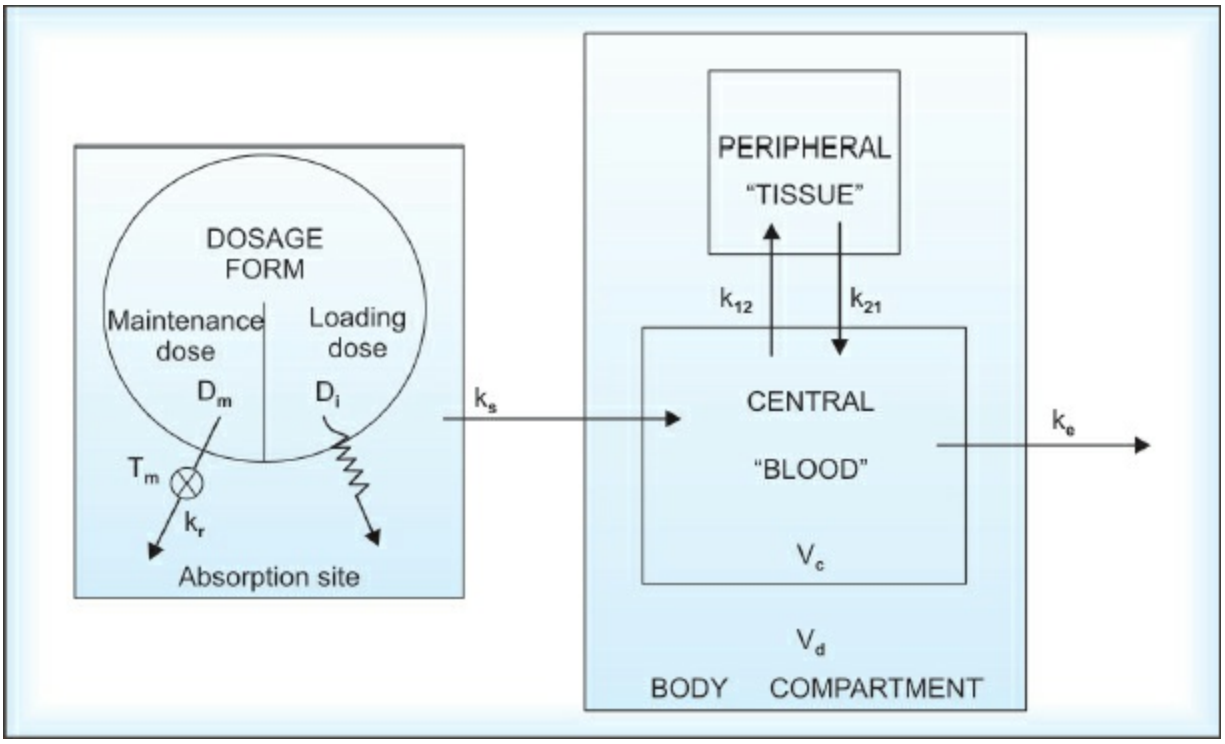



**Fig. 16.2:** A general pharmacokinetic model of an ideal peroral sustained release dosage form

Under the foregoing circumstances, the drug kinetics can be characterized by three parameters: the elimination rate constant ($k_e$) or biologic half-life ($t_{1/2}$= 0.693/$k_e$), the absorption rate constant ($k_a$), and the apparent distribution volume ($V_d$), which defines the apparent body space in which drug is distributed. A large $V_d$ value (e.g. 100 L) means that drug is extensively distributed into extravascular space: a small $V_d$ value (e.g. 10 L) means that drug is largely confined to the plasma. It is best interpreted as a proportionality factor which when multiplied by the blood level gives the total amount of drug in the body. For the two-body-compartment representation of drug kinetics, $V_c$ is the volume of the central compartment, including both blood and any body water in which drug is rapidly perfused.

A diagrammatic representation of a dosage form, which identifies the specific parameters that must be taken into account in optimizing sustained release dosage form designs, is shown in Fig. 16.2 at the absorption site. These are the loading or immediately available portion of the dose ($D_i$), the maintenance or slowly available portion of the dose ($D_m$), the time ($T_m$) at

which release of maintenance dose begins (i.e. the delay time between release of $D_i$ and $D_m$), and the specific rate of release ($k_r$) of the maintenance dose.

Figure 16.3 shows the form of the body drug-level time profile that characterizes an ideal peroral sustained release dosage form after a single administration. $T_p$ is the peak time, and $h$ is the total time after administration in which the drug is effectively absorbed. $C_p$ is the average drug level to be maintained constantly for a period of time equal to ($h - T_p$) hours; it is also the peak blood level observed after administration of a loading dose. The portion of the area under the blood level curve contributed by the loading and maintenance doses is indicated on the diagram. To obtain a constant drug level, the rate of drug absorption must be made equal to its rate of elimination. Consequently, drug must be provided by the dosage form at a rate such that the drug concentration becomes constant at the absorption site.

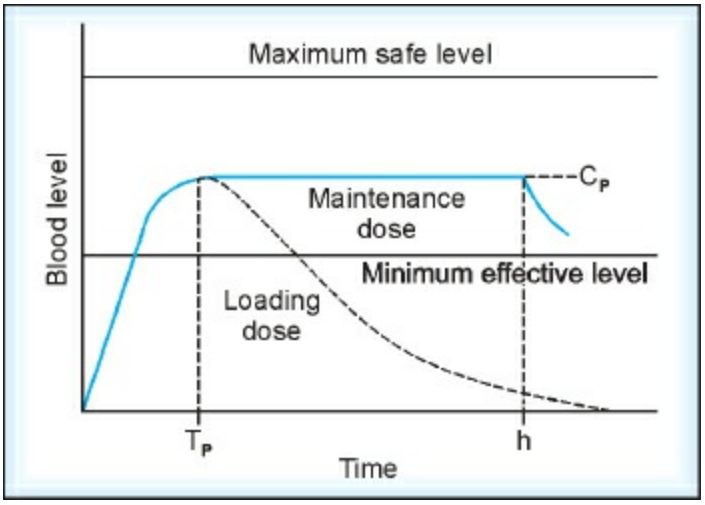



**Fig. 16.3:** A blood-level time profile for an ideal peroral sustained release dosage form

Detailed theoretic treatments of a number of sustained release dosage form designs have been reported. These include systems in which drug is released for absorption by zero-order and first-order processes with and without loading doses. In the former case, designs based on both immediate and delayed release of maintenance dose have been described. The following general assumptions have been made in developing these designs: (1) Drug disposition can be described by a one-compartment open model. (2) Absorption is first-order and complete. (3) Release of drug from the dosage

form, not absorption, is ratedetermining, i.e. the effect of variation in absorption rate is minimized ($k_a > k_e$).

## Zero-order Release Approximation

The profile shown in Fig. 16.3 can most nearly be approximated by a design consisting of a loading dose and a zero-order release maintenance dose, as described by Robinson and Eriksen. If a zero-order release characteristic can be implemented in a practical formulation, the release process becomes independent of the magnitude of the maintenance dose and does not change during the effective maintenance period. Table 16.2 lists the expressions that can be used to estimate the design parameters for an optimized zero-order model, for both simultaneous and delayed release of maintenance dose. Their application is illustrated using procainamide, an important antiarrhythmic agent, as an example.

Table 16.3 lists the pharmacokinetic parameters characterizing the disposition of procainamide, which is described by a two-body-compartment open model, in an average patient based on data reported by Manion et al. for 11 subjects. Conventional formulations are administered every 3 hours for maintenance of therapy, resulting in a maximum-to-minimum blood level ratio >2 at the steady state. Sustained release formulations have been shown to have advantages as an alternate dosage form. A comparison is made between estimates based on three cases (1) the one-compartment model assumption with delayed release of maintenance dose, (2) the actual two-compartment fit of procainamide pharmacokinetic data with delayed release of maintenance dose and (3) the two-compartment model with simultaneous release of loading and maintenance doses. In all cases, the blood level is assumed to be maintained at 1 µg/ml for 8 hours, i.e. $C_p$ = 1 mg/L, and $h - T_p$ = 8. Table 16.4 summarizes the results of the calculation of sustained release design parameters for procainamide, assuming zero-order release kinetics. The following steps are required to estimate the design parameters listed in the table. (Equation numbers refer to equations in Table 16.2.)

**Table 16.2:** Expressions useful for estimation of design parameters for zero-order sustained release dosage form models

| Parameter | Equation | |
|---|---|---|
| Maximum body drug content to be maintained | $A_m = C_p V_d$ | Eq.(1) |

| | | |
|---|---|---|
| Zero-order rate constant | $k_{r0} = k_e A_m$ | Eq.(2) |
| Peak time | $T_p = [2.3/k_a - k_e)] \log(k_a/k_e)$ | Eq.(3) |
| Bioavailability factor | $F = (AUC)_{oral}/(AUC)_{iv}$ | Eq.(4) |
| Fraction of dose ($D_i$) at peak ($F = 1$) | $f = \left(\frac{k_a}{k_e}\right)^{\frac{k_e}{k_e - k_a}}$ | Eq.(5) |
| Maintenance dose | $D_m = k_{r0}(h - T_m)/F$ | Eq.(6) |
| Loading dose ($T_m = T_p$) | $D_i, = A_m/fF$ | Eq.(7) |
| Loading dose ($T_m = 0$) | $D_i = (D_i - k_{r0}T_p)$ | Eq.(8) |

**Table 16.3:** Pharmacokinetic parameters for procainamide in an average subject (weight: 75 kg)

| Parameter | Value | Parameter | Value |
|---|---|---|---|
| $\beta$ | 0.21 hr | $k_{12}$ | 3.15 hr |
| $t_{1/2}$ | 3.4 hr | $k_{21}$ | 1.4 hr |
| $k_e$ | 0.97 hr | $V_c$ | 59 L |
| $k_a$ | 2.0 hr | $V_d$ | 205 L |
| $F$ | 0.83 | $T_p$ | 0.5 hr |

**Table 16.4:** Estimated sustained release design parameters for procainamide ($C_p$ = 1 µg/ml, $h - T_m$ = 8 hours)

| Parameter | Case 1 | Case 2 | Case 3 |
|---|---|---|---|
| $A_m$ (mg) | 205 | 59 | 59 |
| $k_{r0}$ (mg/hr) | 43 | 57 | 57 |
| $T_m$ | 1.2 | 0.5 | 0 |
| $D_m$ (mg) | 414 | 549 | 549 |
| $F$ | 0.768 | 0.274 | 0.274 |
| $D_i$ (mg) | 322 | 259 | ($Di$) 224 |
| $D_m/D_i$ | 1.28 | 2.12 | 2.4 |
| $D_m + D_i$ (mg) | 736 | 808 | 773 |

**1. Estimation of $k_{r0}$:** Eq. (2) is derived by considering that at the steady state the rate of absorption is constant and equal to the rate of elimination, that is: Rate Absorption = $k_a . X_a$ = Rate Elimination = $k_e . X_b$ where $X_a$ is the amount of drug at the absorption site, and $X_b$ is the body drug content, which

is set equal to $V_d$. $C_p$, or $A_m$ in Eq. (1), the body drug content to be maintained constant. If absorption of the loading dose is effectively complete, $k_a X_a = k_{r0}$. For case 1 in Table 16.4, $k_e$ = 0.21 (the beta disposition constant), since the biologic halflife is estimated from the terminal part of the blood level curve if a one-compartment model is used to approximate blood level data. For cases 2 and 3, $k_e$ = 0.97. The apparent volume of the central compartment, $V_c$, rather than $V_d$, is used to calculate $A_m$ for the two-compartment model, i.e. $A_m = C_p V_c$. For example:

Case 1: $k_{r0} = 0.21 \times 1 \times 205$
$= 43$ mg/hr

Cases 2 and 3: $k_{r0} = 0.97 \times 1 \times 59$
$= 57$ mg/hr

**2. Estimation of $T_m$:** Release of maintenance dose is set at the peak time for the loading dose (cases 1 and 2). Eq. (3) is used to calculate the peak time from known value absorption (2 $hr^{-1}$) and elimination (0.21 $hr^{-1}$) rate constants. Since Eq. (3) applies only to the one-compartment model, $T_m$, which is actually 0.5 hours, is significantly overestimated. For example:

$$\text{Case 1: } T_m = T_p \text{ (Eq. 3)} = \frac{2.3 \times \log(2/0.21)}{2 - 0.21}$$
$$= 1.2 \text{ hr}$$
$$\text{Case 2: } T_m = T_p \text{ (actual value)} = 0.5 \text{ hr}$$
$$\text{Case 3: } T_m = 0$$

**3. Estimation of $D_m$:** The maintenance dose is estimated as the product of release rate and maintenance time (Eq. 6), corrected for the bioavailability factor, F (Eq. 4), which is the fraction of the administered dose absorbed from a reference nonsustained release dosage form. The F-value is estimated as the ratio of the area under the plasma level curve (AUC-value) measured after oral administration to the *AUC*-value observed after intravenous administration of the same dose of drug. In the example, F < 1, since procainamide is subject to the first-pass effect, in which a small portion of the absorbed dose is metabolized in the liver. $D_m$ is also a function of the loading dose and an inverse function of the biological half-life, i.e. $D_m = 0.693 f(h - T_m) D_i / t_{1/2}$, a relation obtained by combining Eqs (2), (6), and (7). Practically,

h is not likely to exceed 10 to 12 hours, depending on the residence time in the small intestine. For drugs that are not efficiently absorbed in the stomach, such as procainamide, the gastric emptying rate is an uncertain variable that contributes to $T_m$. For example:

Case 1: $D_m = 43 \times 8/0.83 = 414$ mg
Cases 2 and 3: $D_m = 57 \times 8/0.83 = 549$ mg

Significant increases in dose size are required for drugs with short biologic halflives, e.g. $D_m$ is doubled if the biologic halflife is halved. For case 1, $D_m$ would be 228 mg for $t_{1/2}$ = 6 hr, 456 mg for $t_{1/2}$ = 3 hr, and 685 mg for $t_{1/2}$ = 2 hr.

**4. Estimation of $D_i$:** The loading dose is that portion of the total dose that is initially released as a bolus and is therefore immediately available for absorption. It results in a peak blood level equal to the desired level to be maintained. Eq. (7) allows estimation of $D_i$ if $D_m$ is delayed (cases 1 and 2). If release of $D_m$ is not delayed (case 3), the loading dose calculated using Eq. (7) is adjusted for the quantity of drug provided by the zero-order release process in time $T_m$ as shown by Eq. (8). For example:

Case 1: $D_i = 205/(0.768 \times 0.83) = 322$ mg
Case 2: $D_i = 59/(0.274 \times 0.83) = 259$ mg
Case 3: $D_i = (259 - 57 \times 0.5) = 224$ mg

Figure 16.4 shows the simulated blood level profiles that result from administration of theoretic sustained release dosage forms of procainamide to the average subject for the three cases listed in Table 16.4. Curve A is the profile observed after administration of the loading dose calculated for case 2. Calculations based on the assumption of a one-compartment model (curve B) fail to approximate the desired profile adequately. The procedure suggested for estimation of $k_{r0}$, however, based on the actual two-compartment model that fits procainamide data, gives a reasonable approximation of the optimum profile (curve C). A formulation designed to release loading and maintenance doses simultaneously (case 3) results in a profile (curve D) that does not significantly differ from case 2. The total dose required to maintain a blood level of 1 µg/ml for 8 to 10 hours is about the maximum (<1 g) that can be formulated in a reasonably sized solid peroral dosage form. The usual

minimum therapeutic level required for procainamide is 3 to 4 µg/ml. Multiple units of a sustained release procainamide would have to be administered at each dosing interval to attain a therapeutic level.

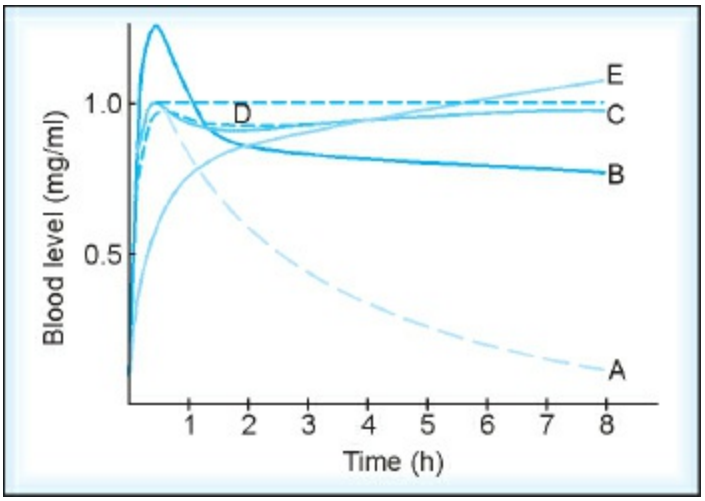



**Fig. 16.4:** Simulated blood level profiles observed after administration of theoretic sustained release formulations of procainamide hydrochloride to an average patient. A, Case 2—loading dose; B, Case 1; C, Case 2; D, Case 3; E, Case 2—patient differs from average

Computer simulation provides a valuable tool for evaluating the performance of sustained release dosage form designs. Curve E in Fig. 16.4 demonstrates another application of simulation, that is, to examine the performance of the dosage form in a patient in which the disposition of the drug (procainamide in the example) differs significantly from the average. In this subject, the pharmacokinetic parameters were as follows: $k_a$ = 1.2, $k_e$ = 0.47, $k_{12}$ = 0.8, $k_{21}$ = 0.77, $V_c$ = 101, and $F$ = 0.7. Lower blood levels are observed initially, and higher blood levels are observed at the end of the maintenance period, since the absorption rate was lower and the biologic half-life higher (approximately 4 hours) than average in this patient. Overall, the difference in response of this patient to the dosage form is not significant.

## First-order Release Approximation

The rate of release of drug from the maintenance portion of the dosage form should be zero-order if the amount of drug at the absorption site is to remain constant. Most currently marketed sustained release formulations, however, do not release drug at a constant rate, and consequently do not maintain the relative constant activity implied by Fig. 16.3. Observed blood levels decrease over time until the next dose is administered. In many instances, the rate of appearance of drug at the absorption site can be approximated by an exponential or first-order process in which the rate of drug release is a function only of the amount of drug remaining in the dosage form. Table 16.5 lists the expressions that can be used to estimate the design parameters for optimized first-order release models. Three different designs are considered: $D_m$ not delayed. $D_m$ delayed where, $T_m = T_p$, and $D_m$ delayed where $T_m > T_p$. Table 16.6 lists the parameters calculated for a drug fitted by a one-body-compartment model, and Fig. 16.5 shows the resulting profiles for each example considered. Doses listed in Table 16.6 are expressed as fractions of loading dose, using the calculation for a zero-order model (case 1, Table 16.4) as a reference.

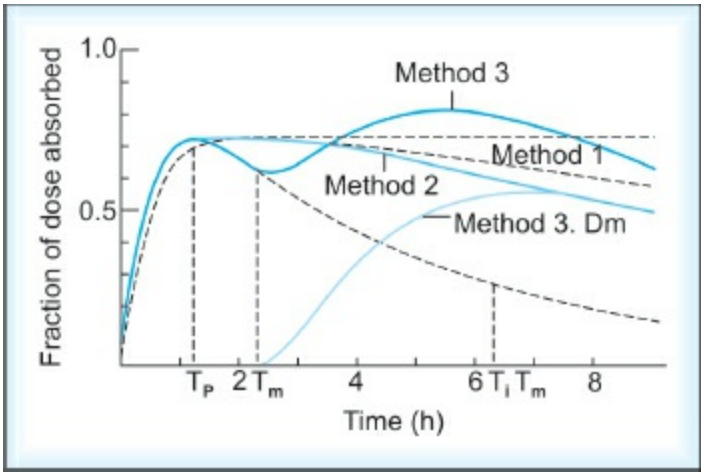



**Fig. 16.5:** Simulated blood level profiles observed after administration of a theoretic sustained release dosage form to an average patient based on different first-order release models. Blood level is plotted as the fraction of dose absorbed (*CpVd/FDi*)

**Table 16.6:** Estimated design parameters for a first-order sustained release

| model* | | | | |
|---|---|---|---|---|
| **Parameter** | **Zero-order** | **Method 1** | **Method 2** | **Method 3** |
| $T_m$ | 1.2 | 0 | 1.2 | 2.3 |
| $T_i$ | – | 8 | – | 4 |
| $k_{r1}$ | ($k_{r0}$ = 0.173) | 0.055 | 0.125 | 0.27 |
| $D_m/D_i$ | 1.4 | 3.1 | 1.4 | 1.45 |
| $D_i$ | 1.0 | 0.8 | 1.0 | 1.0 |
| $(D_i + D_m)/D_i$ | 2.4 | 3.9 | 2.4 | 2.45 |
| * Drug Characteristics: One-Compartment Model | | | | |
| $t_{1/2}$ = 3hr | $k_a$ = 2hr | $k_e$ = 0.23 hr | $f$ = 0.75 | |
| $T_p$ = 1.2 hr | $h$ = 9.2 hr | $F$ = 1 | $CpVd$ = 0.75 | |

**Method 1. Simultaneous release of $D_m$ and $D_i$:** The *crossing time*, $T_i$, is the time at which the blood level profiles produced by administration of separate loading and maintenance doses intersect. The closest approximation to the ideal profile is obtained if the crossing point is made at least equal to the desired maintenance period ($h - T_p$). Eq. (9) shown in Table 16.5, is an approximation of Eq. (11) where, $k_e > k_{r1}$. The maintenance dose is estimated from the initial release rate, i.e. $k_{r1}D_m = k_eA_m = k_{r0}$. The loading dose is estimated by correcting the immediate release dose required to achieve the maintenance level for the quantity of drug delivered by the maintenance dose in the time $T_p$. For example:

$T_i = 9.2 - 1.2 = 8$ hr

$k_{r1} = 0.23 \exp(-0.23 \times 8) = 0.055$ hr

$D_m = 0.173/0.055 = 3.1$

$Di = (0.75/0.75 \times 1) - 0.173 \times 1.2 = 0.8$

**Table 16.5:** Expressions useful for estimation of design parameters for first-order sustained release dosage form models

| Parameter | Method 1 | Method 2 | Method 3 |
|---|---|---|---|
| $T_m$ | 0 | $T_p$ (Eq. 3) | $4.6/k_a$ (Eq. 10) |
| $T_i$ | $(h - T_p)$ | — | $(h - T_p)/2$ |
| $k_{r1}$ | $k_e[\exp(-k_eT_i)]$ (Eq. 9) | $1/(h - T_p)$ | $T_i = \frac{2.3 \log(k_{r1}/k_e)}{(k_{r1} - k_e)}$ (Eq. 11) |
| $D_m$ | $k_{r0}/k_{r1}$ | $k_{r0}/k_{r1}$ | $D_m = \frac{k_e(A_m - A_i)}{k_{r1}} \exp[k_{r1}(2T_i - T_m)]$ (Eq. 12) |
| | | | $A_i = \frac{D_i k_a}{(k_a - k_e)} \exp[-k_e(T_i + T_m)]$ (Eq. 13) |
| $D_i$ | $C_pV_d/fF - k_{r0}T_p$ | $C_pV_d/fF$ | $C_pV_d/fF$ |

**Method 2. Delayed release of $D_m$:** $T_m = T_p$. If $D_m$ is large and $k_r$ is made small, maintenance dose may be released as a pseudozero-order process. As a first approximation, $k_{r1}$ may be estimated as the reciprocal of the maintenance time. $D_m$ is then calculated as in method 1. Better approximation of a zero-order response can be obtained if $D_m$ is increased and $k_{r1}$ is reduced to maintain the product $k_{r1}D_m$ constant. For example:

$T_m = T_p = 1.2$ hr
$k_{r1} = 1/(9.2 - 1.2) = 0.125$ hr
$D_m = 0.173/0.125 = 1.4$
$Di = 0.75/0.75 \times 1 = 1$

Since $k_{r1}D_m = 0.173$, then $k_{r1}$ should be reduced to 0.86 if $D_m$ is increased to 2.0, to maintain this product constant.

**Method 3. Delayed release of $D_m$:** $T_m > T_p$. Increasing the delay time, $T_m$, allows the use of faster release rates. A period equal to the time at which 99% of the loading dose has been absorbed is selected using Eq. (10) in Table 16.5. The release rate constant is iteratively calculated from Eq. (11) such that a peak is obtained from the maintenance dose at the midpoint of the maintenance time. The amount of drug required to produce a second peak at this time is the maintenance dose, calculated from Eqs (12, 13). For example:

$T_m = 4.6/2.0 = 2.3$ hr
$T_i = (9.2 - 1.2)/2 = 4$ hr
$4 = 2.3 \times \log (k_{r1}/0.23)/(k_{r1} - 0.23)$

(Solve this expression iteratively by finding the value of $k_{r1}$ that satisfies the equality: in this case, $k_{r1} = 0.27$.)

$$
\begin{aligned}
D_i &= 0.75/0.75 \times 1 = 1 \\
A_i &= (1 \times 2)\exp(-0.23(4 + 2.3))/(2.0 - 0.23) = 0.267 \\
A_m &= k_{r0}/k_e = 0.173/0.23 = 0.74 \\
D_m &= (0.23/0.27)(0.74 - 0.267)\exp(0.27(2 \times 4 - 2.3)) = 1.45
\end{aligned}
$$

Methods 1 and 2 have the disadvantage that large maintenance doses are required, resulting in a significant loss of drug available for absorption. In the examples plotted in Fig. 16.5, 50% of the dose calculated using method 1 and 30% of the dose calculated using method 2 were not available after 10 hours. Method 1 represents a design that is least efficient in terms of the dose required to achieve the design objective. Drugs characterized by large loading doses and small biologic half-lives could not be practically formulated in these designs. In spite of the fluctuation observed in the profile characterizing method 3, a reasonably average level is maintained with a minimum maintenance dose. Furthermore, only 15% of the dose remains unreleased after 10 hours. The maximum dose, which usually should be formulated in a sustained release dosage form, should not exceed the total dose administered by conventional forms during the maintenance period. If the total dose is released all at once, it should not result in a blood level exceeding the maximum safe level.

Optimization of sustained release dosage form design requires minimization of the total dose delivered and maximization of the duration of drug release. In the case of the zero-order model, adjustment of the maintenance dose is a function only of the duration time. Designs based on simultaneous release of $D_m$ and $D_i$, requires the minimum total dose (e.g. case 3, Table 16.4). Optimum first-order designs are those in which $D_m$ is delayed beyond the peak obtained from the loading dose (e.g. method 3, Fig. 16.5).

From a practical point of view, the procedures summarized here provide a starting point for implementation. The use of mean values for parameters and the inherent variation in the population using the dosage form often lead to larger variations than the errors resulting from the approximate nature of

these calculations. Where therapeutic effect cannot be correlated with measured body drug content, or where such measurements cannot be obtained, a clinically determined dosage regimen can be used as a basis for estimation of $D_m$-values. One may assume that $D_m$ equals the sum of the maintenance doses ordinarily administered in the desired maintenance time. For example, for a drug administered 4 times a day, $D_i$, is the single dose and $D_m$ the sum of two doses for a maintenance period of 12 hours. As a rule of thumb, the total drug dose can be reduced from 5 to 10%.

## Multiple Dosing

Like conventional dosage forms, peroral sustained release forms are administered as multiple dosing regimens in which the objective is to maintain the required average drug level for the duration of therapy with a minimal fluctuation between doses. If the dosing interval is made equal to (or less than) the total anticipated drug release time (Fig. 16.6, curve A), accumulation results from formulations designed with loading doses. Significant blood level peaks may be observed with the zero-order release model; however, minimal fluctuation between administered doses can be obtained if the dosing interval is set equal to ($h + t_{1/2}$), as shown by curve B (Fig. 16.6). Increasing the dosing interval further while diminishing the peak also deepens the trough in the drug level profile, defeating one of the objectives of the dosage form design. Even with the zero-order model, multiple dosing therapy can result in nonideal drug level profiles.

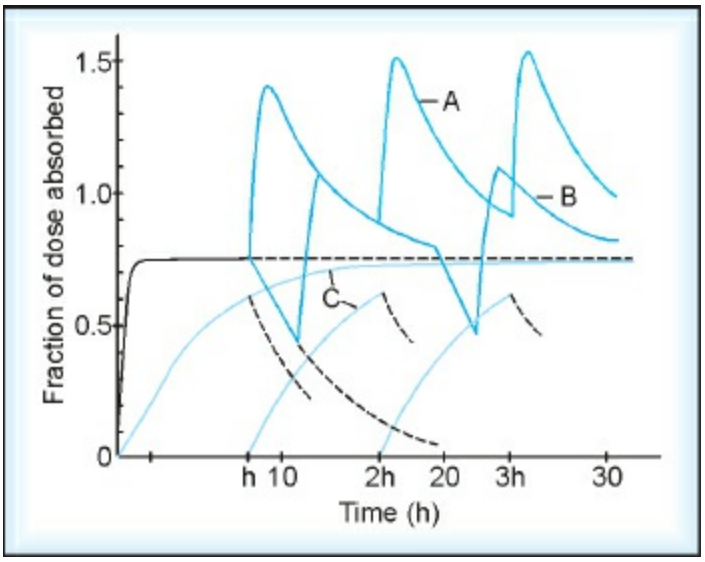



**Fig. 16.6:** Multiple patterns of dosage of sustained release dosage forms where $k_a$ = 2 hr, $k_e$ = 0.23 hr, ($t$ = 3 hr), F = 1, and H = 8hr. Dosage intervals are: A, 8hr; B, 12 hr; and C, 8 hr with no loading dose

In sustained release therapy, only the dosing interval is adjustable. If two units are initially administered followed by single-unit subsequent doses, as is common in therapy with nonsustained forms, a slow fall in overall drug level occurs after several doses to the average level determined by the dosage form

design. Welling has described several strategies for attaining approximations of the ideal profile in the multiple dosing of different sustained release designs based on both cumulative and noncumulative approaches.

Alternately, administration of formulations designed without loading doses can result in minimal fluctuation during long term therapy. In Fig. 16.6, curve C shows the result of this type of multiple dosing regimen where a zero-order based design consisting only of the slow release maintenance dose is administered at intervals of h hours. If absorption is consistent, profiles obtained are equivalent to those resulting from administration of drug by constant rate infusion (curve E, Fig. 16.1), with one dosing interval required to attain nearsteady-state drug levels. Accumulation does not take place unless the dosing interval is made less than the effective maintenance time.

# IMPLEMENTATION OF DESIGNS

## Approaches Based on Drug Modification

Two general sets of methods have been developed for implementation of practical sustained release dosage form designs: methods based on modification of the physical and/or chemical properties of the drug and methods based on modification of the drug release rate characteristics of the dosage that affect bioavailability. The physicochemical properties of a drug may be altered through *complex formation, drug-adsorbate preparation*, or *prodrug synthesis*. These techniques are possible only with drug moieties containing appropriate functional groups (e.g. acidic or basic). The principal advantage of this approach to sustained release is that it is independent of the dosage form design. Drugs so modified may be formulated as liquid suspensions, capsules, or tablets. Loading doses of unmodified drug may also be incorporated in formulations that are ordinarily formulated to release both unmodified and modified drugs without significant delay. Figure 16.7 identifies the mechanisms involved in controlling the release of drug from complexes, adsorbates, and prodrugs.

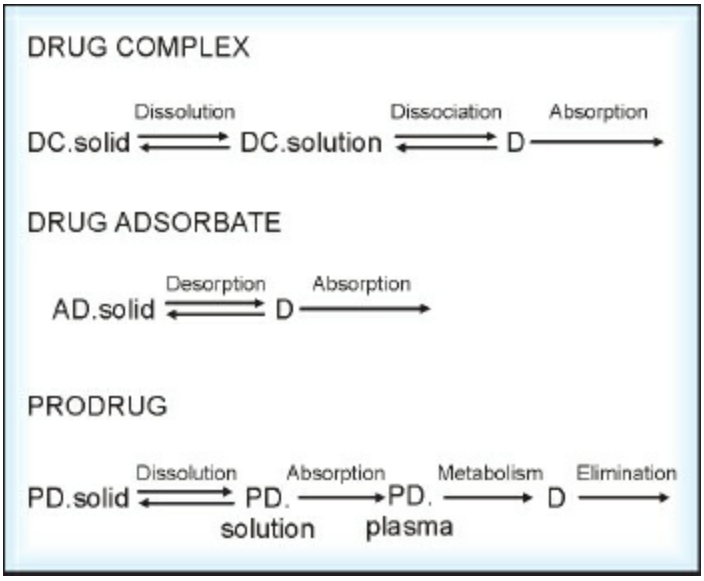



**Fig. 16.7:** Mechanisms of sustained release based on

### *Complex Formation*

In the case of drug complexes, the effective release rate is a function of two

processes: the rate of dissolution of the solid complex into the biologic fluids and the rate of dissociation or breakdown of the complex in solution. In general, the dissolution step may be described by the following expression:

$$\text{Rate dissolution} = k_s\,(\text{Solubility})(\text{Surface area})$$

where, $k_s$ is the dissolution rate constant, a function of the hydrodynamic state as well as factors influencing the diffusion process (e.g. viscosity). The formulator has the option of altering surface area through particle size control and/or solubility of the drug complex through selection of the complexing agent. While both processes are dependent on the pH and composition of the gastric and intestinal fluids, the dissociation step is critically so, since its rate may be pH-dependent, may be determined by the ionic composition of the fluid, and may be affected by the natural digestive processes including enzymatic and bile salt action. The formulator should select the appropriate complex for preparation with knowledge of the specific in vivo processes involved in the control of drug release from the complex. For example, tannate complexes of bases are hydrolyzed in both acidic and basic media, the dissociation of the complex being more rapid at the gastric pH. Drug release from cationic ion-exchange resin complexes depends on sodium ion concentration in GI fluids, and although a stearate salt of a weak base resists the action of gastric fluid, natural digestive processes in the intestine act to dissociate the complex.

If the rate of dissolution is greater than the rate of dissociation, a zero-order release pattern might be realized, because the concentration of complex is maintained at its saturation point if the solubility of the complex is sufficiently low so that excess solid complex is present during most of the effective maintenance time. In this case, high specific surface material should be prepared to promote dissolution. On the other hand, if the rate of dissociation is greater than the rate of dissolution, the dissolution of the complex is rate-determining. Particle size of the complex should be adjusted to establish the most appropriate rate of release. With sufficient excess solid phase, a zero-order release may also be approximated. The complex may be viewed in this instance as simply a means of reducing the solubility of the drug in order to reduce its availability.

Eq. (14) describes the rate (R) at which drug is made available through dissolution under conditions of diminishing surface:

$$R = \frac{4.85DCsW^{2/3}}{hp^{2/3}} = ks'W^{2/3} \quad \text{... (14)}$$

where, *D* is the diffusion coefficient, h is the thickness of the diffusion layer, *Cs* is the solubility, $\rho$ is the density, and W is the weight of undissolved solid. This expression has been derived assuming that the simple diffusion-layer model applies, that the particles are spherical in form, and that a near sink condition is maintained with respect to dissolved complex. Consequently, the rate of drug availability, expressed as the rate of decrease in mass of undissolved complex, diminishes during the effective maintenance period if the surface area decreases as drug is dissolved and absorbed. The amount of drug available at any time can be calculated from the integrated form of equation (14) (the cube-root law):

$$W^{1/3} = (W_0^{1/3} - ks't) \text{ ... (15)}$$

If *ks*' = 3 $mg^{1/3}$/hr and the maintenance dose is 900 mg, then 324 mg of drug would be available after 4 hours, and only 36 mg would be available after 8 hours of dissolution.

## *Drug-adsorbate Preparation*

Drug adsorbates represent a special case of complex formation in which the product is essentially insoluble. Drug availability is determined only by the rate of dissociation (desorption), and therefore, access of the adsorbent surface to water as well as the effective surface area of the adsorbate.

## *Prodrug Synthesis*

Prodrugs are therapeutically inactive drug derivatives that regenerate the parent drug by enzymatic or nonenzymatic hydrolysis. Figure 16.7 shows the scheme that identifies the potential processes for achieving sustained action. The solubility, specific absorption rate, and/or elimination rate constant of an effective prodrug should be significantly lower than that of the parent compound. Kwan has described the pharmacokinetics of a prodrug in which the sustained blood level is determined by the metabolic rate, i.e. by formation of the active moiety after absorption. If the solubility of a drug has been significantly reduced by the formation of prodrug, and if breakdown of the prodrug takes place at the absorption site, then availability is limited by dissolution rate, and the same arguments as in the case of an insoluble drug

complex apply. Examples of drugs from which prodrugs designed for prolonged action have been synthesized include isoproterenol, isoniazid, and penicillin.

Approaches based on drug modification are sensitive to in vivo conditions. An important objective of sustained release formulation is to minimize the effect of in vivo variables on drug release. An alternate approach, which has been advanced by Banker, involves preparation of drug dispersions through “molecular scale drug entrapment” in suitable carrier materials that act to retard release. Compositions of this type can be prepared by induced flocculation of a polymer latex (e.g. acrylic copolymers). Control of drug release is accomplished by varying the nature of the carrier material, the loading dose of drug, and particle size of the product (i.e. surface area). These systems follow a scheme similar to that suggested for drug adsorbates. They also have the advantage of allowing formulation of different dosage forms and may, with appropriate selection of the carrier, be less influenced by in vivo variables.

Since the extended release form of the drug, whether complex, prodrug, or solid dispersion, when formulated as a liquid suspension, is in contact with a fluid medium, an equilibrium is established in the formulation with respect to “free” drug and “bound” drug. The chemical stability of these systems with respect to the conversion of “bound” to “free” drug, in addition to physical stability problems characteristic of suspensions, adds an additional dimension to their overall formulation. The development of injectable depot forms as suspensions of physicochemically modified drugs has been proven to be an effective means of achieving controlled release in antibiotic therapy.

## Approaches Based On Dosage Form Modification

Most peroral sustained release products have been formulated as encapsulations or tablets. Formulations based on modification of the physicochemical properties of these dosage forms can be classed into three product types: *encapsulated slow release beads (or granules), tabletted mixed or slow release granulations*, and *slow release (core) tablets*. Fabrication of tablets allows for direct incorporation of loading doses, by preparation of either multilayered or press-coated tablets. One layer or the outer coat of the tablet is prepared from a potentially rapid disintegrating granulation, leaving the less quickly disintegrating layer or core, which contains the maintenance dose. Systems prepared as tabletted mixed released granulations may or may not be designed to disintegrate quickly, simulating the administration of an encapsulated form in the latter case.

Encapsulated sustained release dosage forms have two specific advantages over core tablet designs. (1) Undisintegrated tablets may remain in the stomach for extended periods of time, excessively delaying absorption of the maintenance dose. Disintegration of the capsule shell in the gastric fluid releases particles that can pass unimpeded through the pyloric valve. (2) There is statistical assurance of drug release with encapsulated forms, since release of drug by a significant fraction of the granules is highly probable. If a core tablet fails to release drug, all of the maintenance dose is lost. Two general principles are involved in retarding drug release from most practical sustained release formulations involving dosage form modification. These are the embedded matrix and the barrier principle, which are schematically shown in Figs 16.8 and 16.9.

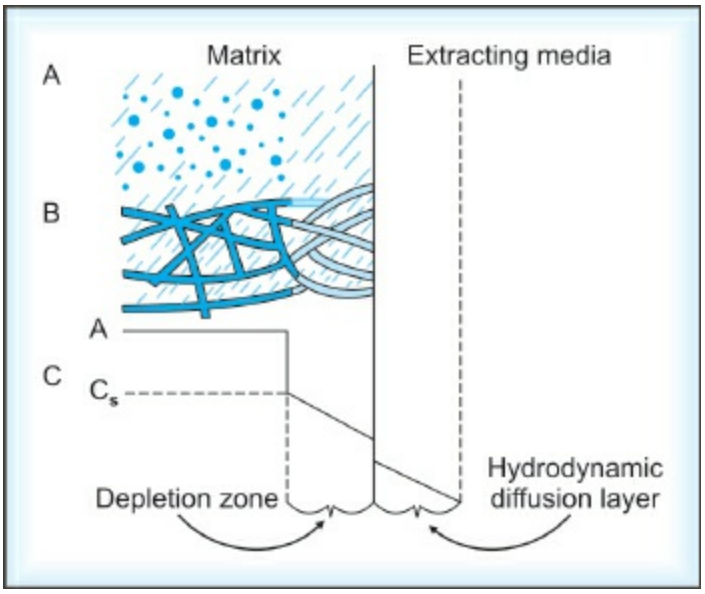



**Fig. 16.8:** Embedded matrix concept as a mechanism of controlled release in sustained release dosage form design. Network model: (A) Drug is insoluble in the retardant material. Dispersion model; (B) Drug is soluble in the retardant material. Diffusion profile; (C) Characterizes drug release from a matrix system

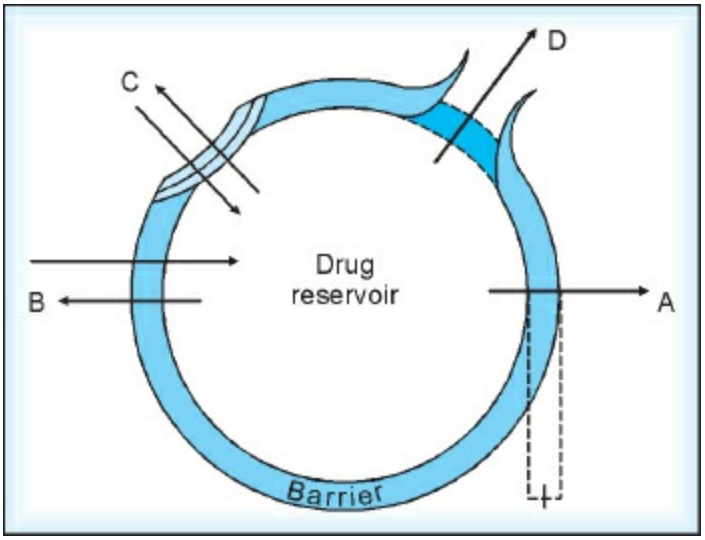



**Fig. 16.9:** Barrier-mediated models of sustained release dosage form designs: A. Drug diffusion through the barrier, B. Permeation of barrier by elution media followed by drug diffusion, C. Erosion of barrier, releasing drug, D. Rupture of barrier as a result of permeation of elution media

## *Embedded Matrix*

In this case, drug is dispersed (embedded) in a matrix of retardant material, which may be encapsulated in particulate form or compressed into tablets. Release is controlled by a combination of several physical processes. These include permeation of the matrix by water, leaching (extraction or diffusion) of drug from the matrix, and erosion of matrix material. Alternately, drug may dissolve in the matrix material and be released by diffusion through the matrix material or partitioned between the matrix and extracting fluid. Matrices may be prepared from insoluble or erodable materials (e.g. silicone polymers or lipids).

Higuchi has provided the theoretic basis for defining drug release from inert matrices. The *Higuchi equation* describing drug release from the planar surface of an insoluble matrix is:

$$Q = [(D \in C_s/\tau)\,(2A - \in C_s)\,t]^{1/2} \quad \ldots (16)$$

where, $Q$ is the amount of drug released per unit surface after time $t$, $D$ is the diffusion coefficient of the drug in the elution medium, $t$ is the tortuosity of the matrix, $\in$ is the porosity of the matrix, $C_s$ is the solubility of the drug in the elution medium, and A is the initial loading dose of drug in the matrix. This expression was derived assuming a linear diffusion gradient as diagrammed in Fig. 16.8C. Drug release is triggered by penetration of eluting media into the matrix, dissolving drug, thereby creating channels through which diffusion takes place (Fig. 16.8B). The depletion zone gradually extends into the core of the matrix. A high tortuosity means that the effective average diffusion path is large. The porosity term takes into account the space available for drug dissolution; an increased porosity results in increased drug release. Both porosity and tortuosity are functions of the amount of dispersed drug, the physicochemical properties of the matrix, and the dispersion characteristics of the drug in the matrix. If the drug is freely soluble in the elution medium, i.e. $C_s >> A$, such that the dissolution rate is rapid, then Eq. (17), which describes the release of drug from a solution entrapped in an insoluble matrix, applies:

$$Q = 2A\,(Dt/\pi\tau)^{1/2} \quad \ldots (17)$$

Release rate is directly proportional to the amount of dispersed drug, $A$; it is proportional to $A$ ½ for insoluble drugs if $2A = C_s$. These expressions predict that plots of $Q$ versus $\sqrt{t}$ be linear. The theory has been extended to

defining matrix-controlled release from spherical pellets as well as from cylindric and biconvex compacts. In Chien's description of the application of a general expression for the case of a drug dispersed in a matrix in which the drug dissolves (Fig. 16.8A), both matrix and partition control are possible. The drug has low solubility in the elution media, partition control dominates, and the release is zero-order, that is:

$$Q = KDC_s t/h \quad \ldots (18)$$

where, $K$ is the partition coefficient ($K = C_s/C_p$), $C_p$ is the solubility in the matrix phase, and $h$ is the thickness of the hydrodynamic diffusion layer. A modified form of Eq. (16) applies in the former case (diffusion takes place in the matrix phase), in which the drug has a high solubility in the elution media.

These expressions have been successfully applied to interpreting drug release from insoluble polymer matrices as well as from such potentially erodable materials as wax-lipid compositions and hydrophilic polymers. With the latter, hydration of the polymer forms a gel, which controls the initial stages of drug release. The variables affecting drug release, which have been studied using these models, have included the nature of retardant, drug solubility, effect of added diluents, drug loading, drug mixtures, and drug-matrix interaction. The rate of drug release from embedded matrices is capable of adjustment by manipulation of the parameters defined by Eqs (16), (17), and (18). The release characteristics of a base formulation can be defined by the slopes of plots of cumulative drug released versus t½. The effect of formulation modifications such as change in drug loading can then be predicted.

Figure 16.10 shows the forms of different drug-release profiles from different dosage form planar models, including zero-order (curve A), first-order (curve B), and square root of time (curve C). Profiles have been adjusted to show 50% release at the same time.

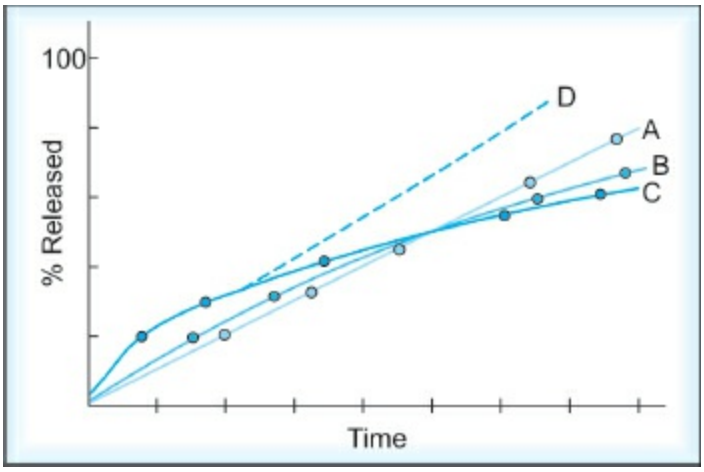



**Fig. 16.10:** Drug-release profiles characteristic of different dosage form models representing embedded matrix systems: A. Zero-order model, B. First-order model, C. Diffusion model, D. Diffusion model with erosion

Analysis of in vitro release data of many different sustained release formulations has demonstrated a pseudo-first-order release characteristic, if the log percentage of unreleased drug was plotted against time. The observed apparent first-order rate constants, however, could not be interpreted in terms of the fundamental properties of the dosage form as in the case for systems characterized by curve C. Curve D represents a situation in which erosion is superimposed on matrix-controlled release. In the planar case, erosion should be zero-order, a function of the product of the drug concentration in the matrix and the effective dissolution rate of the retardant. Release due to erosion is, in general, more rapid than matrix-controlled release. This has been demonstrated with dispersions of chlorpheniramine maleate in methylcellulose matrices.

## *Barrier Principle*

The barrier concept of controlled release implies that a layer of retardant material is imposed between the drug and the elution medium. Drug release results from diffusion of drug through the barrier, permeation of the barrier by moisture, and/or erosion of the barrier. In addition to barrier composition and physicochemical properties, thickness and integrity of the barrier are important variables in controlling drug release. Figure 16.9 summarizes the more significant models of barrier-mediated release. Figure 16.11 shows the form of the drug-release profiles characteristic of these models.

For case A, the barrier is impermeable to the elution medium; drug is present in the reservoir as a solution or suspension. At the steady state, the release rate into a sink is:

$$R = SD_mC_{sm}/l \ldots (19)$$

where, $S$ is the surface area, $D_m$ is the diffusion coefficient of drug in the membrane, $l$ is the thickness of the membrane barrier, and $C_{sm}$ is the solubility of drug in the membrane, assuming constant activity of drug in the reservoir. For the membrane-encapsulated solution, release is first-order. If membrane diffusion is slower than dissolution, release is zero-order for membrane-encapsulated suspensions. Two forms of release profiles may be observed: a burst effect if the membrane is saturated with drug (Fig. 16.11, curve A), and a time lag if drug has not penetrated the membrane (Fig. 16.11, curve B). Drug release approximates first-order kinetics during the depletion phase. This principle has been successfully applied in the development of ophthalmic, intravaginal, and transdermal controlled release devices.

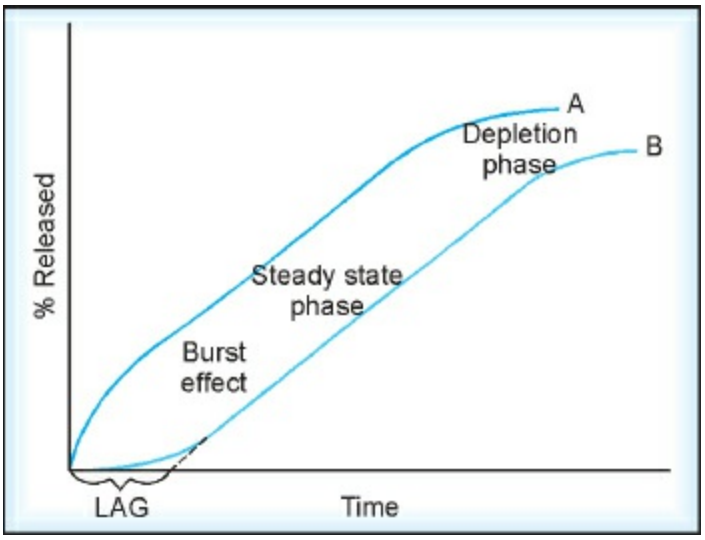



**Fig. 16.11:** Drug-release profiles characteristic of barrier-mediated models: A and B Membrane-controlled diffusion

For case B (Fig. 16.9), in which the barrier is permeable to the elution media, a time lag is involved since drug is not released until moisture has penetrated the barrier, dissolving drug in the reservoir. Additional mechanisms might involve timed erosion of the barrier (case C) or rupture of the barrier after sufficient moisture has permeated the membrane (case D).

The pharmaceutical formulator can select from a variety of potential sustained release dosage form designs. Those based on drug modification are limited to drugs with appropriate structural characteristics. In principle, dosage form designs may be applied to all drug types; however, the selection of a particular dosage form may be limited by the specific drug properties (e.g. solubility, dissociation constant, stability, etc.), the manufacturing technology available, and the methodology needed to establish the validity of the design. For example, water-insoluble drugs may not be suited to designs based on the embedded matrix principle using insoluble matrices, but may be suited to barrier controlled release, which is applicable to a wide variety of drug characteristics.

# SUSTAINED RELEASE FORMULATIONS

## Encapsulated Slow Release Granules

The first significant marketed sustained-release dosage forms were encapsulated mixed slow release beads, to which was applied the *barrier principle* of controlling drug release, based on model D (Fig. 16.9). For low-milligram potency formulations, nonpareil seeds (20/25-mesh sugar-starch granules) are initially coated with an adhesive followed by powdered drug, and the pellets are dried. This step is repeated until the desired amount of drug has been applied. The resultant granules are subsequently coated with a mixture of solid hydroxylated lipids such as hydrogenated castor oil or glyceryl trihydro-xystearate mixed with modified celluloses. The thickness of the barrier was regulated by the number of applied coatings to obtain the desired release characteristic. The original formulations utilized glyceryl monostearate beeswax compositions, which tended to be physically unstable, showing altered release patterns on aging.

A unit of this type contains hundreds of color-coated pellets divided into 3 to 4 groups, which differ in the thickness of the time-delay coating. A typical mix consists of uncoated pellets providing the loading dose and pellets designed to release drug at 2 or 3 hours, 4 or 6 hours, and 6 or 9 hours. The key factor controlling drug release is moisture permeation of the barrier, which depends on coating thickness. Absorption of moisture by the core and subsequent swelling rupture the coating, releasing drug. Some pellets within each group “release drug at intervals overlapping other pellet groups, resulting in a smooth rather than discontinuous release profile. Variables that can be manipulated to alter the release pattern include the amount of drug per pellet, the composition and thickness of the coating, and the number of pellets included in each group. Regulation of coating and maintenance of the proper mix of beads during encapsulation present specific production difficulties.

In the case of high-milligram potency formulations, individual crystals of drug or pelletized drug may be coated by pan or f luidized-bed processes with a retardant barrier. Combinations of waxes, fatty acids, alcohols, and esters can be applied using fluidized-bed technology. Such enteric materials as cellulose acetate phthalate and formalized gelatin, as well as lipid compositions, have been used to control release by erosion. 

Microencapsulation by spray-drying or coacervation has also been used to prepare sustained release encapsulations.

Many encapsulation formulations that implement barrier controlled release have been designed to produce granules with uniform rather than mixed release characteristics, thus eliminating some of the production complexities associated with the manufacture of mixed release pellets. One group of products consists of pellets coated with a hydrolyzed sytrene maleic acid copolymer, which produces a pH-sensitive barrier. Formulations of methylpredriisolone have been described as releasing 5% drug in 2 hours at pH 1.2, and 90 to 100% drug in 4 hours at pH 7.5. Delayed release ascorbic acid consisting of ascorbic acid crystals encapsulated in partially hydrogenated cottonseed oil has been marketed for the food industry. Typical formulations contain up to 50% lipid. Pseudo-latexes of ethylcellulose modified by the addition of dibutyl sebacate and plasticized by triethyl citrate have been shown to produce effective retardant barriers whose permeability can be altered by varying the additive concentration. A near zero-order release is claimed with a coating containing 24% sebacate.

A unique application of the barrier concept involves the preparation of granules described as microdialysis cells (model B, Fig. 16.9). Drug-containing pellets might be coated with ethylcellulose, a water-insoluble and pH-insensitive polymer, modified by the addition of suspended sodium chloride particles or other water-soluble materials (e.g. polyethylene glycol). Reaching the salt from the film results in a dialytic membrane formed in situ. This allows permeation of water, dissolution of drug, and diffusion of drug through the essentially intact membrane. These microdialysis cells have been applied to produce a sustained release form of nitroglycerin, which is claimed to be independent of the composition of the GI fluids; however, release of acidic or basic drugs from this system is generally pH-dependent. Pellet cores containing propoxyphene in a buffer system are claimed to produce a pH-independent drug release if formulated as microdialysis cells. Drug may also be microencapsulated in ethylcellulose using a thermally induced phase separation technique in which polyisobutylene is used to control particle size and release rate.

Alternately, placebo pellets may be coated with polyethylene glycol modified ethylcellulose, shellac, or cellulose acetate phthalate containing suspended drug, which is leached from the coating as it is eroded by the

action of intestinal fluid. This approach combines both barrier and embedded matrix models of drug release. Release profiles conform to patterns C or D in (Fig. 16.10), depending on the relative importance of erosion.

The embedded matrix principle has also been applied to prepare sustained release encapsulations. Medicaments are dispersed in molten lipid materials to form a slurry, which may be spray-congealed, or after solidification, granulated. Preferred retardant materials include hydrogenated oils, glyceryl stearates, fatty alcohols, and microcrystafiine wax. The addition of 2 to 10% wicking agents, i.e. finely divided powders of methylcellulose, alginic acid, or carboxymethylcellulose, is claimed to produce a greater uniformity of drug release, which more nearly approximates a zero-order process in vitro testing. These agents promote the permeation of moisture into the matrix, facilitating its erosion. Time-delay matrix materials may constitute 45 to 75% by weight of the formulation. Spray-congealing is advantageous since spherical pellets varying in size from 250 to 2000 microns can be obtained. Sustained release is achieved in part by the random mix of different particle sizes and random dispersal of drug in the matrix.

Suspension or emulsion polymerization has been used to produce resin beads containing drug for sustained release. For example, methyl methacrylate monomer containing drug and mediacrylic acid, a water soluble monomer that is added to contribute a swelling characteristic to the final product, are dispersed in an aqueous solution containing appropriate suspending and deflocculating agents. Benzoyl peroxide is added to catalyze the polymerization. The resulting product consists of uniformly sized beads that can be encapsulated. Vinyl acetate and epoxy resins have also been used successfully to produce sustained release beads. With the latter, curing agents are required, and the reaction is carried out in a silicone oil phase in which the curing agent and the drug dispersed in the liquid resin are insoluble. Drug release from such systems is controlled by both matrix diffusion and partitioning as defined by Eq. (19).

## Tabletted Slow Release Granulations

Compression of timed-release granulations into tablets is an alternate to encapsulation. Such tablets should be designed to disintegrate in the stomach so as to simulate the administration of a capsule form having the advantages associated with sustained release encapsulations, while retaining the advantages of the tablet dosage form. Three examples, each utilizing a different process, illustrate this type of formulation. The first is a tabletted mixed release granulation in which binders with different retardant properties are used to prepare three different granulations, which are color coded for identification, blended and tabletted. The first is a conventional nonsustained release granulation prepared using gelatin as a binder; the second uses vinyl acetate, and the third uses shellac, as binders. Drug release is controlled by erosion of the granulation in intestinal fluid—the vinyl acetate granulation disintegrates at a faster rate than the shellac granulation.

The second example is illustrated by a sustained release aspirin formulation based on the microdialysis cell principle. Aspirin crystals are microencapsulated in a retardant barrier and are compressed to form a tablet that rapidly disintegrates into sustained release granules. The barrier approach is particularly advantageous for formulation of high-milligram-potency drugs such as aspirin, since only a relatively small amount of retardant is required in the formulation. The third example is represented by a sustained release form of theophylline, which is claimed to release drug zero-order for a 12 hour dosing interval. The tablet is formulated as a matrix of loading dose of theophylline containing theophylline pellets encapsulated in a semipermeable coating. Disintegration of the matrix in the stomach releases the extended action pellets. The loading dose granulation should have physical characteristics, e.g. size, similar to the maintenance dose pellets to ensure homogeneous mixing of the two granulations during compression.

## Matrix Tablets

One of the least complicated approaches to the manufacture of sustained release dosage forms involves the direct compression of blends of drug, retardant material, and additives to form a tablet in which drug is embedded in a matrix core of the retardant. Alternately, retardant-drug blends may be granulated prior to compression. Table 16.7 identifies examples of the three classes of retardant material used to formulate matrix tablets, each class demonstrating a different approach to the matrix concept. The first class consists of retardants that form insoluble or "skeleton" matrices; the second class represents water-insoluble materials that are potentially erodable; and the third class consists of polymers that form hydrophilic matrices. Loading doses are best included as the second layer of a two-layer tablet or in a coating applied to the matrix core.

**Table 16.7:** Materials used as retardants in matrix tablet formulation

| Matrix characteristics | Material |
|---|---|
| Insoluble, inert | Polyethylene<br>Polyvinyl chloride<br>Methyl acrylate-methacrylate copolymer<br>Ethylcellulose |
| Insoluble, erodable | Carnauba wax<br>Stearyl alcohol<br>Stearic acid<br>Polyethylene glycol<br>Castor wax<br>Polyethylene glycol monostearate<br>Triglycerides |
| Hydrophilic | Methylcellulose (400 cps, 4000 cps)<br>Hydroxyethylcellulose<br>Hydroxypropylmethylcellulose (60 HG, 90 HG, 25 cps, 4000 cps, 15,000 cps)<br>Sodium carboxymethylcellulose<br>Carboxypolymethylene |

| | |
|---|---|
| | Galactomannose |
| | Sodium alginate |

Insoluble, inert polymers such as polyethylene, polyvinyl chloride, and acrylate copolymers have been used as the basis for many marketed formulations. Tablets prepared from these materials are designed to be egested intact and not break apart in the GI tract. Tablets may be directly compressed from mixtures of drug and ground polymer; however, if ethyl cellulose is used as the matrix former, a wet granulation procedure using ethanol can be employed. The rate-limiting step in controlling release from these formulations is liquid penetration into the matrix unless channeling (wetting) agents are included to promote permeation of the polymer matrix by water, which allows drug dissolution and diffusion from the channels created in the matrix. Formulations should be designed so that pore diffusion becomes rate-controlling, release is defined by Eqs (16) or (17), and the release profile is represented by curve C in Fig. 16.10. Drug bioavailability, which is critically dependent on the drug: polymer ratio, may be modified by inclusion of diluents such as lactose in place of polymer in low-milligram-potency formulations.

Egested tablets contain unreleased drug in the core. In one study of polyvinyl chloride matrix tablets containing prednisolone disodium phosphate, egested tablets contained 72% of the maintenance dose for matrices containing 87% plastic and 2% drug, and 28% drug for matrices containing 84% plastic and 3% drug. These forms of matrix tablets are not useful for high-milligram-potency formulations in which the polymer content would be insufficient to form a matrix, or for highly water-insoluble drugs in which dissolution in the matrix would become rate-limiting. Release of water-soluble drugs, however, should be unaffected by the amount of liquid, pH-value, enzyme content, and other physical properties of digestive fluids, unless the drug is in a salt form that precipitates within the matrix pores on dissolution when penetrated by acid or basic media.

Waxes, lipids, and related materials form matrices that control release through both pore diffusion and erosion (curve D, Fig. 16.10). Release characteristics are therefore more sensitive to digestive fluid composition than to the totally insoluble polymer matrix. Total release of drug from wax-lipid matrices is not possible, since a certain fraction of the dose is coated

with impermeable wax films. Release is more effectively controlled by the addition of surfactants or wicking agents in the form of hydrophilic polymers, which promote water penetration and subsequent matrix erosion.

Carnauba wax in combination with stearyl alcohol or stearic acid has been utilized as a retardant base for many sustained release matrix formulations. Mixtures of (1:1) hydrogenated castor oil and propylene glycol monostearate and of carnauba wax and stearyl alcohol or stearic acid have been extensively studied as retardants for both water-soluble and water-insoluble compounds. Materials with melting points that are too low or materials that are too soft cannot be readily processed to form tablets with good physical stability. Such retardants as carnauba wax or hydrogenated castor oil provide the necessary physical characteristics to form an easily compressible stable matrix. If used singly, these materials excessively delay drug release.

Three methods may be used to disperse drug and additive in the retardant base. A solvent evaporation technique can be used, in which a solution or dispersion of drug and additive is incorporated into the molten wax phase. The solvent is removed by evaporation. Dry blends of ingredients may be slugged and granulated. A more uniform dispersion, however, can be prepared by the fusion technique, in which drug and additive are blended into the molten wax matrix at temperatures slightly above the melting point (approximately 90°C for carnauba wax). The molten material may be spray-congealed, solidified and milled, solidified and flaked, or poured on a cold rotating drum to form sheets, which are then milled and screened to form a granulation.

In the absence of additives, drug release is prolonged and nonlinear. Apparent zero-order release can be obtained by addition of additives such as polyvinyl pyrrolidone or polyoxyethylene lauryl ethers. In a study by Dahkuri et al., 10 to 20% hydrophilic polymer effectively controlled release from carnauba-wax/stearyl-alcohol matrices of tripelennamine hydrochloride. Matrices prepared from carnauba-wax/polyethylene-glycol compositions have also been used to prepare sustained release theophylline tablets. The wax:Glycol ratio could be adjusted to vary the release characteristics.

A novel approach to the development of a lipid matrix utilizes pancreatic lipase and calcium carbonate as additives, with triglycerides as retardants. The lipase is activated on contact with moisture and thus promotes erosion

independent of intestinal fluid composition. The release profile is controlled by the calcium carbonate, since calcium ions function as a lipase accelerator. In another technique, drug is mass, blended with stearyl alcohol at a temperature above its glass transition (approximately 60°C), and the mass is cooled and granulated with an alcoholic solution of zein. This formulation is claimed to produce tablets with stable release characteristics. Since natural waxes and lipids are complex mixtures, and a fusion process is, usually required for processing, hardening with decrease in effective drug release on aging may be observed, owing to polymorphic and amorphous to crystalline transitions.

The third group of matrix formers represents nondigestible materials that form gels in situ. Drug release is controlled by penetration of water through a gel layer produced by hydration of the polymer and diffusion of drug through the swollen, hydrated matrix, in addition to erosion of the gelled layer (curve D, Fig. 16.10). The extent to which diffusion or erosion controls release depends on the polymer selected for the formulation as well as on the drug:polymer ratio. Low-molecular-weight methylcelluloses release drug largely by attrition, since a significant intact hydrated layer is not maintained. Anionic polymers such as carboxymethyl cellulose and carpolene can interact with cationic drugs and show increased dissolution in intestinal fluid. Carboxypolymethylene does not hydrate in gastric fluid. The best matrix former in this group is hydroxymethylcellulose 90 HG 15,000 cps, an inert polymer that does not adversely interact with either acidic or basic drugs, and that on contact with water slowly forms a gel that is more resistant to attrition. Release rates can be adjusted for low-milligram potency formulations by replacing polymer with lactose. High drug: polymer ratios result in formulations from which drug release is controlled by attrition.

The process used to prepare formulations for compression depends on the polymer and drug: polymer ratio, With high drug: polymer ratios, a wet granulation process is required. Low-milligram-potency formulations may be directly compressed or granulated using alcohol if the polymer is not in a form amenable to direct compression. Formulations of this type are designed to release 100% of drug in vivo, unlike the other matrix forms, which may be partially egested and consequently must be formulated to contain drug in excess of that required to attain the desired therapeutic effect.

Pseudo-latex forms of the normally water-insoluble enteric polymers,

such as cellulose acetate phthalate and acrylic resin, can be used as granulating agents for high-milligram-potency drugs. Erodable matrix tablets can be prepared from these granulations. This approach has been tested with theophylline.

## Drug Complexes

The principal advantage of preparing drug derivatives for sustained release is that such materials can be formulated into diverse dosage forms. This approach has proven effective in the development of injectable depot forms, in which release profiles are not subject to the variability characteristic of the gastrointestinal tract. Sensitivity to in vivo variables is a definite disadvantage of perorally administered forms: in vivo studies may not consistently support sustained release claims.

If an alcoholic solution of a basic drug and tannic acid are mixed in a 5:1 drug:Tannic acid ratio at reduced temperature, tannate complexes containing one amine per digallyl moiety are precipitated. These complexes are split by hydrolysis in gastric and intestinal fluid. Amphetamine and antihistamine tannates were once marketed in both tablet and suspension forms with sustained release claims. Breakdown of the tannate complex depended on pH, being somewhat faster in gastric than intestinal fluid, as well as on the low solubility of the complex. Other complex acids used to prepare relatively insoluble and degradable complexes of basic drugs have included polygalacturonic acid, alginic acid, and arabogalactone sulfate. Products obtained by interaction of montmorillonite clays (e.g. bentonite) with cationic drugs or amine salts and certain nonionic drugs have also been investigated. Both cation exchange and strong chemisorption contribute to the interaction. The release profiles can be varied by altering the drug:clay ratio. A 1:20 complex of amphetamine, for example, is reported to release drug effectively in intestinal fluid, with little release in gastric fluid. Since the clays are anionic, effective adsorbates cannot be prepared from anionic drugs.

## Ion-activated Systems

Ion-exchange resin complexes, which potentially can be prepared from both acidic and basic drugs, have been more widely studied and marketed. Salts of cationic or anionic exchange resins are insoluble complexes in which drug release results from exchange of "bound" drug ions by ions normally present in GI fluids ($Na^+$, $H^+$, $C1^-$, $OH^-$). Resins used are special grades of styrene/divinyl benzene copolymers that contain appropriately substituted acidic groups (carboxylic or sulfonic for cation exchangers) or basic groups (quaternary ammonium for anion exchangers) on the styrene moiety of the resin. Ion-active sites are distributed uniformly throughout the resin structure. Variables relating to the resin are the degree of cross-linking, which determines the permeability of the resin, its swelling potential, and the access of the exchange sites to the drug ion; the effective pKa of the exchanging group, which determines the exchange affinity; and the resin particle size, which controls accessibility to exchange ions. Drug-resin salts, for example, may be prepared by percolation or equilibrium of the resin in acid form with a concentrated solution of drug hydrochloride salt. The resin is washed with ion-free water and partially dried. The resulting product can be encapsulated, tabletted, or suspended in ion-free vehicles. The following equations represent the preparation and exchange reaction affecting drug release in vivo:

$$\text{Resin} \sim SO_3Na + \text{Drug HCl} = NaCl + \text{Resin} \sim SO_3.\ \text{Drug H}$$
$$\text{Resin} \sim SO_3.\ \text{Drug H} + NaCl = \text{Drug HCl} + \text{Resin} \sim SO_3Na$$

A strong acid resin must be used to minimize exchange of drug by hydrogen ion, to avoid excessive drug release in the gastric fluid. The percentage of cross-linking is the most important variable affecting release profiles. Resins with minimum cross-linking show maximum swelling when converted from free acid to salt forms. Subsequent contact with acid can cause shrinkage and reduction of pore volume at the resin periphery, thus entrapping large ions. However, more drug is available for release at higher pH-values. Increased cross-linking decreases resin porosity, not only reducing drug availability but also limiting access of exchange groups to drug ions during preparation. Particle size control becomes important when pore and molecular size are important. Similar observations apply to drug resinates, which are prepared by reaction of sodium salts of acidic drugs with

resin chlorides.

The amount of drug that can be incorporated in these systems is limited to a maximum of 200 to 300 mg, since larger doses require too much resin. Release profiles characteristic of resin complexes are sensitive to variation of in vivo ion concentrations, Coating of the resin beads with appropriate polymers, which act as a diffusion barrier to both exchange ions and exchanged drug and water, provides a controllable rate-limiting factor that minimizes the effect of in vivo variables. In one process, the resin particles are pretreated with polyethylene glycol to provide a base for a subsequent coating of ethylcellulose plasticized with a refined vegetable oil. Release profiles can be controlled by appropriate mixing of both coated and uncoated drug-resin complexes. The complexes can be formulated in encapsulated or suspension forms. Different drugs might be readily combined as coated resin forms with independently controlled release characteristics in the same product.

## pH-Dependent Systems

For a drug labile to gastric fluid pH or irritating to gastric mucosa, this type of Controlled drug delivery systems has been developed to target the delivery of the drug only in the intestinal tract, not in the stomach. It is fabricated by coating a core tablet of the gastric fluid-sensitive drug with a combination of intestinal fluid—insoluble polymer like hydroxylmethyl cellulose phthalate, and ethyl cellulose (Fig. 16.12). In the stomach, the coating membrane resists the degrading action of gastric fluid (pH <3), and the drug molecules are thus sheltered from the acidic degradation.

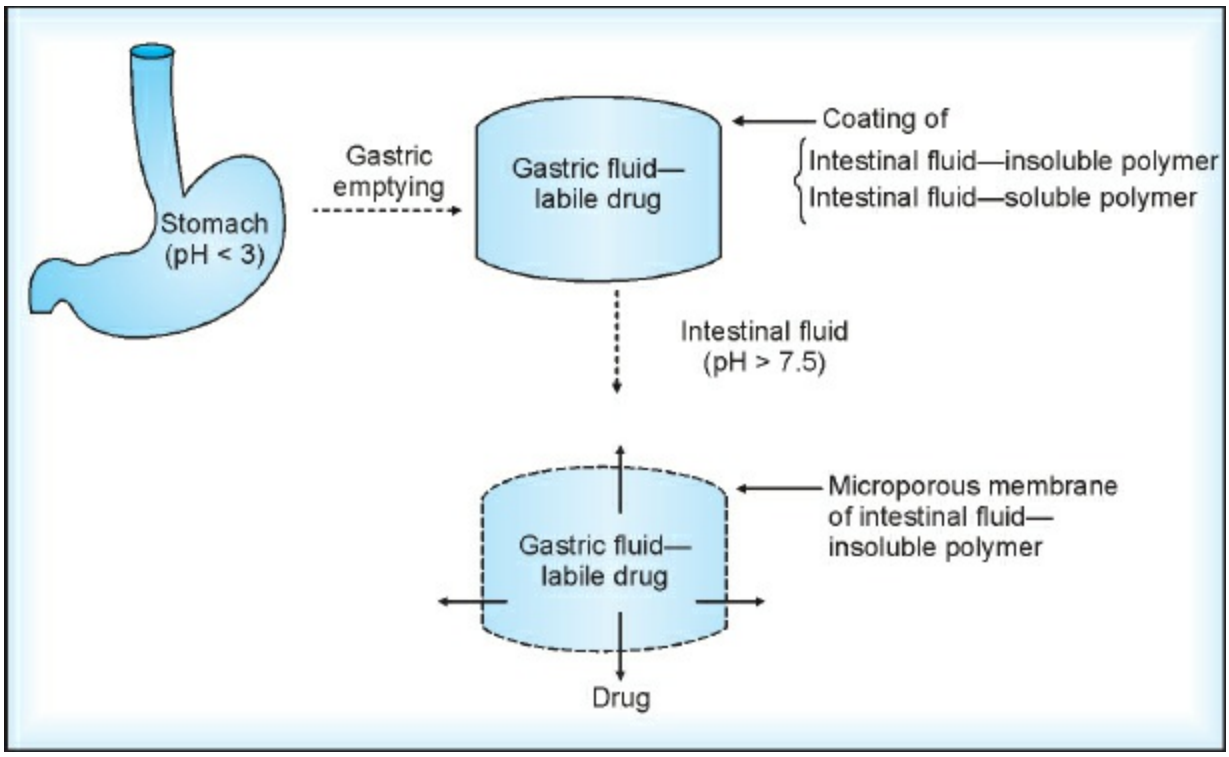



**Fig. 16.12:** Schematic repesentation of the pH-dependent systems

After gastric emptying, the controlled drug delivery systems travels to the small intestine, and the intestinal fluid-soluble component in the coating membrane is dissolved away by the intestinal fluid (pH > 7.5). This produces a microporus membrane of intestinal fluid-insoluble polymer to control the release of drug from the core of the tablet. The drug is thus delivered in a controlled mode in the intestine by a combination of drug dissolution in the core and diffusion through the pore channels. By bending the ratio of the

intestinal fluid-soluble polymer to the intestinal fluid-insoluble polymer in the membrane, the rate of drug delivery can be regulated. Typical example of this type of controlled drug delivery systems is in the oral controlled delivery of potassium chloride, which is highly irritating to gastric epithelium.

## pH-Independent Systems

When a formulation is administered orally, it encounters several pH environments until absorbed or excreted. If the formulation is chewed, first environment of the drug exposed will be pH 7. The drug will then be exposed to a pH of 1 to 4 in the stomach, depending on the amount and type of food, followed by a pH of 5 to 7 in the intestine. Many reports show a pH dependence of drug release from a sustained-release formulation. For example, the release of papaverine from a commercial sustained-release preparation of papaverine is signicantly affected by pH of the dissolution media. Most of the drug is released in the stomach from this preparation and little is released in the intestine due to low solubility. To achieve pH-independent drug release, buffers can be added to the drug to help maintain a constant pH. Salts of tartaric acid, phosphoric acid, citric acid, phthalic acid, or amino acids are preferred because of physiological acceptability. The rate of availability/release of propoxyphene after administration of a buffered controlled-release formulation showed signicantly increased reproducibility, probably due to lower sensitivity of its release rate to the surrounding pH.

## Altered Density Systems

### *Low Density Micropellets*

Empty globular shells, which have a visible density lower than that of gastric juice, can be used as carriers of drugs for sustained or controlled release purposes. Conventional polystyrol, gelatin capsules, and poprice are candidates as carriers. The surface of the empty shells is undercoated with polymers, such as acrylic and methacrylic copolymer, cellulose acetate phthalate, or sugar. This undercoated shell is further coated with a drug-polymer mixture and polymeric material that shows dissolution-controlled drug release (e.g. hydroxypropylcellulose, ethylcellulose, or cornstarch). This type of carrier oats on the gastric juice for an extended period with slow release the drug. This same principle can be applied to formulate buoyant capsules. The particles of a drug-hydrocolloid mixture will swell to form a soft gelatinous mass on the surface when in contact with gastric juice. In vitro dissolution studies with this formulation show a good correlation with plasma concentration levels. Hydrocolloids that are suitable for this purpose are hydroxyalkylcellulose, alginate, carrageenan, carboxymethylcellulose, agar, guar gum, gum arabic, gum karaya, gum tragacanth, locust bean gum, pectin, and the like. It has been reported that multiple-unit drug formulations have an advantage over single-unit preparations. The subunits of the multiple-unit formulation are distributed throughout the GI tract, and their transport is less affected by transit time of food. Further, difference in specific density of the multiple-unit dose signicantly influence the transit time of the subunits through the GI tract.

### *High Density Micropellets*

An increase in density from 1.0 to 1.6 extended the average transit time from 7 to 25 h. The pellets are dispersed throughout the small intestine at a rate that depends principally on their density. Titanium dioxide, zinc oxide, barium sulphate, and iron powder are the substances used to increase pellet density. Density of the pellets must exceed the normal stomach contents and should therefore be at least 1.4. Furthermore, a diameter of 1.5 mm is considered maximal for a true multiple-unit formulation. The drug can be coated on a heavy core or mixed with heavy inert materials, and then the weighted pellet can be covered with a diffusion-controlling membrane.

## Colonic Release Systems

Drugs are poorly absorbed through colon, but may be delivered to this site for two reasons, i.e. *local action* as needed in treatment of ulcerative colitis, infection or diarrhoea and systemic absorption of proteins and peptides (insulin and vasopressin). Design principles for these delivery systems take advantages of:

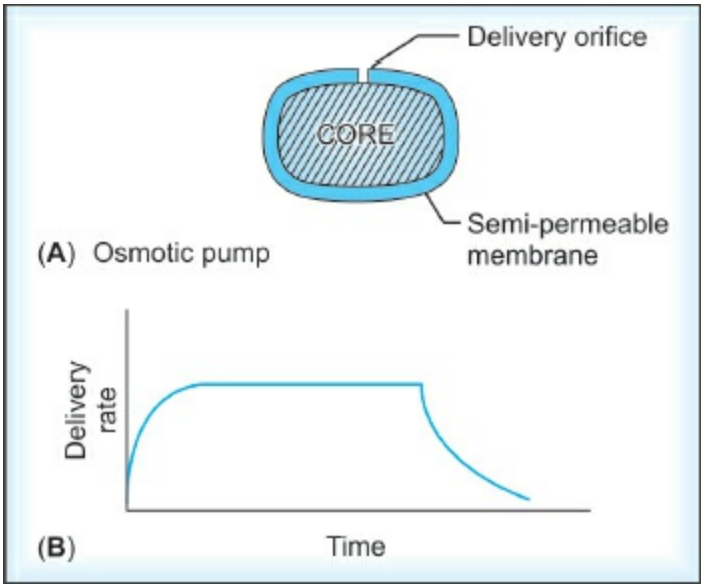
- The *specific pH of colon* – pH sensitive bioerodible polymers like polymethacrylates (e.g. combination of Eudragit 10055 and Eudragit S) release the medicament only at the alkaline pH of the colon.
- The *colonic microflora* – coating of drug delivery system with polymers, which can be cleaved only by the azoreductase of colonic bacteria to release the medicament. Polymers used include divinylbenzene cross-linked polymer, pectin cross-linked with calcium with different degree of methoxylation or glassy amylase mixed with ethylcellulose.
- The *small intestine transit time* – swelling-induced time-controlled drug delivery systems lead to a delay in drug release. Polymers used are poly(ethyl-acrylate-methylmethacrylate) i.e. Eudragit NE 30 D or Eudragit RS.
- The *temporary increase in intra-luminal pressure* – hard gelatin capsules with inner ethylcellulose coating disintegrate in the colon by intra-luminal pressure resulting from strong peristaltic wave and release the drug.

# CONTROLLED RELEASE FORMULATIONS

Controlled release dosage forms are designed to release drug in vivo according to predictable rates that can be verified by in vitro measurements. Of the many approaches to formulation of sustained-release medication described in this chapter, those fabricated as insoluble matrix tablets come closest to realization of this objective, since release of water-soluble drug from this form should be independent of in vivo variables. Controlled release technology implies a quantitative understanding of the physicochemical mechanism of drug availability to the extent that the dosage form release rate can be specified. Potential developments and new approaches to oral controlled release drug delivery include osmotic pressure activated systems and hydrodynamic pressure activated systems.

## Osmotic Pressure-activated Systems

One example of a dosage form design that demonstrate the appliance of controlled release technology to pharmaceutical formulation is the orally administered elementary osmotic pump shown in Fig. 16.13A. This delivery system is fabricated from a tablet that contains water-soluble osmotically active drug, or that is blended with an osmotically active substance as diluent, by coating the tablet with a cellulose triacetate barrier, which functions as a semipermeable membrane. A laser is used to form a precision orifice in the barrier. Since the barrier is permeable only to water, initial penetration of water dissolves the outer part of the core, resulting in the development of an osmotic pressure difference across the membrane. The system imbibes water at a rate proportional to the water permeability and effective surface area of the membrane and to the osmotic gradient of the core formulation. The device delivers a volume of saturated solution equal to the volume of water uptake through the membrane. After an initial lag time (approximately 1 hour) during which the delivery rate increases to its maximum value, drug release is zero-order, as shown in Fig. 16.13B, until all solid material is dissolved. Thereafter, the delivery rate decreases parabolically to zero.




**Figs 16.13A and B:** Controlled release dosage forms: (A) Cross-section of osmotic pump; (B) Release rate profile characteristic of osmotic pump

The diameter of the orifice must be smaller than a maximum size to

minimize drug delivery by diffusion through the orifice, and larger than a minimum size to minimize the hydrostatic pressure in the system, which acts in opposition to the osmotic pressure. For devices containing potassium chloride, orifices can range from 75 to 275 mm in diameter. The device can be used as a drug delivery system for any water-soluble drug and solubility of drug in water plays a critical role in its functioning. Typically the solubility of drug delivered by these pumps should be at least 10 to 15% w/v. Osmotic pumps can be designed to deliver significant fractions of the total dose at zero-order rates unaffected by in vivo conditions. Since ions do not diffuse into the device, release of acidic and basic drugs is independent of gastrointestinal pH.

The choice of a rate-controlling membrane is an important aspect in the formulation development of oral osmotic systems. Since the membrane in osmotic systems is semi permeable in nature, any polymer that is permeable to water but impermeable to solute can be selected. Some of the polymers that can be used for above purpose include cellulose esters such as cellulose acetate, cellulose diacetate, cellulose triacetate, cellulose propionate, cellulose acetate butyrate, etc.; cellulose ethers like ethyl cellulose and eudragits. Cellulose acetate is commonly employed semipermeable polymer for the preparation of osmotic pumps. It is available in different acetyl content grades. Degree of substitution (DS), is the average number of hydroxyl groups on the anhydro-glucose unit of the polymer replaced by substituting group such as acetyl. Particularly acetyl content 32 and 38% are widely used. Cellulose diacetate having a DS of 1–2 has an acetyl content of 21–35%. Cellulose triacetate is having a DS of 2–3 and an acetyl content of 35–44.8%. Apart form cellulose derivatives, some other polymers such as agar acetate, amylose triacetate, betaglucan acetate, poly (vinyl methyl) ether copolymers, poly (orthoesters), poly acetals and selectively permeable poly (glycolic acid) and poly (lactic acid) derivatives can be used as semipermeable film forming materials. The permeability is the important criteria for the selection of semipermeable polymers.

A design that provides zero-order release of potassium chloride consists of the soluble tablet core coated with a microporous membrane, which controls the diffusion rate. The membrane is produced in situ by leaching out sucrose that has been suspended in a polyvinyl chloride membrane.

## Hydrodynamic Pressure-activated Systems

A hydrodynamic pressure-activated drug-delivery system can be fabricated by putting a liquid drug formulation inside a collapsible, impermeable barrier to form a drug reservoir compartment. This is then contained inside a rigid, shape-retaining housing. A shield of an absorbent layer and a swellable, hydrophilic polymer layer is sandwiched between the drug reservoir compartment and the housing. In the gastrointestinal tract, the laminate will swallow the gastrointestinal fluid through the annular openings at the lower end of the housing and become swollen. This creates a hydrodynamic pressure inside the system. The hydrodynamic pressure, thus developed, forces the drug reservoir compartment to reduce in volume and causes the liquid drug formulation to release through the delivery orice. The drug release rate is expressed by:

$$\frac{Q}{t} = \frac{P_f A_m}{h_m}(q_s - q_e)$$

where, $P_f$, $A_m$ and $h_m$ are the fluid permeability, the effective surface area, and the thickness of the wall with annular openings, respectively; $q_s$ and $q_e$, is the difference in hydrodynamic pressure between the drug delivery system ($q_s$) and the environment ($q_e$). The release of drug from this type of systems is activated by hydrodynamic pressure and controlled rate is determined by the fluid permeability and effective surface area of the wall with annular openings as well as by the hydrodynamic pressure gradient.

## Hydrodynamically Balanced Systems/Intra-gastric Floating Tablets

Several approaches are currently utilized in the prolongation of the gastric retention time (GRT), including floating drug delivery systems (FDDS), also known as hydrodynamically balanced systems (HBS), swelling and expanding systems, polymeric bioadhesive systems, modified-shape systems, high-density systems, and other delayed gastric emptying devices. Floating drug delivery systems either float due to their low density than stomach contents or due to the gaseous phase formed inside the system after they come in contact with the gastric environment. Depending upon the working principle of floating drug delivery systems they are divided into two main categories: noneffervescent and effervescent floating drug delivery systems.

***Non-effervescent FDDS*** are usually prepared from gel-forming or highly swellable cellulose type hydrocolloids, polysaccharides or matrix forming polymers like polyacrylate, polycarbonate, polystyrene and polymethacrylate. In one approach, gel-forming hydrocolloid swells in contact with gastric fluid after oral administration and maintains a relative integrity of shape and a bulk density of less than unity within gastric environment. The air thus trapped by the swollen polymer imparts buoyancy to the dosage form. Nevertheless, the gel structure acts as a reservoir for sustained drug release. When these dosage forms come in contact with an aqueous medium, the hydrocolloids imbibe water and start to hydrate thereby forming a gel at the surface. The resultant gel layer subsequently controls the trafficking of drug out and passage of solvent in to the dosage form. With the passage of time the exterior surface of the dosage form goes into solution and immediate adjacent hydrocolloid layer becomes hydrated and maintains the gel structure. The drug in dosage form dissolves in and diffuses out with the diffusing solvent forming a ‘receding boundary’ within the gel structure.

***Effervescent FDDS*** employ matrices from swelling polymers like Methocel® or chitosan and effervescent components such as sodium bicarbonate and tartaric or citric acid or matrices having chambers of liquid components that gasify at body temperature. The matrices are prepared in such a manner that when they come in contact with stomach fluid, carbon dioxide is generated, and retained entrapped in the hydrocolloid gel. This leads to an upward drift of the dosage form and maintains it in a floating

condition. A single layer tablet can be produced by intimately mixing the carbon dioxide generating component in tablet matrix. A bilayer tablet may be compressed in which gas liberating component is present in hydrocolloid layer and the drug is compressed in other layer for sustained release. The concept has been judiciously utilized to develop floating capsule system consisting of a mixture of sodium alginate and sodium bicarbonate. Recently, a multiple unitfloating pill has been developed based on the concept of effervescence In this system, carbon dioxide gas is generated from reaction of sodium bicarbonate and tartaric acid. The system consisted of sustained release pill surrounded by an effervescent layer. This coated system is further coated with swellable polymers like polyvinyl acetate and purified shellac. Moreover, the effervescent layer is divided into two sub layers to prevent direct contact between tartaric acid and sodium bicarbonate.

## The Synchron System

This system is a homogeneous mixture of cellulose and noncellulose material with a drug in a tablet from. When the Synchron system tablet comes in contact with water, the outer layer of the matrix softens to a gel-like consistency, which allows the trapped drug to release at a controlled rate. The Synchron system is Forest Laboratories patented procedure which has formulated, Theochron, a controlled-release theophylline product for market.

## The Pennkinetic System

This system makes use of two controlled-release technologies: ion exchange and membrane diffusion control. It is formed by reacting a drug in its ionic state with a suitable polymer matrix. This polymer-drug complex is then subjected to PEG 4000, to impart plasticity and stability to the complex. A coating of ethylcellulose is applied to the preparation to form an ionic and drug-permeable but water-insoluble coating. To be effective, this system requires that a drug interact ionically with the ion-exchange polymer. Drug release from the Pennkinetic system is quite precise and unaffected by variations in pH, temperature, or contents in the stomach or intestine, since the ion concentration in the human GI tract is very consistent. Additionally the Pennkinetic matrix system makes the drug tasteless, which is helpful in masking the bitter taste of many drugs, especially in pediatric formulations.

The Pennkinetic system is Pennwalt's proprietary liquid system, which has manufactured a variety of long-lasting nonprescription drug products like Delsym (a 12-h cough product containing dextro-methor-phan), Corsym and Cold Factor 12 (12-h cold preparations containing chlorpheniramine and phenylpropanolamine respectively).

## Bioadhesive Systems

One of the simplest concepts for prolonging the duration of drug presence in the GI tract and locating it in a specific region involves binding the product to the mucin/epithelial surface of the GI tract. This is the hypothesis of bioadhesion. Although the concept is old, it is receiving transformed attention because of a better understanding of polymers and the GI tract. Now it is possible to attach a number of polymeric substances non covalently to mucous tissue and keep them localized for an extended period of time. One of the early researchers in this area was Nagai et al., who used the bioadhesion principle to treat inflammation of the mouth. He used an antiinflammatory drug mixed with a bio-adhesive polymer that would attach to the cheek or tongue and remain for many hours. He has since extended the clinical application of this concept to the nasal delivery of peptides and the treatment of cervical cancer. Extended, local release of the drug has yielded good clinical results in the treatment of both aphtha and cervical cancer. Experimental work at University of Wisconsin, with drugs in the eye and GI tract, has led to once-daily administration and revealed a number of other advantages, including improved duration of blood or local drug levels, improved fraction of dose absorbed, improved local drug targeting, strategies for platforms for enzyme inhibition (peptidases), drug-polymer prodrugs, and membrane permeability for enzyme change in a restricted area. One of the great prospective of an oral bioadhesive is in its use with peptide drugs. Protein and polypeptide drugs, which are expected to increase substantially in number as a result of genetic engineering, are subject to peptidase inactivation in the GI tract and commonly have difficulty because of their size while crossing the intestinal barrier. To overcome this unease, it is necessary to localize a dosage form in a specific region of the GIT to inhibit local peptidase activity and perhaps modify intestinal membrane permeability. Bioadhesives offer the potential to partially bring about this goal and also offer the best and most signicant opportunity to improve controlled oral drug delivery.

# PRODUCT EVALUATION AND TESTING

## In vitro Measurement of Drug Availability

It is not possible to simulate in a single in vitro test system the range of variables that affect drug release during the passage of sustained release medication through the GI tract. Properly designed in vitro tests for drug release serve two important functions, however. First, data from such tests are required as a guide to formulation during the development stage, prior to clinical testing. Second, in vitro testing is necessary to ensure batch-to-batch uniformity in the production of a proven dosage form. Different methods are usually required by these two distinctly different testing situations. Although attempts to correlate in vitro release profiles with clinical performance are useful once sufficient clinical testing has been completed, in-vitro/in-vivo correlation must not be assumed. In vitro studies are not sufficient to establish the efficacy of a new preparation.

Tests developed for the purpose of quality control are generally limited to USP dissolution testing methods, using either the rotating basket (apparatus 1), the paddle (apparatus 2), or the modified disintegration testing apparatus (apparatus 3). In many instances in which USP test procedures are followed, upper and lower limits are specified for drug release in simulated gastric and/or intestinal fluid. Measurements are made at specified time intervals appropriate to the specific product. Complete release profiles are not measured unless automated techniques are used. At present, there are no specific USP specifications for sustained release dosage forms. Procedures are determined by nature of the dosage form (e.g. tablet or capsule), the principle utilized to control drug release (e.g. disintegrating or nondisintegrating), and the maintenance period.

During formulation development, testing methods should be designed to provide answers to the following questions.

1. Does the product “dump” maintenance dose before the maintenance period is complete? Sustained release products are subject to either of two modes of failure: Insufficient dose is released, or too much drug is made available too quickly.
2. What fraction of the dose remains unavailable, i.e. what fraction will not be released in the projected time of transit in the GI tract?

3. What is the effect of physiologic variables on drug release? For example, delayed gastric emptying, interaction between drug and GI constituents, composition and volume of GI fluids, and variation in intensity of agitation should be considered.
4. Is the loading dose (if present) released immediately? Is release of the maintenance dose delayed? If so, is the delay time within the desired range?
5. What is the unit-to-unit variation? How predictable is the release profile?
6. What is the sensitivity of the drug release profile to process variables?
7. What is the stability of the formulation with respect to its drug release profile?
8. In short, does the observed release profile fit expectations?

The methods used to measure drug release profiles should have the following characteristics. The analytic technique should be automated so that the complete drug release profile can be directly recorded. Allowance should be made for changing the release media from simulated gastric to simulated intestinal fluid at variable programmed time intervals, to establish the effect of retention of the dosage form in gastric fluid as well as to approximate more closely the pH shifts that the dosage form is likely to encounter in vivo. In addition, the hydrodynamic state in the dissolution vessel should be controllable and capable of variation. The apparatus should be calibrated using a nondisintegrating dissolution standard (e.g. salicylic acid compacts).

Besides the USP dissolution testing apparatus, testing equipment used for in vitro testing of sustained action formulations have included the rotating bottle, stationary basket/rotating filter, Sartorius absorption and solubility simulator, and column-type flow-through assembly. The rotating bottle method was developed for evaluation of sustained release formulations. Samples are tested in 90 ml bottles containing 60 ml of fluid, which are rotated end over end in a 37° bath at 40 rpm. The method is not adaptable to automated analysis, however, or to easy manipulation of the dissolution media. The Sartorius device includes an artificial lipid membrane, which separates the “dissolution” chamber from a simulated plasma compartment in which drug concentrations are measured. Alternately, a dialysis type membrane may be used. Systems of this type are advantageous in measuring release profiles of disintegrating dosage units, and suspension, granular, and

powdered material, if the permeability of the membrane is properly defined. The column flow-through apparatus possesses similar advantages since drug release is confined to a relatively small chamber by highly permeable membrane filters. This apparatus is flexible, well-defined, and meets all the necessary requirements for measurement of drug release profiles from sustained release dosage forms. It can also be adapted to measurements under near sink conditions if the release medium is passed only once through the dissolution chamber, directly measuring the rate of release. Alternately, the dissolution fluid might be recirculated continuously from the reservoir, allowing measurement of the cumulative release profile. The composition of the release media as well as the flow rate can readily be altered.

The time of testing may vary from 6 to 12 hours, depending on the design specifications of the dosage form. If formulations contain retardants whose function depends on the action of normal constituents of the GI fluids (e.g. bile salts, pancreatin, and pepsin), then the appropriate materials must be included in the simulated release media. Apparatus of the Sartorius type would be advantageous in these circumstances if the analytic procedure for the drug would be adversely affected by the presence of these substances. Otherwise, the simulated fluids consisting of pH 1.2 and 7.2 buffers, as well as intermediate pH-values, which represent the transition between gastric and intestinal pH, would suffice at 37°. For insoluble dosage forms, flow rates should be set at the practical maximum to minimize diffusion in the hydrodynamic layer at the dosage form interface as a significant factor affecting release. For disintegrating or erodable units, measurements should be obtained at several rates of fluid flow to establish the effect of this variable, so that conditions can then be established to generate a release profile encompassing the length of time the dosage form is designed to release drug in vivo. Encapsulated products should be removed from the capsule shell for testing.

Figure 16.14 shows the observed release profile of an encapsulated slow release bead form of papaverine hydrochloride, measured in a flow through apparatus at three pH-values. No significant effect of flow rate was observed for this product. The profiles show that the amount of drug released depends on the length of time the formulation is in contact with gastric fluid, and more significantly, on the length of time it is exposed to media of pH 4 to 5, i.e. the transition between gastric and intestinal pH. In effect, release is controlled by the dissolution characteristics of the drug and apparently not by the dosage

form design. Papaverine has maximum solubility at pH 4.5. These data demonstrate the importance of measuring release profiles from dosage units exposed to a variety of conditions in a single test.

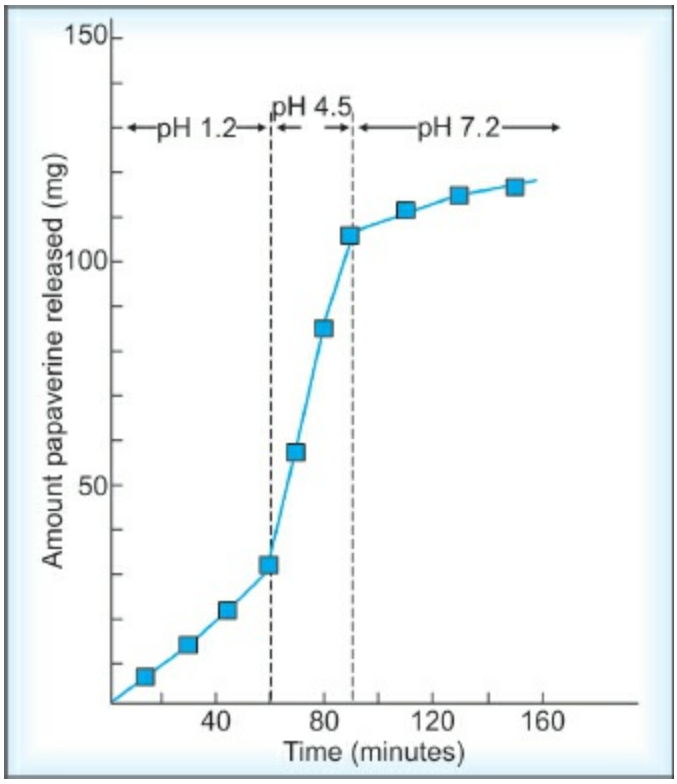



**Fig. 16.14:** In vitro release profile of a papaverine hydrochloride sustained release capsule

As with all pharmaceutical dosage forms, stability testing is an important aspect of the development stage. The same standards that apply to conventional dosage forms with respect to active ingredient stability and dosage form integrity should be used. The stability testing program includes storage of the formulation under both normal (shelf) and exaggerated temperatures so that appropriate extrapolations for long-term stability can be made. The stability of the release profile in addition to that of the active ingredient must be assessed.

Most sustained release formulations are complex. They may be formulated with ingredients that often present special problems regarding their physical stability upon storage. Furthermore, accelerated stability testing may induce changes in some systems (e.g. polymorphic or amorphous to crystalline transitions); these changes would not be observed under normal

shelf storage conditions. In addition, observed release profiles measured after storage at elevated temperatures reflect loss of drug due to degradation. Consequently, predictions of long-term release profile stability based on accelerated tests could lead to erroneous conclusions. The stability testing program for a sustained release product cannot be outlined specifically. It depends on the dosage form and its composition.

There are many advantages to treating release-profile data kinetically by using Eqs (16), (17), and (18), or Eq (19), or a first-order approximation, to obtain a rate constant. Confidence limits for the kinetic parameters can be calculated, allowing establishment of limits for the percentage of released drug under limited testing conditions established for purposes of quality control. Comparison of results obtained with the same product using different testing methods as well as comparisons between multiple runs, different lots, samples in stability, and different products can be made more readily.

## In vivo Measurement of Drug Availability

Validation of sustained release product designs can be achieved only by in vivo testing. The basic objective is to establish the bioequivalence of the product for which a controlled release claim is to be made with conventional dosage forms of the formulated drug.[18] Since no unnecessary human testing should be done, animal models, such as dogs, should be used initially during the product development stage to tune the formulation to the desired specifications. It is necessary to verify that dumping or insufficient drug availability are not observed in vivo. Tests in both animal and subsequent human trials should include periodic blood level determinations, comparison of urinary excretion patterns, serial radiophotographs (in humans) to follow the course of the dosage form in the GI tract, and sequential observations of pharmacologic activity. In some instances (e.g. with insoluble core tablets), egested dosage forms should be recovered and assayed for drug content. If drug level cannot be measured in biologic fluids, then the pharmacologic effect must be observed as a function of time, or clinical trials must be designed, to establish the effectiveness of the drug product.

The FDA has promulgated the general bioavailability and bioequivalence requirements for drug products. These are made to ensure that the new drug product meets its controlled release claims that no dose dumping occurs, that performance is consistent between individual dosage units, and that steady-state drug levels obtained with the product are equivalent to currently marketed products with approved new drug applications (NDAs). Reference materials can include the pure drug substance in solution or suspension as well as conventional dosage forms administered according to their usual dosage schedules or according to the dosage schedule of the controlled release product. Bioavailability studies are ordinarily single-dose comparisons of tested drug products in normal adults in a fasting state. A crossover design in which all subjects receive both the product and reference material on different days is preferred. Guidelines for clinical testing have been published for multiple-dose steady-state studies as well as for single-dose studies. Correlation of pharmacologic activity or clinical evidence of therapeutic effectiveness with bioavailability may be necessary to validate the clinical significance of controlled release claims.

Figure 16.15 shows one example of the type of data required in an in vivo

study designed to demonstrate the validity of a sustained release product design. Comparison is made between blood level profiles observed after administration of a single unit of the sustained release product (curve A), a conventional tablet form containing the usual single dose of active ingredient (curve B), the total dose in the sustained release form as a tablet (curve C), and the total dose administered as three divided doses at the recommended dosing interval (curve D). Samples should be taken over a period of 24 hours or three half-lives of the active ingredient at sufficient frequency to permit reasonable estimation of peak concentrations. Pharmacokinetic model-independent methods are best employed to quantitate the data. The data represented in Fig. 16.15 show the equivalence between administration of the sustained release form and that of three divided doses of drug. Equivalence is demonstrated by comparison of measured blood levels and the AUC-values (area under blood level curve).

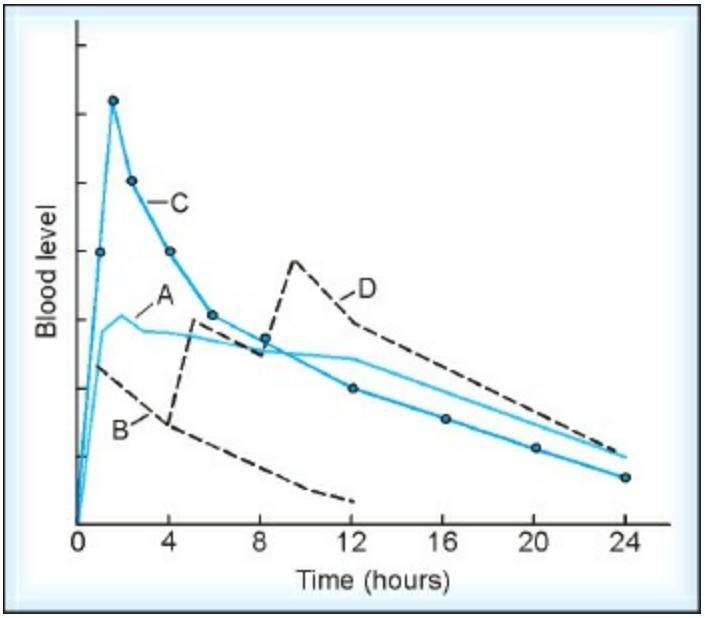



**Fig. 16.15:** In vivo validation of a sustained release tablet of phendimetrazine tartrate: A. 105 mg sustained release tablet, B. 35 mg nonsustained release tablet, C. 105 mg nonsustained release tablet, D. 35 mg nonsustained release tablet administered q4h for three doses

While single-dose studies are usually sufficient to establish the validity of sustained release dosage form designs, multiple-dose studies are required to establish the optimum dosing regimen. They are also required when

differences may exist in the rate but not the extent of absorption, when there is excessive subject-to-subject variation, or when the observed blood levels after a single dose are too low to be measured accurately. A sufficient number of doses must be administered to attain steady-state blood levels. According to an extensive study of sustained release theophylline products, for example, encapsulated forms showed less peaking during multiple dosing, and therefore, better control of blood level within the desired limits.

Attempts to correlate in vivo performance with in vitro availability tests generally have been based on “single-point” measurements. For example, AUC-values, peak blood levels or peak times might be correlated with the time required for 50% of drug to be released in vitro. The best that can be expected from this approach is a rank-order correlation. Significant bioavailability difference between formulations might be masked by improper in vitro methods, or drug release studies might indicate a greater difference than is actually seen in vivo.

Two general approaches to interrelating in vivo and in vitro measurements of drug release have been suggested. In one approach, an in vitro release profile is transformed into a predicted in vivo response. A weighting function characterizing a reference product is determined between the release profile and the average in vivo response, which is measured in a panel of human subjects by the mathematical operation of deconvolution. The in vivo response, predicted in vitro, of the dosage form undergoing testing is obtained by convolution of the observed release profile and the weighting function. The method has been successfully applied to prediction of plasma levels of warfarin and acetazolamide from tablet dissolution data. The technique is computationally complex, but maximizes the amount of information derived from in vitro dissolution testing. Alternately, a reference blood level profile is used as the input to a feedback-controlled dissolution testing apparatus, which is subsequently forced to yield a release profile close to the standard by dynamically changing release media and flow rates. The conditions established using the reference product are used for testing other formulations. Application of these techniques to sustained release products requires a similar formulation as the reference.

In the second approach, the apparent in vivo drug release profile is computed from smoothed blood level or urinary excretion data. This technique requires knowledge of the pharmacokinetic model of the drug. The

in vivo data are used as input to a computer simulation of the pharmacokinetic model; the output represents the amount of drug released at the absorption site as a function of time. Beckett, in applying this method to a sustained release form of phendimetrazine, found that measured in vitro release rates were significantly faster than computed in vivo release rates.

In vivo testing involves a number of simplifying assumptions regarding the uniformity of the absorption process and the suitability of using average data points to represent the population. Since the formulator has no control over physiologic variables, it is essential that clinical studies be based on sufficiently large cross-sections of the population to provide meaningful results. Both in vivo and in vitro testing methods play a major part in validating the effectiveness of sustained release formulations.

# 17:
# Monophasic Liquids

Various types of formulations, both aqueous and nonaqueous, have been placed under the rubric of “liquid preparations,” including solutions, suspensions, and emulsions. In pharmaceutical terms, solutions can be defined as liquid preparations that contain one or more soluble chemical substances, dissolved in a suitable solvent (usually water) or mixture of mutually miscible solvents and they do not, by reasons of their ingredients, method of preparation, or use, fall into another group of products. The oral use of liquid pharmaceuticals has generally been justified on the basis of ease of administration to those individuals who have difficulty swallowing solid dosage forms. A more positive argument can be made for the use of homogeneous liquids (systems in which the drug or drugs are in solution). With rare exceptions, a drug must be in solution in order to be absorbed. A drug administered in solution is immediately available for absorption, and in most cases, is more rapidly and efficiently absorbed than the same amount of drug administered in a tablet or capsule.

The formulation of solutions presents many technical problems to the industrial pharmacist. Some drugs are inherently unstable; this property is magnified when the drug is in solution. Special techniques are required to solubilize poorly soluble drugs. The final preparation must satisfy the requirements of pharmaceutical elegance with regard to taste, appearance, and viscosity. This chapter discusses those factors particularly important in the formulation and manufacture of solutions. The various dosage forms that fall under this general classification are treated as a group. No attempt is made to treat each dosage form individually. Whenever possible, however, specific examples are given to illustrate the application of the principles discussed.

## FORMULATION CONSIDERATIONS

To solve the formulation problems encountered with pharmaceutical liquids, an interesting dichotomy of investigative skills is required. On the one hand, solubility and stability factors can be approached with the precision long associated with the exact sciences; on the other hand, flavouring and other organoleptic characteristics remain subjective factors for which the application of the scientific method still plays a distressingly minor role. Thus, the successful formulation of liquids, as well as other dosage forms, requires a blend of scientific acuity and pharmaceutical “art.”

## Solubility

Whether or not a substance dissolves in a given system and the extent to which it dissolves depend largely on the nature and intensity of the forces present in the solute, the solvent, and the resultant solute-solvent interaction. The nature of these interaction energies and the interplay of electronic and steric factors in determining the solubility of substances in various classes of solvents has been clearly presented by Martin et al. The equilibrium solubility of the drug of interest should be determined in a solvent that is similar to the one intended for use in the final product. This can readily be done by placing an excess of drug (the drug should be finely powdered to minimize the time required to attain equilibrium) in a vial along with the solvent. The tightly closed vial is then agitated at constant temperature, and the amount of drug in solution is determined periodically by assay of a filtered sample of the supernate. Equilibrium is not achieved until at least two successive samplings give the same result.

Solubility studies are generally conducted at fixed temperatures, preferably at temperatures somewhat higher than room temperature (e.g. 30°C), so that constant conditions can be maintained regardless of normal laboratory temperature variations. During the normal distribution process, however, it is possible and even likely that the product will be exposed to a wide range of temperature conditions. For this reason, information relative to the influence of temperature on solubility should be generated. As a rule, a solution should be designed in which the solubility of the solute is not exceeded even at temperatures as low as 4°C.

The approach used when the required concentration of drug exceeds the aforementioned solubility criteria depends on the chemical nature of the drug and the type of product desired.

## *pH*

A large number of modern chemotherapeutic agents are either weak acids or weak bases. The solubility of these agents can be markedly influenced by the pH of their environment. Through application of the law of mass action, the solubility of weakly acidic or basic drugs can be predicted, as a function of pH, with a considerable degree of accuracy. Consider, for example, the reactions involved in the dissolution of a weakly acidic drug, *DH*:

$$DH \text{ (solid)} \rightleftarrows DH \text{ (solution)} \quad \text{... (1)}$$

where, *DH* (solution) is equal to the solubility of the undissociated acid in moles per liter and is a constant generally referred to as $K_S$.

The undissociated acid is also in equilibrium with its dissociation products:

$$K_a = \frac{[D^-][H^+]}{[DH]} \quad \text{... (2)}$$

$$[D^-] = K_a \frac{[DH]}{[H^+]} \quad \text{... (3)}$$

The total amount of drug in solution is the sum of the ionized form [$D^-$] and the unionized form [*DH*]. The equation for total solubility, $S_T$, therefore can be written as:

$$S_T = [DH] + [D^-] = [DH] + K_a \frac{[DH]}{[H^+]} \quad \text{... (4)}$$

since *DH* has previously been defined as equal to $K_S$:

$$S_T = K_S + K_S \frac{[K_a]}{[H^+]} + K_S \left(1 + \frac{K_a}{[H^+]}\right) \quad \text{... (5)}$$

This equation is most useful one for determining the total solubility of a weak acid at a specific hydrogen ion concentration. Since the question most frequently asked is “What must the pH of the formulation to maintain X amount of drug in solution?” a modified form of Eq. (4) is frequently useful:

$$S_T - K_S = \frac{K_S K_a}{[H^+]} \quad \text{... (6)}$$

$$[H^+] = \frac{K_S K_a}{S_T - K_S} \quad \text{... (7)}$$

*For example:* What must the pH of an aqueous formulation to maintain in solution 10 mg/ml of a weakly acidic drug, molecular weight (MW) = 200, $K = 1\times10^{-5}$, $K_S$ = 0.001 M/L?

The desired molar $\left(\frac{g/L}{MW}\right)$ concentration of;

$$\text{Drug} = \frac{0.010\times1000}{200} = 0.05\,\text{M}$$

$$[H^+] = \frac{(1\times10^{-3})(1\times10^{-5})}{0.05-0.001} = \frac{1\times10^{-8}}{0.049}$$

$$[H^+] = 2.04\times10^{-7}$$

$pH = 7 - \log 2.04 = 7–0.31$

$pH = 6.69$

An equation that is useful for poorly soluble, weakly basic drugs can be similarly derived:

$$\text{DOH (solid)} \rightleftarrows \text{DOH (solution)}$$

$K_S$ = DOH (solution) and DOH (solution) is equal to the solubility of the undissociated base in moles/liter.

The dissociation of the weak base can be written as:

$$\text{DOH (solution)} \rightleftarrows D^+ + OH^-$$

$$K_b = \frac{[D^+][OH^-]}{DOH\,(solution)} \qquad \ldots (8)$$

$$D^+ = \frac{K_b\,[\text{DOH}\,(solution)]}{OH^-} \qquad \ldots (9)$$

The total solubility of the base, $S_T$, is the sum of the ionized form $[D^+]$ and the unionized form [DOH]:

$$S_T = [\text{DOH}] + [D^+]$$

$$= [\text{DOH}] + \frac{K_b\,[\text{DOH}]}{OH^-} \qquad \ldots (10)$$

$$S_T = K_S + \frac{K_b K_S}{OH^-} \qquad \ldots (11)$$

Since;

$$K_W = [H^+][OH^-]$$

$$[OH^-] = \frac{K_W}{[H^+]}$$

Then;

$$S_T = K_S + \frac{K_b K_S}{\frac{K_W}{[H^+]}} = K_S + \frac{K_b K_S}{K_W}[H^+] \qquad \ldots (12)$$

Rewriting to solve for $[H^+]$:

$$\frac{S_T}{[H^+]} - \frac{K_S}{[H^+]} = \frac{K_b K_S}{K_W} \qquad \ldots (13)$$

(OR)

$$[H^+] = \frac{K_W}{K_b K_s}(S_T - K_s) \qquad \ldots (14)$$

In practice, these equations hold reasonably well; however, there are limitations that the reader should be aware of:

1. The values for the solubility constant $K_S$ and the dissociation constants $K_a$ or $K_b$ that are reported in the literature (or determined in preformulation studies) are usually for the drug in distilled water. These values may be considerably different in a pharmaceutical dosage form such as an elixir, which contains a high percentage of solids and cosolvents. In general, cosolvents such as alcohol or glycerin have the effect of increasing $K_S$ and decreasing the dissociation constant, as shown in Figs 17.1 and 17.2.

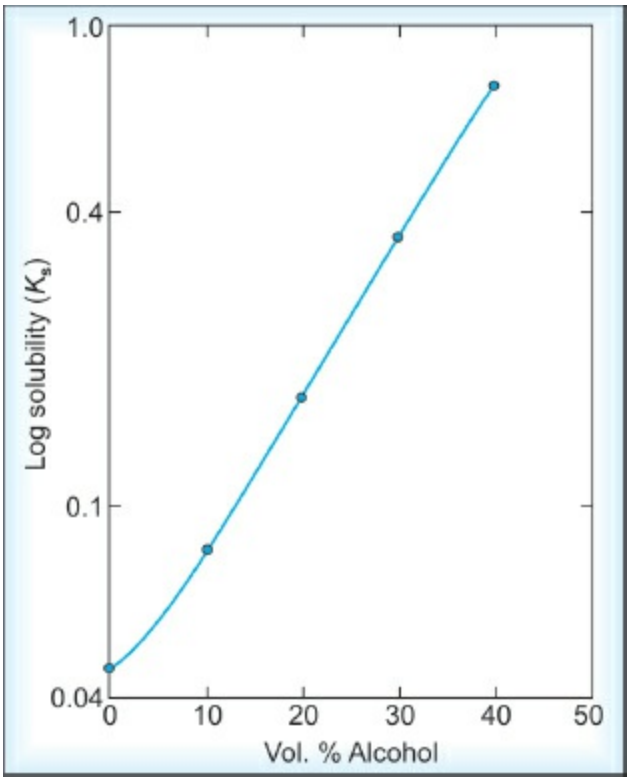



**Fig. 17.1:** Effect of alcohol concentration on the solubility ($K_S$) of un-ionized sulfathiazole

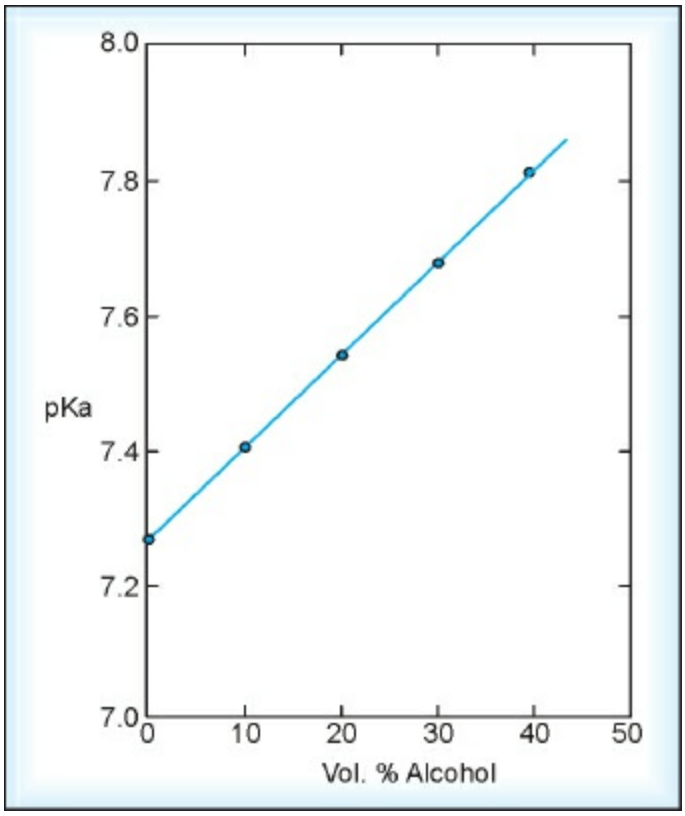



**Fig. 17.2:** Effect of alcohol concentration on the dissociation constant ($K_a$) of sulfathiazole

2. The equations assume little or no interactions between the solute and itself or between the solute and other formulation components. At low concentrations of solute (below several percent), this assumption is generally valid.

In selecting the pH environment for adequate solubility, several other factors should be considered. The pH that satisfies the solubility requirement must not conflict with other product requirements, such as stability and physiologic compatibility. In addition, if pH is critical to maintaining drug solubility, the system must be adequately buffered. The selection of a buffer must be consistent with the following criteria:

1. The buffer must have adequate capacity in the desired pH range.
2. The buffer must be biologically safe for the intended use.
3. The buffer should have little or no deleterious effect on the stability of the final product.
4. The buffer should permit acceptable flavouring and colouring of the

product.

The first three points have been discussed by Windheuser; the last needs no further elaboration. Figure 17.3 is a graphic representation of a number of pharmaceutically useful buffer systems and their effective buffer ranges. In general, a buffer system has adequate capacity within one pH unit of its pK. As an example of research in this area, Wang and Paruta have recently studied the effect of aqueous buffer systems and temperature on the solubility of commonly used barbiturates. Basic information of this type is valuable to the formulator of liquid products.

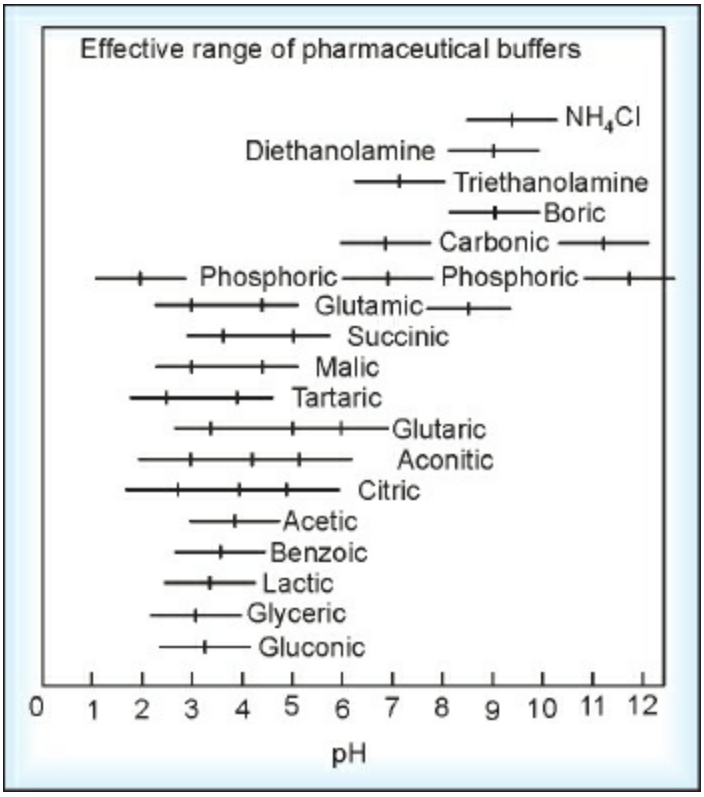



**Fig. 17.3:** Commonly used pharmaceutical buffers and their effective buffer ranges

For many drugs, a pH adjustment does not provide an appropriate means for effecting solution. In the case of very weak acids or bases, the required pH may be unacceptable in terms of physiologic considerations or owing to the effect of pH extremes on the stability of formulation adjuvants (such as sugars and flavours) or of the drug itself. The solubility of nonelectrolytes will, for all practical purposes, be unaffected by hydrogen ion concentration.

In these cases, if solution is to be achieved, it must be done by the use of cosolvents, solubilization, complex phenomenon, or in special circumstances, chemical modification of the drug to a more soluble derivative.

## *Cosolvency*

Weak electrolytes and nonpolar molecules frequently have poor water solubility. Their solubility usually can be increased by the addition of a water-miscible solvent in which the drug has good solubility. This process is known as *cosolvency*, and the solvents used in combination to increase the solubility of the solute are known as *cosolvents*. The mechanism responsible for solubility enhancement through cosolvency is not clearly understood. It has been proposed that a cosolvent system works by reducing the interfacial tension between the predominately aqueous solutions and the hydrophobic solute. The solubility of a substance in a blend of solvents is usually not equal to the value predicted on the basis of its solubility in the pure solvents. For example, undissociated phenobarbital has a solubility of approximately 1.2 g/L in water and 13 g/L in ethyl alcohol. The ratio of solvents, as well as pH, can alter solubility (Fig. 17.4).

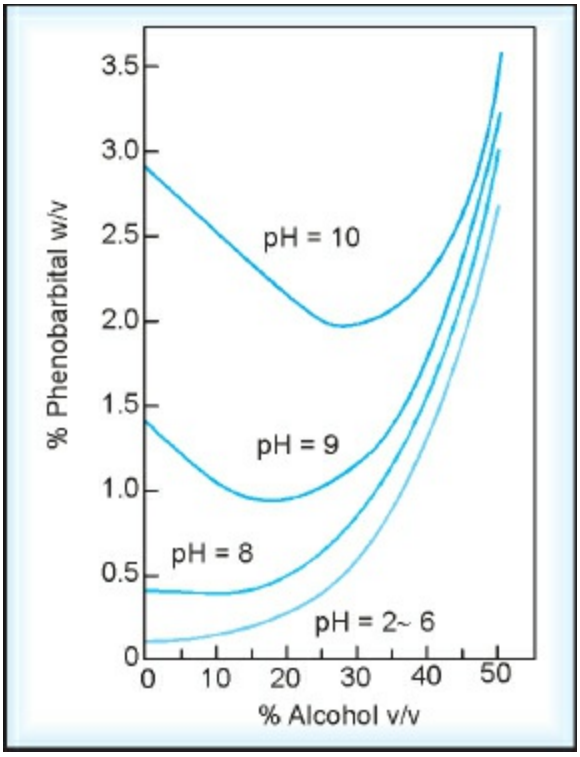



**Fig. 17.4:** Interdependence of pH and alcohol concentration on the solubility

of phenobarbital

Ethanol, sorbitol, glycerin, propylene glycol, and several members of the polyethylene glycol polymer series represent the limited number of cosolvents that are both useful and generally acceptable in the formulation of aqueous liquids. Spiegel and Noseworthy, in their review of non-aqueous solvents used in parenteral products, cited a number of solvents that might also be useful in oral liquids. These include glycerol dimethylketal, glycerol formal, glycofurol, dimethylacetamide, N-(β-hydroxyethyl)-lactamide, ethyl lactate, ethyl carbonate and 1,3-butylene glycol. It should be emphasized, however, that with the possible exception of dimethylacetamide, all of these solvents are unproven with respect to their acceptability for systemic use. Dimethylacetamide has been used as a cosolvent in parenteral products, but its use in oral liquids is seriously limited, owing to the difficulty of masking its objectionable odor and taste. Thus, the spectrum of solvents from which one may make a selection is extremely narrow. Nevertheless, the frequency of their use is high, as can readily be seen by reviewing the formulas for a variety of official and proprietary oral liquids.

Cosolvents are employed not only to affect solubility of the drug, but also to improve the solubility of volatile constituents used to impart a desirable flavour and odor to the product.

Much of the early data on the solubility of pharmaceutical solutes in mixed solvents have been reported as a function of solvent composition; no attempt has been made to explain the data. In recent years, much more emphasis has been placed on cultivating a basic understanding of this phenomenon, with the objective of developing a mathematical approach to interpreting and predicting solubility behavior. Hildebrand and Scott have developed an equation that yields a thermodynamic measure of the cohesive forces that exist within a homogeneous substance. This number is often referred to as *Hildebrand’s solubility parameter*.

There are several serious limitations to the practical application of the solubility parameter concept to pharmaceutical systems. The approach is restricted to what Hildebrand terms “regular solutions”. A regular solution has been defined as one in which there are no interactions between the various solvents present and between the solute and the solvents. All molecules are randomly distributed and oriented in the system. In

thermodynamic language, this may be stated as “a solution involving no entropy change when a small amount of one of its components is transferred to it from an ideal solution of the same composition, the total volume remaining unchanged”.

Martin and coworkers attempted to use this theory in their study concerning the solubility of benzoic acid in mixed solvent systems. The solubility of benzoic acid was found to be in general agreement with the values predicted by the Hildebrand equation, particularly when the solubility parameter of the mixed solvent was approximately equal to the solubility parameter of benzoic acid. The same general conclusions were reached when the solubility of a series of p-hydroxybenzoic acid esters were studied, and the experimental data were compared with the value predicted by the solubility parameter approach. Other authors have extended solubility studies with benzoic acid to other binary and ternary solvent systems.

## *Dielectric Constant*

A more practical, although admittedly less rigorous, approach to the solubility problem is “dielectric requirement” for solubility. According to this theory, every solute shows a maximum solubility, in any given solvent system, at one or more specific dielectric constants.

The absolute solubility of a solute may vary considerably in two different solvents of the same dielectric constant, but the solubility profile, as a function of dielectric constant, appears to be similar for a solute in a wide variety of solvent systems. Solubility profiles as a function of dielectric constants have been reported for numerous pharmaceuticals in a variety of liquid solvent systems. Examples of substances studied include barbiturates, parabens, xanthine derivatives, antipyrine, and aminopyrine.

The dielectric constants of most pharmaceutical solvents are known; values for a number of binary and tertiary blends have been reported, and if not reported, can be readily estimated. Molal boiling point and dielectric constant equations may be used to estimate solubility of pure solvents and miscible solvent blends. The use of each varies, depending upon the literature values and/or laboratory equipment available. To determine the dielectric requirement of the substance of interest, dioxane-water blends having known dielectric constants are used, and the dielectric constant at which maximum solubility is attained is noted. Pharmaceutical formulations of comparable

dielectric constant(s) can then be prepared, and the most appropriate system can be selected on the basis of the solubility requirements, stability, and organoleptic characteristics.

## *Solubilization/Micellization*

Solubilization has been defined by McBain as the spontaneous passage of poorly water-soluble solute molecules into an aqueous solution of a soap or a detergent, in which a thermodynamically stable solution is formed. The mechanism for this phenomenon has been studied quite extensively and involves the property of surface-active agents to form colloidal aggregates known as micelles. When surfactants are added to a liquid at low concentrations, they tend to orient at the air-liquid interface. As additional surfactant is added, the interface becomes fully occupied, and the excess molecules are forced into the bulk of the liquid. At still higher concentrations, the molecules of surfactant in the bulk of the liquid begin to form oriented aggregates or micelles; this change in orientation occurs rather abruptly, and the concentration of surfactant at which it occurs is known as the *critical micelle concentration (CMC)*. Solubilization is thought to occur by virtue of the solute dissolving in or being adsorbed onto the micelle. Thus, the ability of surfactant solutions to dissolve or solubilize water-insoluble materials starts at the critical micelle concentration and increases with the concentration of the micelles.

Solubilizing agents have been used in pharmaceutical systems for many years. As early as 1868, it was reported that cholesterol was markedly more soluble in aqueous soap solutions than in pure water.

In recent years, the application of solubilization phenomenon to pharmaceutical systems has greatly increased. Table 17.1 shows the type of solubilizing agents most frequently used in pharmaceutical systems and the types of drugs for which these agents have been effective. The acceptability of these surfactants for oral use should be determined on an individual basis.

It is readily apparent from this tabulation that a wide variety of substances can be solubilized. McBain has stated, “Any material can be solubilized in any solvent by proper choice of solubilizing agent.” This may well be true, but the questions that must be asked and answered are: To what extent can the substance be solubilized? How is the proper solubilizing agent selected? What effect will the solubilizing agent have on the stability, efficacy, and

physical characteristics of the product?

**Table 17.1:** Solubilizing agents used in pharmaceutical systems

| Solubilizer | Solubilizate |
|---|---|
| Polyoxyethylene sorbitan Fatty acid esters (Tween® series) | Acetomenaphthone<br>21-Acetoxypregnenolone<br>Barbital<br>Caffeine<br>Benzocaine<br>Chloramphenicol<br>Chloroform<br>Chlorotrianisene<br>Cortisone acetate<br>Cyclocoumarol<br>Desoxycorticosterone acetate<br>Dicoumarol<br>Dienestrol<br>Diethyl stilbestrol<br>Digitoxin<br>Volatile oils<br>Essential oils<br>Estrone<br>Ethylbiscoumacetate<br>Hexestrol<br>Menthol<br>Methyltestosterone<br>Phenobarbital<br>Progesterone<br>Reserpine<br>Salicyclic acid<br>Testosterone<br>Vitamin A (alcohol and esters)<br>Vitamin D<br>Vitamin E (alcohol and esters)<br>Vitamin K |
| Polyoxyethylene monoalkyl ethers | Essential oils |

| | |
|---|---|
| (BRIJ* and MYRJ® series) | Volatile oils<br>Benzocaine<br>Benzoic acid derivatives<br>Chloroxylenol<br>Iodine |
| Sucrose monoesters | Vitamin A (alcohol and esters)<br>Vitamin D<br>Vitamin E (alcohol and esters) |
| Lanolin esters and ethers | Essential oils<br>Volatile oils<br>Hexachlorophene<br>Vitamin A palmitate |

® Registered trademarks of ICI United States, Inc.

It has generally been observed that lyophilic surface active agents with hydrophilic-lipophilic balance (HLB) values higher than 15 are the best solubilizing agents. Final selection of solubilizing agents should be based on phase solubility studies in a manner similar to that employed by Guttman et al. in their studies concerning the solubilization of prednisolone, methyl prednisolone, and fluorometholone with Triton WR-1339. They determined the equilibrium solubility of the steroids at 25°C as a function of surfactant concentration. Figure 17.5 is a plot showing the apparent solubility of steroid as a function of Triton WR-1339. A similar plot could be constructed in which the solubility of a specific substance is determined as a function of surfactant concentration, and several surfactants of interest are included in the study. The appropriate surfactant can then be selected on the basis of its efficiency as a solubilizer and its effect on other product characteristics.

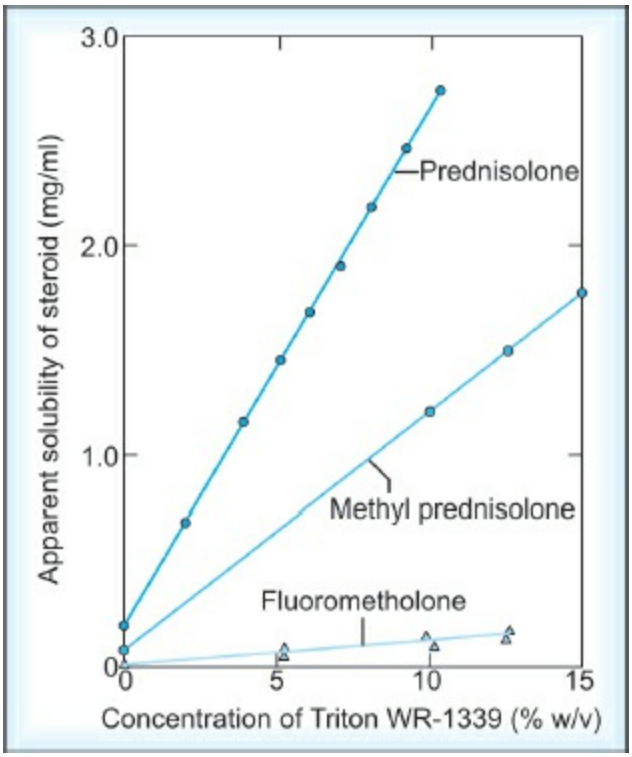



**Fig. 17.5:** The effect of varying concentrations of Triton WR-1339 in water on the solubility of some anti-inflammatory steroids

The major producers of surface active agents have carried out extensive studies on the physiologic effects of products that they recommend for pharmaceutical use. Although the agents are, in themselves, generally free of toxicity, their use must be tempered with a full understanding of the secondary effects that they may produce. While nontoxic surfactants have been shown to improve the stability of vitamin A, other surfactants have been shown to have a deleterious effect on formulation components such as dyes. Addition of surface active agents to drug systems has in some instances enhanced gastrointestinal absorption and pharmacologic activity, and in other cases inhibited the same. Of comparable importance is the effect that surfactants can have on formulation adjuvants. The activity of several preservatives, for example, has been found to be significantly decreased in the presence of a number of surface active agents.

## *Complexation*

Every substance has a specific, reproducible equilibrium solubility in a given solvent at a given temperature. Any deviation from this inherent solubility

must be due to the formation of a new species in solution.* In the case of weakly acidic and basic compounds, the total solubility is equal to the *inherent* solubility of the undissociated compound plus the concentration of the dissociated species. Similarly, when complex formation occurs, the total solubility is equal to the inherent solubility of the uncomplexed drug plus the concentration of drug complex in solution. Consider the interaction between a drug, D, and a complexing agent, C:

$$xD + yC \rightleftarrows D_x C_y \quad \text{... (16)}$$

where, $x$ and $y$ denote the stoichiometry of the interaction. For simplicity, only the case in which one species of complex is formed is considered here; it is possible for several species of complexes to coexist.

The total solubility of drug in this case is:

$$S_T = [D] + X\,[D_x C_y] \quad \text{... (17)}$$

where:

$[D]$ = *the solubility of uncomplexed drug* = $K_S$

$X[D_x C_y]$ = *concentration of drug in complexed form*

By use of the solubility analysis technique, the stoichiometry of this interaction, as well as its equilibrium constant, can be determined. This is carried out by placing excessive quantities of the drug, together with solutions containing various concentrations of complexing agent, in well-closed containers. The containers are agitated at a constant temperature until equilibrium is achieved. Aliquot samples of the supernatant liquid are then removed and assayed for total concentration of drug.

The extent to which the solubility of the drug can be increased is limited by the solubility of the complex. In other cases, the limitation may be imposed by the solubility of the complexing agent. To the pharmaceutical investigator, however, the major concerns are how much drug can be put into solution by a specific complexing agent and how the resultant complex affects the safety, stability, and therapeutic efficacy of the product.

The formulator must also be aware of potential detrimental interaction between ingredients. An example of such interaction is the complexation of nonionic surfactants, such as polysorbate 80 with parabens resulting in the inactivation of the preservatives. Certain polyols have been shown to inhibit

this complexation, thus maintaining paraben antimicrobial activity. Unfortunately, a commonly used polyol, sorbitol, does not inhibit these complexation reactions, probably because it is too polar to partition into the surfactant micelle (Fig. 17.6).

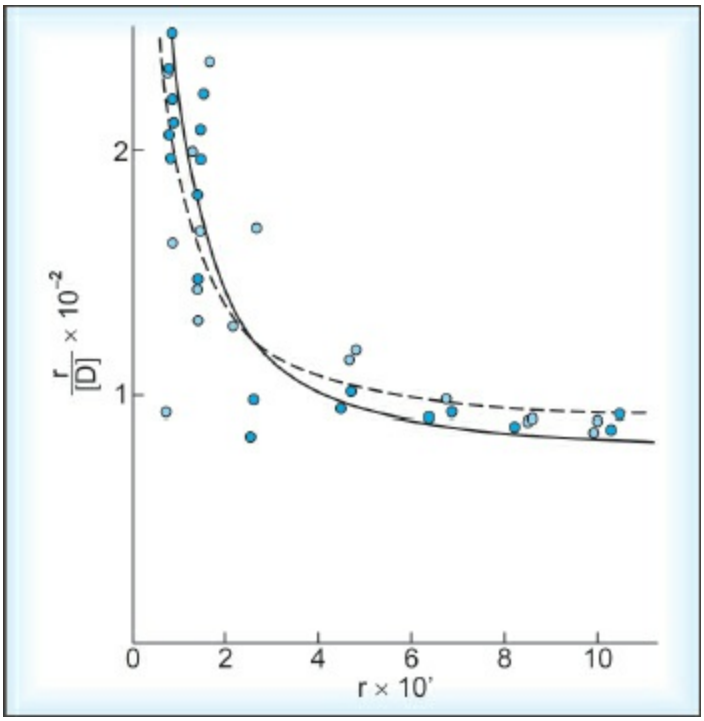



**Fig. 17.6:** Influence of sorbitol on the binding of methylparaben to 10% (w/v) polysorbate 80 to 30. Key: (----) methylparaben alone; and (—) methylparaben plus sorbitol (1:1)

Wolfson and Banker have reported significant increases of barbiturate solubilities as a result of complexation with poly-N-vinyl-5-methyl-2-oxazolidone. Polyvinylpyrrolidone participates in complex formation with a variety of organic and inorganic pharmaceuticals, the most dramatic of which is the well known PVP-iodine complex.

## *Hydrotrophy*

The term hydrotrophy has been used to designate the increase in solubility in water of various substances due to the presence of large amounts of additives. The mechanism by which this effect occurs is not clear. Some workers have speculated that hydrotrophy is simply another type of solubilization, with the solute dissolved in oriented clusters of the hydrotrophic agent. Hydrotrophic

solutions do not show colloidal properties, however, others feel that this phenomenon is more closely related to complexation involving a weak interaction between the hydrotrophic agent and the solute. Still others reason that the phenomenon must be due to a change in solvent character because of the large amount of additive needed to bring about the increase in solubility.

The influence on large concentrations of sodium benzoate on the solubility of caffeine is a classic example of this phenomenon applied to a pharmaceutical system. Other examples include the solubilization of benzoic acid with sodium benzoate, and of theophylline with sodium acetate and sodium glycinate. Except for these examples, little use has been made of hydrotrophy in pharmaceutical systems, probably due to the large amount (in the range of 20 to 50%) of additive necessary to produce modest increases in solubility.

The principle of hydrotrophy has been effectively applied in the solubilization of adrenochrome monosemicarbazone with sodium salicylate (Adrenosem ampoules and oral liquid, Massengill).

## *Chemical Modification of the Drug*

Many poorly soluble drugs can be chemically modified to water-soluble derivatives. This approach has been highly successful in the case of corticosteroids. The solubility of betamethasone alcohol in water, for example, is 5.8 mg/100 ml at 25°C. The solubility of its disodium phosphate ester is greater than 10 g/100 ml, an increase in solubility greater than 1500-fold. In general, however, this approach has severe practical limitations. New derivatives must be subjected to essentially the same testing protocol as the parent compound, including biologic activity studies, acute and chronic toxicity, pharmaceutical evaluation, and clinical testing. An undertaking of this magnitude can be justified only if no other reasonable approach is available.

## *Salt Formation*

Various poorly soluble, ionizable drugs can be readily solubilized by converting them to their salt form since salt formation can drastically alter the physicochemical properties of a drug in the solid state, which govern solubility. For example, an antimalarial agent a[-(2-Piperidyl) b-3,6-bis-trifluoromethyl]-9-Phenanthrenemethanol, is only slightly soluble in water.

However, its lactate salt is approximately 200 times more soluble. This enhanced aqueous solubility is attributed in part to the decrease in crystal lattice energy, as indicated by a reduction in the melting point. If a particular salt form cannot be isolated because of its high solubility, the desired aqueous solubility can be achieved by in situ salt formation. This is accomplished by using an appropriate acid or base to adjust the pH level while formulating the drug product solution.

## Appearance

The overall appearance of liquid products depends primarily on their colour and clarity. Colour selection is usually made to be consistent with flavour, i.e. green or blue for mint, red for berry. The types of colourants available for pharmaceutical use, their relative stabilities, and areas of application have been reviewed by Swartz and Cooper.

The dyes permitted for use in pharmaceuticals vary from country to country. Each regulatory agency also revises its approved list from time to time. Before formulating a product that may be marketed in several countries, it would be wise to check the current status of each dye. Suppliers of the dyes are usually excellent sources of information on this subject.

A purification step invariably is required to achieve maximum clarity. Particulate matter may be introduced through lint and fibers from the solvent or trace quantities of insoluble contaminants in one or more of the formulation components. It is quite common, for example, for alcoholic solutions of natural flavours to precipitate pectins and resins on addition to the bulk aqueous solution. Their removal and removal of other particulate matter is referred to as “polishing,” and technically may be accomplished (1) by settling and subsequent decantation, (2) by centrifugation and (3) by filtration. Filtration is the only practical method when large volumes of liquid are involved. A discussion of clarification and filtration procedures can be found in Chapter 5. It is in order at this point, however, to mention some of the effects that this process can have on the product. The filter pads most frequently used for oral liquids were formerly composed either entirely of asbestos or of mixtures of asbestos and cellulose. With the finding that asbestos fibers can cause cancer, liquids are now filtered, whenever possible, through membrane filters. A number of manufacturers make available membrane filters in a variety of materials and pore sizes. Combined with filter aids and prefilters, they can be used to filter most pharmaceutical liquids.

Studies should be carried out before and after filtration to determine the extent, if any, to which actives, preservatives, flavours, colourants, and other important product components are adsorbed. Production conditions should be simulated as closely as possible, and particular attention given to filtration rate and liquid-to-filter surface area ratio.

Adsorption observed in the filtration of small batches in which the ratio of adsorptive surface to liquid volume is large may be misleading. Under production conditions, when the ratio of adsorptive surface to liquid volume is small, this effect may be insignificant. If adsorption is a significant problem and a non-adsorptive filter cannot be found, a satisfactory filtration process may require the use of an appropriate over-charge or preequilibration of the filter medium with the formulating component(s) being adsorbed.

A much less common problem is that of extraction of materials from the filter pad. This is only of concern in so far as the extracted material may affect the physical and chemical stability of the product. (Pharmaceutically acceptable filter pads and aids contain no biologically active components.) For this reason, stability studies must be carried out on product made by the same process as the one to be used for ultimate production.

## Stability

### *Chemical Stability*

Techniques for predicting chemical stability of homogeneous drug systems are well-defined. Chemical instability of a drug invariably is magnified in solution, as opposed to solid or suspension systems. This liability, however, is to a large extent offset by the rapid and accurate stability predictions, which are possible with homogeneous systems but are extremely risky with heterogeneous dosage forms.

Studies involving evaluation of stability in liquid drug systems include the effect of amino acids on the stability of aspirin in propylene glycol solutions, and a systematic study of the auto-oxidation of polysorbates.

### *Physical Stability*

A physically stable oral liquid retains its viscosity, colour, clarity, taste, and odor throughout its shelf-life. All of these characteristics can and should be evaluated subjectively, and objectively if possible, during the course of stability assessment. A freshly made sample should serve as a reference standard for subjective evaluations.

Objective measurements should be made when such tests are practical. Colour can readily be measured spectrophotometrically, and the absorbance at the appropriate wave length of aged samples can be compared with the initial value to determine the extent of colour change. Clarity is best determined by shining a focused beam of light through the solution. Undissolved particles scatter the light, and under these conditions, the solution appears hazy. Light-scattering equipment is available to give a quantitative measure of turbidity. In general, however, this type of measurement is not necessary in the evaluation of oral liquids. In most cases, a liquid that becomes noticeably turbid with age is unacceptable. A quantitative measure of the turbidity is of little importance except as a tool in determining factors that influence the rate at which the liquid becomes turbid.

Taste and odor continue to be subjectively evaluated by the formulation chemist. This is done either by the pharmaceutical investigator, or preferably, by a panel of unbiased, taste-sensitive individuals. Coded aged samples are submitted to each panel member along with a similarly coded reference

sample. The odor and taste of the reference sample should be known to be intact. Product that has been stored in the refrigerator frequently is used as the reference sample. Each panel member is asked to taste and compare the coded samples. If most of the panel members cannot detect a difference between samples, obviously the organoleptic character of the aged sample has not significantly changed. Some time should be allowed to pass between the preparation of a flavoured sample and taste testing. Aging for about two weeks to one month is recommended to permit flavour blending.

Some attempts have been made, mainly by the perfume and flavour industry, to characterize products by vapor phase chromatography. This approach has proven useful with respect to pure flavours and perfumes, but has rarely been successfully applied to a finished pharmaceutical product.

An integral part of any stability study should include consideration of the possible effect of the package on the contents as well as the effect of the contents on the package. For this reason, stability studies that are intended to support a New Drug Application and/or marketing of a drug must be carried out on the package intended for ultimate use. A well-designed stability protocol should call for a thorough evaluation of both the package and its contents at various conditions of storage, including exposure to both natural and artificial light of known amounts.

It is important to store the final product in the same container in which it was marketed until its expiration date. Flavours and colours often change with time, owing to adsorption by plastic containers or closures, or evaporation of the solvent, with a resultant concentrating of the product or chemical breakdown. One example of chemical breakdown is oxidative breakdown induced by repeated opening of a pint or gallon container to dispense prescriptions. This results in a new source of oxygen being introduced and a larger headspace with each dispensing.

Most oral liquids are packaged in either amber or flint glass containers with plastic or metal caps. Fortunately, glass is generally inert to aqueous solutions in the pH range appropriate for oral liquids. The same is not necessarily true for the cap and liner. Plastic caps may undergo stress cracking on contact with some liquids, whereas under some conditions, corrosion may be a problem with metal caps. In both cases, it is important to select a liner that on the basis of actual testing, is compatible with the package contents.

The integrity of the seal between a cap and container depends on the geometry of the cap and container, the materials used in their construction, the composition of the cap liner, and the tightness with which the cap is applied. Torque is a measure of the circular force, measured in inch-pounds, applied in closing or opening a container. The application and removal torque should be considered an integral part of any pharmaceutical development project involving a threaded package and is of particular importance with respect to liquid formulations. An inadequate cap seal may result in excessive loss of volatile components or leakage of product from the container. Extreme tightening may deform, or break closures and make the cap excessively difficult to remove. Product prepared for stability evaluation should be capped with essentially the same torque as anticipated for use in production.

The optimum application torque for closures and containers varies, depending on the material used in their manufacture. The proper application torque should be determined experimentally by the use of containers and closures of the same size, neck finish, and composition as those intended for use in the final product. The recommended application torque is generally a compromise between the torque providing maximum product protection and the torque that allows for convenient cap removal.

## ADDITIVES AND COMPONENTS

### Preservatives

In recent years, adequate preservation of liquid products has increased in importance. Reports of clinical complications arising from microbial contamination of oral and topical products have originated in several European countries and the United States. Numerous product recalls and tightened regulatory and compendia limits have re-emphasized the need for the formulator to carefully and thoroughly consider all aspects of the preservative system chosen for a particular formula. In addition to presenting a health hazard to the user, microbial growth can cause marked effects on product stability.

Numerous sources of contamination exist. Including among these are raw materials, processing containers and equipment, the manufacturing environment, operators, packaging materials, and the user.

Manufacturing techniques to minimize microbial contamination are presented under the heading “Manufacturing Considerations.” The remainder of this section deals with preservative systems for liquid products.

An ideal preservative can be qualitatively defined as one that meets the following three criteria:

1. It must be effective against a broad spectrum of microorganisms.
2. It must be physically, chemically, and microbiologically stable for the lifetime of the product.
3. It must be nontoxic, nonsensitizing, adequately soluble, compatible with other formulation components, and acceptable with respect to taste and odor at the concentrations used.

No single preservative exists that satisfies all of these requirements for all formulations. The selection of a preservative system must be made on an individual basis, using published information and “in house” microbiologic studies for guidance. Frequently, a combination of two or more preservatives are needed to achieve the desired antimicrobial effect.

The antimicrobial agents that have been used as preservatives can be classified into four major groupings: acidic, neutral, mercurial, and quaternary ammonium compounds. Table 17.2 lists some representative

members of these groupings and the concentration ranges at which they have been used.

**Table 17.2:** Some pharmaceutically useful preservatives

| **Class** | **Usual Concentration %** |
|---|---|
| *Acidic* | |
| Phenol | 0.2–0.5 |
| Chlorocresol | 0.05–0.1 |
| o-phenyl phenol | 0.005–0.01 |
| Alkyl esters of parahydroxy-benzoic acid | 0.001–0.2 |
| Benzoic acid and its salts | 0.1–0.3 |
| Boric acid and its salts | 0.5–1.0 |
| Sorbic acid and its salts | 0.05–0.2 |
| *Neutral* | |
| Chlorbutanol | 0.5 |
| Benzyl alcohol | 1.0 |
| o-phenylethyl alcohol | 0.2–1.0 |
| *Mercurial* | |
| Thimerosal | 0.001–0.1 |
| Phenylmercuric acetate and nitrate | 0.002–0.005 |
| Nitromersol | 0.001–0.1 |
| *Quaternary ammonium compounds* | |
| Benzalkonium chloride | 0.004–0.02 |
| Cetylpyridinium chloride | 0.01–0.02 |

The phenols are probably the oldest and best known pharmaceutical preservatives, but are little used in oral pharmaceuticals, owing to their characteristic odor and instability when exposed to oxygen. The more useful members of the series, for this application, are the p-hydroxybenzoic acid

esters, and the salts of benzoic and sorbic acid. They are adequately soluble in aqueous systems and have been demonstrated to possess both antifungal and antibacterial properties.

Frequently, a combination of two or more esters of parahydroxybenzoic acid are used to achieve the desired antimicrobial effect. Methyl and propyl parahydroxybenzoic acid, for example, are often used together in a ratio of 10 to 1, respectively. The use of more than one ester makes possible a higher total preservative concentration, owing to the independent solubilities of each, and according to some researchers, serves to potentiate the antimicrobial effect. The solubilities of a series of parabens have been studied at four temperatures. The solubilities were expressed in terms of ideal, actual, and excess free energies.

The remaining three classes of preservatives have been widely used in ophthalmic, nasal, and parenteral products, but have been little used in oral liquids. The neutral preservatives are all volatile alcohols, and their volatility introduces odor problems as well as concern for preservative loss on aging. The mercurials and quaternary ammonium compounds are excellent preservatives. They are, however, subject to a variety of incompatibilities, with mercurials being readily reduced to free mercury and the quaternary compounds being inactivated by a variety of anionic substances. The incompatibilities common to these and other preservatives are discussed by Lachman.

Syrups containing approximately 85% sugar resist bacterial growth by virtue of their exosmotic effect on microorganisms. Syrups that contain less than 85% sucrose, but a sufficient concentration of polyol (such as sorbitol, glycerin, propylene glycol, or polyethylene glycol) to have an exosmotic effect on microorganisms, similarly resist bacterial growth. It is possible, however, for surface dilution to take place in a closed container as a result of solvent evaporation followed by condensation, with the condensate flowing back onto the liquid surface. The resulting diluted surface layer makes an excellent medium for bacterial and fungal growth. These products, therefore, should be designed so that even after dilution, they do not support microbial growth. This can be done either by incorporating a sufficient concentration of preservative, so that a diluted sample of the product resists microorganism growth, or by including approximately 5 to 10% ethanol in the formulation. The vapor pressure of ethanol is greater than that of water and normally

vaporizes to the surface of the liquid and the cap area, preventing, or at least minimizing, the potential for microorganism growth as a result of surface dilution.

An effectively designed preservative system must retain its antimicrobial activity for the shelf-life of the product. To ensure compliance with this precept, the preservative characteristics of the product in its final form (including formulation and package) must be studied as a function of age. The best method of demonstrating preservative characteristics is by microbiologic evaluation.

To determine whether a specific organism is hazardous, one must consider the nature of the product and its dose, the state of health of the user, and clinical reports on the frequency and severity of infections caused by the microorganism.

The FDA distinguishes between organisms that are “always objectionable” and “usually objectionable.” The former designation is based on only two factors-pathogenicity of the organism and site of use. The latter designation is based on an additional determinant, the state of health of the user. The official compendia are continually reevaluating their standards based on the latest FDA data and guidelines.

Specific organisms generally recognized as undesirable in oral liquids include *Salmonella* species, *Escherichia coli, Enterobacter* species, *Pseudomonas* species (commonly *P. aeruginosa*), proteolytic species of *Clostridium*, and *Candida albicans*. Some liquid pharmaceuticals (i.e. ophthalmic solutions) must be processed aseptically and rendered sterile.

Chemical analysis for the antimicrobial constituent frequently provides a helpful guide but can be misleading. Molecular interactions involving preservatives and commonly used pharmaceutical adjuvants, such as surfactants and cellulose derivatives, have been observed. For example, it has been shown that Tween 80 interacts to a significant extent with the methyl and propyl esters of parahydroxybenzoic acid, and that the preservative-surfactant complex is essentially devoid of antibacterial activity. Chemical analysis for the parahydroxybenzoate esters would not differentiate between the unbound substance (microbiologically active) and the bound substance (microbiologically inactive).

## Sweetening Agents

Sweetening agents generally constitute a major portion of the solid content in those dosage forms requiring them. Sucrose has had a long history of use. It is soluble in aqueous media (solutions containing approximately 85% sucrose can be prepared); it is available in highly purified form at reasonable cost, and it is chemically and physically stable in the pH range of 4.0 to 8.0. It is frequently used in conjunction with sorbitol, glycerin, and other polyols, which are said to reduce the tendency of sucrose to crystallize. One of the manifestations of sucrose crystallization is caplocking, which occurs when the product crystallizes on the threads of the bottle cap and interferes with cap removal. This phenomenon has been studied, and several vehicles containing sucrose, glucose, sorbitol, and glycerin have been reported as acceptable in terms of product characteristics and resistance to cap-locking.

Liquid glucose is an extremely viscid substance that imparts both body and sweetness to liquid formulations. It is prepared by the partial hydrolysis of starch with strong acid, and contains, as its main component, dextrose with smaller amounts of dextrins and maltose. In a manner similar to that of honey and molasses, but to a lesser degree, this agent imparts a characteristic odor and flavour to the formulations in which it is used. Although liquid glucose is not a pure chemical entity, its method of manufacture can be well controlled, and batch-to-batch variability is usually not a significant problem. The same is not true for such materials as honey and molasses. The quality of these substances varies, depending on the source from which they are obtained, and if the source is held constant, depending on the time of year they are produced and on other natural factors over which there is little or no control. The use of these and other naturally occurring materials should be predicated on a rigorous quality control regimen, which gives maximum assurance of product uniformity.

Saccharin is used to supplement sugars and polyols as sweeteners. It is approximately 250 to 500 times as sweet as sugar, but it can have a bitter aftertaste if not properly used in the formula. A new synthetic sweetener, aspartame, has been approved in numerous countries for use as a food and/or drug ingredient. Aspartame is the methyl ester of aspartic acid and phenylalanine (Fig. 17.7). It is approximately 200 times sweeter than sucrose and has none of the aftertaste of saccharin.

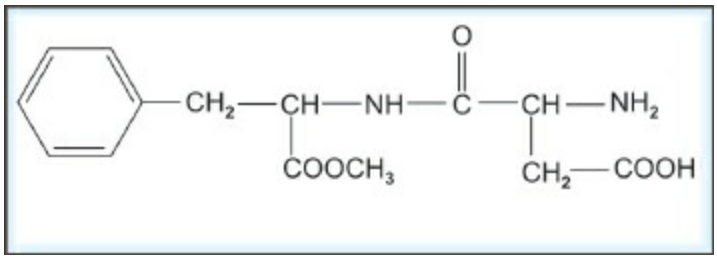



**Fig. 17.7:** Structure of aspartame

Aspartame's aqueous solubility is quite adequate for formulation purposes (Fig. 17.8).

Although it is very stable as a dry powder, its stability in aqueous solutions is quite pH-and temperature-dependent. Aspartame's greatest stability is between pH 3.4 and 5.0 and at refrigerated temperatures (Fig. 17.9 and Table 17.3). Sweetness enhancement by aspartame is synergistic with saccharin, sucrose, glucose, and cyclamate. In addition, its taste properties have been improved using sodium bicarbonate, gluconate salts, and lactose.

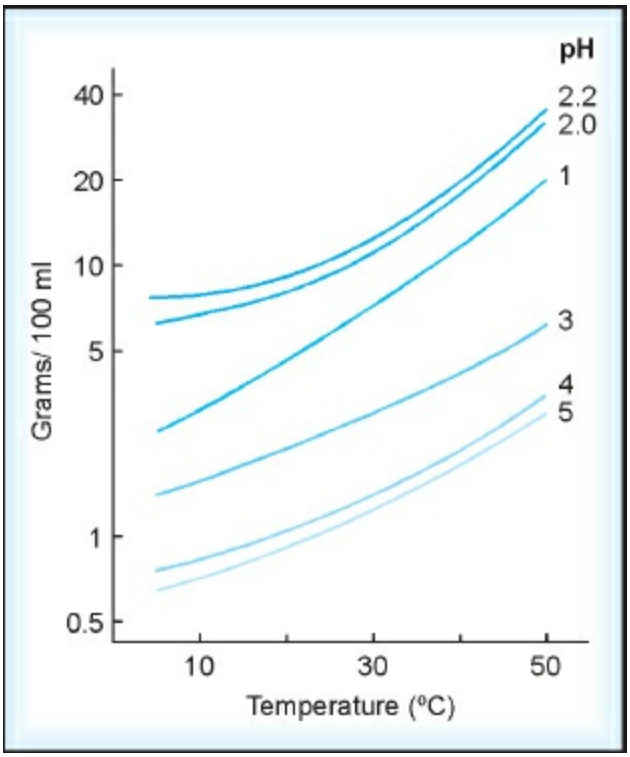



**Fig. 17.8:** Effect of pH and temperature on aspartame solubility in water

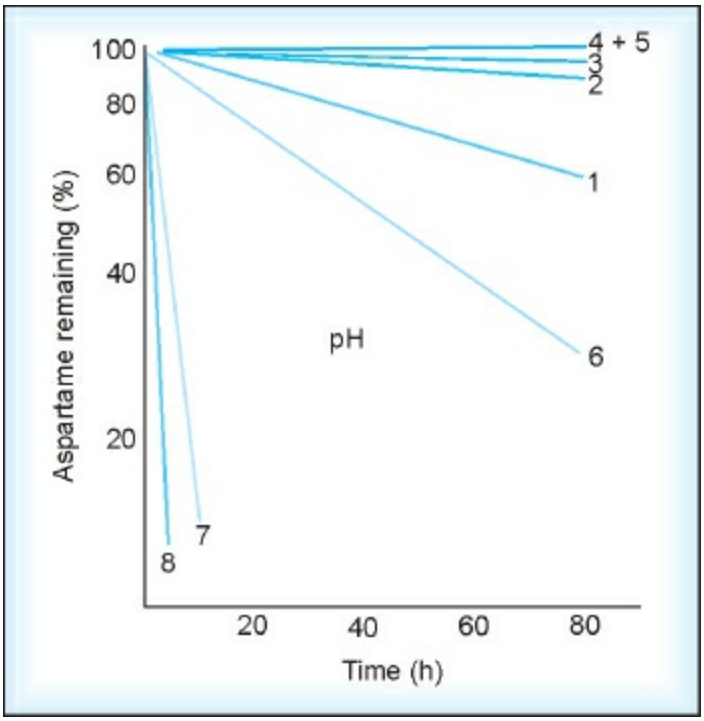



**Fig. 17.9:** Effect of pH and time on aspartame stability in aqueous buffer systems at 40°C

**Table 17.3:** Effect of storage temperature on aspartame stability in aqueous solutions at pH 4.0

| Temperature storage (°C) | Calculated time for 20% decomposition (days) |
|---|---|
| 10 | 387 |
| 20 | 134 |
| 30 | 51 |
| 40 | 22 |
| 55 | 5 |
| 68 | 2 |
| 80 | 1 |
| 90 | 0.15 |

## Viscosity Controlling Agents

It is sometimes desirable to increase the viscosity of a liquid, either to serve as an adjunct for palatability or to improve pourability. This can be achieved by increasing the sugar concentration or by incorporating viscosity-controlling agents such as polyvinylpyrrolidone or various cellulosic derivatives (e.g. methylcellulose or sodium carboxymethylcellulose). These compounds form solutions in water that are stable over a wide pH range. Methylcellulose and carboxy-mediylcellulose are available in a number of different viscosity grades. Carboxymethylcellulose may be used in solutions containing high concentrations of alcohol (up to 50%) without precipitating. It is precipitated, however, as an insoluble salt of a number of multivalent metal ions such as $Al^{+++}$, $Fe^{+++}$, and $Ca^{++}$. Methylcellulose polymers do not form insoluble salts with metal ions, but can be salted out of solution when the concentration of electrolytes or other dissolved materials exceed certain limits. These limits may vary from about 2 to 40%, depending on the electrolyte and the type of methylcellulose involved.

Viscosity-inducing polymers should be used with a degree of caution. They are known to form molecular complexes with a variety of organic and inorganic compounds and thus may influence the activity of these compounds. It is conceivable that highly viscid systems that resist dilution by gastrointestinal fluids might impede drug release and absorption.

## Buffers

During storage of liquid preparations, degradation of the product, interactions with container components or dissolution of gases and vapors causes change in their pH level, which can be prevented by addition of buffer. A suitable buffer system should have adequate buffer capacity to maintain the pH level of the product. Commonly used buffer systems are phosphates, acetates, citrates, and glutamates. Although buffers ensure pH stability, the buffer system can affect other properties such as solubility and stability. The ionic-strength contributions of the buffer systems can affect stability. Buffers can also act adversely as general-acid or general-base catalysts and cause degradation of the drug substance. Therefore, before selecting any buffer system, the effect of buffer species should be studied.

## Antioxidants

Various drugs in solution are subject to oxidative degradation. Oxidation is defined as a loss of electrons from a compound leading to change in the oxidation state of the molecule. Such reactions are mediated by free radicals or molecular oxygen, and are often catalyzed by metal ions. Moreover, oxidation often involves the addition of oxygen (or other electronegative atoms like halogens) or the removal of hydrogen. Drugs possessing favorable oxidation potential are especially vulnerable to degradation. Agents with an oxidation potential lower than that of the drug in question are called antioxidants. Additionally, certain properties of the selected primary packaging (such as polymer degradation, oxygen transmission rates, impurities, etc.) can readily lead to oxidation of drug molecules in solution and hence may require the addition of antioxidants to maintain product stability. They are added to solutions alone or in combination with a chelating agent or other antioxidants and function by being preferentially oxidized and gradually consumed or by blocking an oxidative chain reaction where they are not consumed.

Salts of sulfites are the most common antioxidants in aqueous solutions and their antioxidant activity depends on their final concentration and the final pH level of the formulation. Generally sodium metabisulfite is used at low pH, sodium bisulfite at near neutral pH, and sodium sulfite is used at basic pH. A combination is often used since single antioxidant may provide incomplete protection. Certain compounds (e.g. citric and ascorbic acids) have been found to act as synergists, increasing the effectiveness of antioxidants, particularly those that block oxidative reactions. Often, chelating agents such as edetic acid derivatives (EDTA) are used in formulations containing trace amounts of heavy metals that would otherwise catalyze oxidative reactions. Moreover, synthetic phenolic compounds, such as butylated hydroxytoluene (BHT) and butylated hydroxyanisole (BHA) serve as hydrogen atom donors and can successfully prevent oxidation of oils and fats in oral liquid formulations.

## Flavours

Flavouring can be divided into two major categories: selection and evaluation. Much has been written on both phases of pharmaceutical flavouring, but selection remains a totally empiric activity.

The four basic taste sensations are salty, bitter, sweet, and sour. Some generalizations concerning the selection of flavours to mask specific types of taste have been suggested by Janovsky, and by Wesley (Table 17.4).

**Table 17.4:** Flavor selection

| Taste Sensation | Recommended Flavor |
|---|---|
| Salt | Butterscotch, maple, apricot, peach, vanilla, wintergreen mint |
| Bitter | Wild cherry, walnut, chocolate, mint spice, mint combinations, passion fruit, anise |
| Sweet | Fruit and berry, vanilla |
| Sour | Citrus flavors, liquorice, root beer, raspberry |

A combination of flavouring agents is usually required to mask these taste sensations effectively. Menthol, chloroform, and various salts frequently are used as flavour adjuncts. Menthol and chloroform are sometimes referred to as de-sensitizing agents. They impart a flavour and odor of their own to the product and have a mild anesthetic effect on the sensory receptor organs associated with taste. Monosodium glutamate has been widely used in the food industry, and to a lesser extent, in pharmaceuticals, for its reported ability to enhance natural flavours. A carefully selected panel reported this substance to be effective in reducing the metallic taste of iron-containing liquids, as well as the bitterness and aftertaste of a the variety of other pharmaceutical preparations. It cannot be used in pediatric products, however.

Chemburkar and Joslin have reported that the partitioning of parabens into flavouring oils from aqueous systems depends on the concentration of the flavouring oil, the nature and concentration of the additives, and pH.

Wesley’s Pharmaceutical Flavour Guide contains suggestions for

flavouring over 51 types of pharmaceutical preparations. Similar reports provide some guidance for the formulation chemist, but the final selection must result from a trial and error approach. Inherent in this approach is what is referred to as taste fatigue. Repeated samplings of strong tasting substances soon result in decreased flavour acuity, and therefore, impaired ability to evaluate flavour properly. Preliminary flavouring should be carried out on diluted samples. This is done by preparing flavoured vehicles and adding increments of the medicament or other formulation components responsible for the taste problem. The concentration at which the taste of the medicament is perceptible is referred to as the *minimum threshold level*. The vehicles that are most effective in masking low levels of drug are candidates for full-strength flavour evaluation.

Flavour evaluation techniques have progressed to a much greater extent than flavour selection. Taste panels can be useful in selecting one of several candidate formulations. This subject, as well as other flavour considerations, has been surveyed in an excellent book assembled by Arthur D. Little, Inc.

## MANUFACTURING CONSIDERATIONS

The basic principles involved in the preparation of homogeneous liquids are the same regardless of the quantity of material involved. The solubility of the solute and intramolecular and intermolecular interactions in the final solution at equilibrium are independent of the manner in which the solution is made. This assumes, of course, that the method of compounding does not affect the final composition of the system, as would be the case if a volatile component were charged to a heated solution. The rate at which equilibrium is achieved, however, is highly dependent on the details of the compounding procedure and the equipment used.

Liquid processing lends itself to computer-controlled automation. A few pharmaceutical firms have already instituted automated or semi-automated processes for several large-selling liquid products.

## Raw Materials

The raw materials used in manufacturing liquids should conform to well thought out specifications. These specifications should assure identity, purity, uniformity, and freedom from excessive microbial contamination. Incoming raw materials should be impounded and thoroughly tested before they are released for manufacturing. Additional processing may be necessary to obtain a desirable property, such as particle size or freedom from microorganisms. With regard to microbial contamination of raw materials, it is usually much easier to begin with low counts in the raw materials than to try to reduce these counts substantially during processing.

Aside from the active ingredient, water is usually the most important constituent in a liquid product. It should meet the USP requirements for purified water. It may be obtained by distillation or ion-exchange treatment. In recent years, manufacturers have devoted considerable effort to upgrading the microbial purity of the water supply used in oral liquids. Techniques employed include reverse osmosis purification, ultraviolet sterilization, membrane filtration, and constant circulation in piping systems that have no "dead ends" where microorganisms can thrive. In general, the most difficult microbes to remove from a purified water system are the *Pseudomonas* (Fig. 17.10) shows a purified water system designed to minimize microbial growth.

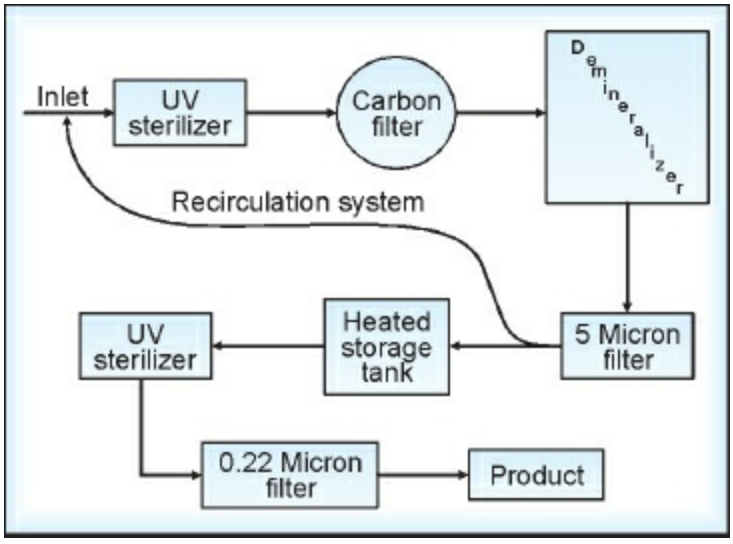



**Fig. 17.10:** Schematic drawing of a deionized water system for the

Many of the ideas incorporated in Fig. 17.10 have been taken from experience gained in preparing and storing pyrogen-free water for injection and adapted to the manufacture of high-quality de-ionized water for use with liquid pharmaceutical products.

The ion-exchange resins used in the water system are important to the successful maintenance of low bacteria counts. An example of an appropriate mixed resin bed would be Ambergard XE-352 and Amberlite IR-120. Ambergard XE-352 is a large-pore, macroreticular, Type 1 quaternary ammonium anion-exchange resin. It is effective for a wide range of flow rates and for many different bacterial strains. Amberlite IR-120 is a strongly acidic, cation-exchange resin that balances the chemical equilibrium of the water.

The use of ultraviolet sterilization of water has been discussed in some detail in the literature. Two factors should be emphasized at this point (1) The flow rate of the water should not exceed the capability of the sterilizing unit and (2) even if sterility is achieved, a filter should still be used downstream to remove the dead microorganisms and particulate matter. Although the first point seems self-evident, it is violated surprisingly often-compromising the effectiveness of the entire system.

A number of in-line filtration units are commercially available. These filters are discussed in Chapter 5.

## Equipment

In the most general terms, the type of equipment used in the manufacture of oral solutions consists of mixing tanks equipped with a means of agitation, measuring devices for large and small amounts of solids and liquids, and a filtration system for the final polishing and/or sterilization of the solution. In addition, most production facilities are equipped with systems for bulk material handling, such as tote bins and tote bin discharging equipment.

All equipment must be thoroughly cleaned and sanitized (sterilized if possible) before use. Appropriate disinfectants include dilute solutions of hydrogen peroxide, phenol derivatives, and peracetic acid. Equipment and lines can be sterilized by such methods as alcohol, boiling water, autoclaving, steam, or dry heat.

Tanks are usually constructed of polished stainless steel and are usually jacketed to allow for heating or cooling of the contents. They can be obtained in a number of different sizes, and are completely covered and equipped with see-through charging ports and illumination for easy observation of the contents. If tanks are used for the compounding of the bulk liquid, they have a built-in agitation system.

Water condensate that forms on the lid of mixing tanks and similar processing equipment during heating and chilling steps may provide a source of microbial contamination that is often overlooked.

The liquid is then clarified by cycling through a filtration system, and the polished solution is stored in an adjacent tank until released by the quality control department. The liquid may then be transported to the filling line, either manually by filling into portable transport tanks or by pumping (or gravity flow) through a suitable liquid delivery conduit.

The distance the product travels between the holding tank and the filling line should be held to a minimum to reduce the chance of microbial contamination. All lines should be easy to disassemble, clean, and sanitize.

A major source of microbial contamination is often the processing operators. Head covering should be worn at all times. Gloves and face masks should be worn as necessary.

An ongoing education program is recommended to maintain operator interest and concern for good work habits.

In additions, the use of portable laminar flow units can be an aid in certain operations (such as the addition of ingredients to a tank).

## Compounding Procedure

Dilute solutions, prepared from rapidly dissolving materials, are simply prepared by charging the solute to the solvent and agitating until the solution is homogeneous. When more concentrated solutions are being made, or when the solute is slowly dissolving, it may be advantageous to employ heat. The syrup formula and manufacturing method presented in Table 17.5 illustrate some of the steps involved in compounding a complex liquid formulation.

**Table 17.5:** Syrup formula and manufacturing method

| Formula | Per ml | Per batch (5000 L) |
|---|---|---|
| Drug | 2.00 mg | 10.0 kg |
| Sodium benzoate USP | 1.00 mg | 5.0 kg |
| Menthol, USP | 0.10 mg | 0.5 kg |
| Alcohol, USP | 0.05 ml (40.8 mg) | 250.0 liters (204.0 kg) |
| Flavor | 0.005 ml (4.5 mg) | 25.0 liters (22.5 kg) |
| Dye FD and C Yellow No. 6 | 0.10 mg | 0.5 kg |
| Glycerin | 0.05 ml (62.45 mg) | 250.0 liters (312.250 kg) |
| Sorbitol solution, USP | 0.10 ml (128.5 mg) | 500.0 liters (642.5 kg) |
| Standard granulated sugar | 550.00 mg | 2750.0 kg |
| Purified water, q.s. to | 1.0 ml | 5000 liters |

## *Compounding Instructions*

1. Charge 2000 L of purified water through the water meter into the compounding tank. Check the volume against the outage chart. Heat to approximately 50°C.
2. To the water in the compounding tank, charge the following materials in the amounts specified in the batch sheet. Dissolve each one, with agitation, before adding the next (a) drug, (b) sodium benzoate, (c)

standard granulated sugar. Agitate the contents of the compounding tank until homogeneous, and then cool to 30°C.

3. Charge the specified amount of glycerin to the compounding tank. Agitate until, batch is homogeneous.
4. Charge the specified amount of sorbitol solution to the compounding tank. Agitate until the batch is homogeneous.
5. Measure 20 L of alcohol into a suitable stainless steel container. Add and dissolve the specified charge of menthol. Add and dissolve the specified charge of flavour.
6. Charge the alcoholic solution of menthol and flavour to the batch in the compounding tank. Agitate until homogeneous.
7. Charge the balance of the specified amount of alcohol to the batch. Agitate until homogeneous.
8. Charge 10 L of purified water to a clean stainless steel container. Add to the water and dissolve the specified amount of FD and C Yellow No. 6.
9. Charge the dye solution to the batch in the compounding tank, and agitate until homogeneous.
10. Add to the compounding tank sufficient purified water to bring the batch volume to 5000 L.
11. Weight out 2.5 kg of filter aid, and charge it to the contents of the compounding tank. Agitate for 10 min. The batch is now ready to filter.
12. Cycle the batch through the filter and back to the compounding tank until the filtrate is clear. At this point, the filtrate may be discharged and collected in the designated holding tank.
13. Sample the batch, and submit for testing in accordance with standard procedure.

The rationale for most of the steps cited in this procedure is obvious. Several steps, however, warrant some discussion. Step 1 calls for the metering of a specific amount of purified water into the compounding tank. The precise quantity of water in this case is not critical, but in spite of this, a confirmatory volumetric check is desirable, to protect against the consequences of a malfunctioning metering device. The purified water is heated in step 2 primarily to facilitate the solution of sucrose, the other solutes being soluble even in cold water. In step 5, the menthol and flavour

are dissolved in an aliquot of the alcohol, and this alcohol solution is charged to the batch (step 6). As previously mentioned, the equilibrium solubility of all solutes is the same regardless of the manner in which they are charged. The rate, at which solution is achieved, however, can be markedly influenced by the compounding procedure. In this case, pre-dissolving the menthol and flavour in alcohol, in which both solutes are highly soluble, and then charging the resultant alcoholic solution to the main part of the batch effect rapid approach to equilibrium conditions.

Solutes present in small concentrations, particularly dyes and other intensely coloured materials, should be predissolved prior to mixing with the main portion of the batch, as is indicated in step 8. This is done to ensure complete solution of the substance before the batch is further processed. If the solutes were charged directly to the bulk mixing tank, it would be extremely difficult to determine the presence of a small amount of undissolved material at the bottom of the tank. As a rule, complete solution should usually be confirmed at every stage in the manufacture of a homogeneous liquid.

Step 11 calls for the addition of a specified amount of filter aid to the contents of the compounding tank. The amount and type of filter aid must be determined during the development of the product. The amount used does not usually exceed 0.5 g/L.

In the laboratory, liquids are usually measured by volume. When large quantities of liquid materials are handled, however, it is frequently more convenient and accurate to use gravimetric means of measurement. For this reason, all liquid components of the cited formula are expressed in units of both volume and weight.

## Packaging

The specific method used for filling a pharmaceutical liquid varies greatly depending on the characteristics of the liquid (e.g. viscosity, surface tension, foam-producing qualities, and compatibility with the materials used in the construction of the filling machine), the type of package into which the liquid is placed, and the required production output. Three basic filling methods—gravimetric, volumetric, and constant level are used for most liquid filling operations. The latter two methods are used most frequently in the filling of pharmaceutical liquids. Filling containers to a given weight (gravimetric filling) is generally limited to large containers or to highly viscous products. The process does not readily lend itself to high-speed, automatic equipment.

Volumetric filling is usually accomplished by positive displacement piston action. Each filling station is equipped with a measuring piston and cylinder. The fill accuracy is controlled by the close tolerances to which the pistons and cylinders are manufactured. The fill amount is measured by the stroke of the piston, which on all machines can be varied to a limited degree. Major changes in fill amount usually necessitate changing the piston and cylinder assembly. This type of device is capable of accuracy to within fractions of a milliliter. There are, however, several significant problems associated with volume filling. Highly viscous liquids may cause the pistons to seize, resulting in either loss of fill accuracy or line breakdown. On the opposite side of the spectrum, thin liquids may flow past the piston, causing uncontrollable dripping from the filling spout and associated fill inaccuracies. These problems can be controlled to a large extent by proper engineering of the filling machine. An inherent problem with volumetric filling, however, is encountered when containers are used that are not dimensionally uniform. In this case, even though the fill amount is accurate, the fill height varies inversely with the container capacity, i.e. an oversized package appears to have a slack fill, whereas an undersized package appears to have an excessive fill.

Constant-level filling uses the container as the means for controlling the fill of each unit. The fill amount is varied by adjusting the height to which the container is filled. Any dimensional variations in the containers result in comparable variations in the net fill per unit. The oldest form of a constant-level filler involves the use of a siphon; however, this method of filling is

usually slow and is rarely used when high production rates are required. The high-speed, automated, constant-level filling machines in use today are generally based on the siphon principle, with the major modification being the induced pressure differential between the liquid discharge nozzle and the constant-level overflow system. The most widely used methods can be broadly classified into three categories: vacuum filling, gravity-vacuum filling, and pressure-vacuum filling.

The principle of vacuum filling is illustrated in Fig. 17.11. To fill by vacuum, a seal must be made between the filling head and the container. A vacuum is then developed within the container, which causes the liquid to flow from the bulk liquid tank to the container. The liquid level rises until it reaches the vacuum tube, which is positioned at the desired constant level. Excess liquid is drawn through the vacuum tube and can be recycled to the bulk liquid tank. In gravity-vacuum filling, the bulk liquid tanks are a level above the filling stem, so that the driving force for liquid flow results from both the negative pressure in the container and the force of gravity. Similarly, in pressure-vacuum filling, a positive pressure is applied to the bulk liquid, which in combination with the vacuum developed in the container, results in a pressure differential that allows for rapid filling of even highly viscous liquids. The latter two methods require some valve mechanism that is responsive to the presence of the container, to open and subsequently close a valve device in the filling stem assembly. Vacuum filters do not require such a mechanism, since a pressure differential to promote liquid flow can only be achieved by the vacuum formed when the filling stem forms a seal with the container.

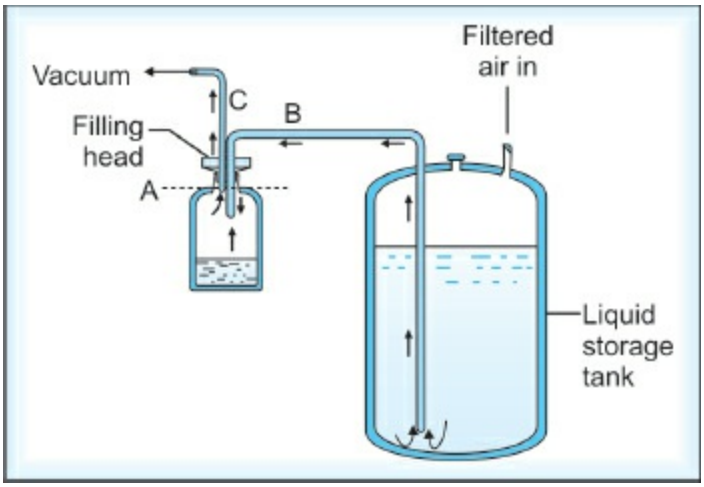


**Fig. 17.11:** Schematic diagram of the principle used in vacuum filling. The vacuum drawn through tube C reduced pressure in the container. Pressure differential causes liquid in storage tank to flow through tube B into container. When liquid level reaches point A, excess is carried out through tube C and may be recycled to liquid storage tank

A problem that is common to all types of liquid filling machines, but that is particularly bothersome with high-speed automatic equipment, is excessive foam. Foaming during the filling operation often can be decreased by filling equipment that minimizes product turbulence, closed system filling to limit the introduction of air or other gases that participate in the formation of foam, mechanical defoaming devices, and reduction in the speed of the filling line.

All of these methods introduce considerable engineering and production difficulties. It would be preferable to formulate the product with careful consideration of the problems that eventually might be encountered in large-scale production and high-speed filling operations.

A microbial survey should be performed on all packaging materials that come into contact with the product to ensure that microbial contamination is not introduced at this point. Attention must also be paid to details during packaging operations. For example, on small volume orders in which bottle closures or tips for plastic squeeze-spray containers are often placed on the product by hand, this procedure can be a source of microbial contamination from the hands of operators unless gloves are used that are presterilized and periodically disinfected during use.

# PHARMACEUTICAL LIQUIDS

## Oral Preparations

### *Aromatic Waters*

Aromatic Waters are saturated solutions of volatile oils or other aromatic substances in water, usually employed for their flavouring rather than their medicinal properties. Aromatic Waters are normally prepared by diluting a concentrated, ethanolic solution of the aromatic substance with water. Their tastes and odors are similar to those of the drugs or volatile substances from which they are prepared, and they are free of empy-reumatic and other foreign odors. Peppermint Water USP and Stronger Rose Water USP are examples of aromatic waters.

Sometimes a problem of *salting out* certain very soluble salts during compounding the aromatic water is observed. Aromatic waters should be made in small quantities and kept in airtight, light resistant containers as they will deteriorate with time, intense light, and excessive heat. Their label states (1) the date after which the aromatic water is not intended to be used, (2) the conditions under which the aromatic water should be stored.

### *Elixirs*

Elixirs are clear, flavoured oral liquids containing one or more active ingredients dissolved in a vehicle that usually contains a high proportion of sucrose or a suitable polyhydric alcohol or alcohols and may also contain ethanol (96 percent) or a dilute ethanol.

### *Linctuses*

Linctuses are viscous oral liquids that may contain one or more active ingredients in solution. The vehicle usually contains a high proportion of sucrose, other sugars or a suitable polyhydric alcohol or alcohols. Linctuses are intended for use in the treatment or relief of cough, and are sipped and swallowed slowly without the addition of water.

## Cutaneous Preparations

### *Collodions*

Collodions are liquids for cutaneous application, usually containing pyroxylin, a partially nitrated cellulose, in a mixture of ether and ethanol. When they are allowed to dry, a flexible film is formed at the site of application. Salicylic acid collodion USP is a example of collodion, which is used as a keratolytic agent, to treat corns and warts. Collodion is made flexible by adding castor oil and camphor.

Collodions should be stored remote from fire. The label states (1) that the collodion is intended for external use only, (2) the date after which the collodion is not intended to be used, (3) the conditions under which the collodion should be stored, (4) the directions for using the collodion, (5) any special precautions associated with the use of the collodion.

### *Liniments*

Liniments are liquids for cutaneous application that are intended to be applied to the unbroken skin with friction. They are alcoholic or oil based solutions or emulsions containing topically active therapeutic agent. Alcoholic liniments are used generally for their counterirritant, mildly astringent, rubefacient, and penetrating effects. The oily liniments are milder in there action as they penetrate lesser, but are more useful when massage is required. White liniment BP and methyl salicylate BP are examples of liniment.

Certain plastic containers, such as those made from polystyrene, are unsuitable for liniments. The label states (1) the names and concentrations of the active ingredients, (2) that the liniment is intended for external use only, (3) if appropriate, that the contents of the container should be shaken before use, (4) the date after which the liniment is not intended to be used, (5) the conditions under which the Liniment should be stored, (6) the directions for using the liniment and (7) any special precautions associated with the use of the liniment.

### *Lotions*

Lotions are liquids for cutaneous application that are intended to be applied to the unbroken skin without friction. They are aqueous, alcohol or polyols

based solution or suspensions containing topically active therapeutic agent. Calamine lotion USP and Amphotericin B lotion USP are examples of lotions. The label states (1) the names and concentrations of the active ingredients, (2) that the lotion is intended for external use only, (3) that the Lotion should be shaken before use, (4) the date after which the lotion is not intended to be used, (5) the conditions under which the lotion should be stored, (6) the directions for using the lotion and (7) any special precautions associated with the use of the lotion.

## *Douches*

Douches are aqueous solutions, directed against a part or into a cavity of the body to give cleansing or antiseptic action. They are often dispensed in the form of powder or a tablet, with directions for dissolving in a specified quantity of warm purified water. Douches are not official as a class of preparation in any pharmacopoeia but several substances are used frequently in this form. *Vaginal douches*, containing benzalkonium chloride, the parabens or chlorothymol, are most common type of douches used to clean the vagina and hygienic purposes.

## Mucosal Preparations

### *Gargles*

Gargles are aqueous solutions, containing antiseptics, antibiotics and/or anesthetics, intended for gargling to obtain a local effect at pharynx and nasopharynx. They are not to be swallowed. They are supplied as ready-to-use solutions or concentrated solutions to be diluted. They may also be prepared from powders or tablets to be dissolved in water before use. Gargles may contain excipients to adjust the pH which, as far as possible, is neutral. Cepacaine solution, a local anesthetic is used as gargle to get rapid relief from pharyngeal and oral pain.

### *Mouthwashes*

Mouthwashes are aqueous solutions intended for use in contact with the mucous membrane of the oral cavity. They are not to be swallowed. They can be used for two purposes-therapeutic and cosmetic. Therapeutic washes, containing antihistamine, nystatin, hydrocortisone or tetracycline, can be formulated to reduce plaque, stomatitis, dental caries and gingivitis. Cosmetic washes, containing antimicrobial and/or flavouring agent, can be formulated to reduce bad breath. They are supplied as ready-to-use solutions or concentrated solutions to be diluted. They may also be prepared from powders or tablets to be dissolved in water before use. Mouthwashes may contain excipients to adjust the pH which, as far as possible, is neutral. Cepacol, Listerine and Micrin are commercially available mouthwashes.

---

* Apparent solubility can be influenced by the size and shape of solute particles when the particles are in the micron size range. The observed solubility increased with decreasing particle size in accordance with the equation:

$$\log \frac{S}{S_o} = \frac{2\gamma v}{2.303RTr}$$

where, $S$ is the observed solubility. $S_o$ is the inherent equilibrium solubility, $\gamma$ is the surface tension of the particles, $\upsilon$ is the molar volume, $R$ is the gas constant ($8.314 \times 10^7$ ergs/deg mole), $T$ is the temperature absolute, and $r$ is the radius of the particles.

# 18:
# Biphasic Liquids

Biphasic liquids such as suspensions and emulsions are unique dosage forms because many of their properties are due to the presence of a boundary region between two phases. In suspensions, a liquid and an insoluble solid meet to form an interface. In the case of emulsions, two immiscible liquids, usually oil and water, form an interface. An interface between the liquid and air is also present. Although the terms *interface* and *surface* are often used interchangeably, the latter term usually indicates boundaries in which one phase is a gas. The most important and fundamental property of any interface is that it possesses a positive free energy. Essentially, this means that the molecules at the interface are in a higher energy state than if they were located in the bulk phase (Fig. 18.1). The greater the preference of the molecule of interest for the bulk, as compared to the interface, the higher the interfacial free energy.

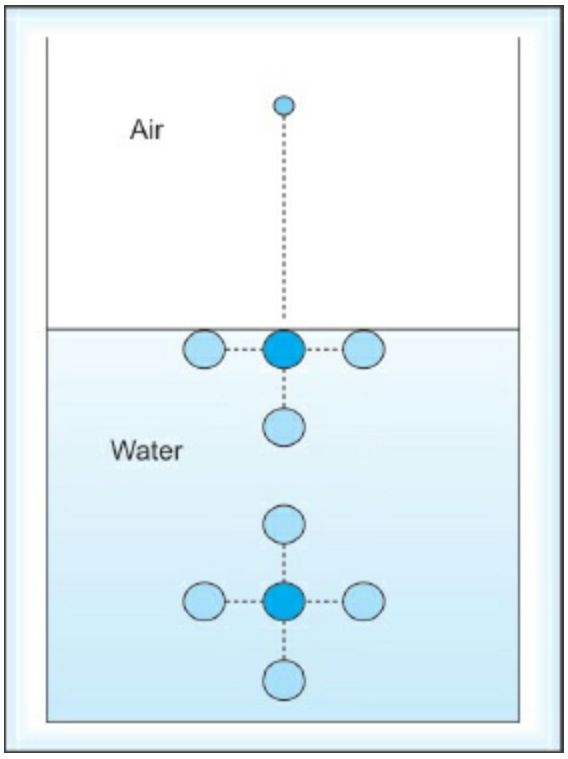


**Fig. 18.1:** Diagrammatic representation of the positive free energy of interface

The primary objective of the formulator is to reduce this positive interfacial free energy value to zero by various means. One approach is simply to reduce the amount of interface i.e. via flocculation or aggregation of particles (in case of suspensions) or via coalescence of the globules to form one macrophase (in case of emulsions). Another method to reduce the interfacial free energy is to vary the composition of the interface to make it rich in surface active material.

Figure 18.1 illustrates the types of boundary regions that are discussed in this chapter. The boundary regions are often complex. *Surface active agents*, which are molecules with special properties, may be contained within a system in various forms: they may be present as single molecules in solution (Fig. 18.2A); they may also be adsorbed at the air-liquid surface (Fig. 18.2B); they may form a layer at the oil-water interface (Fig. 18.2C); or they may form oriented clusters in the aqueous phase, which are called micelles (Fig. 18.2D). Attractive and repulsive forces exist between particles, and the outlined region surrounding the particles in (Fig. 18.2E) indicates a region of potential interaction.

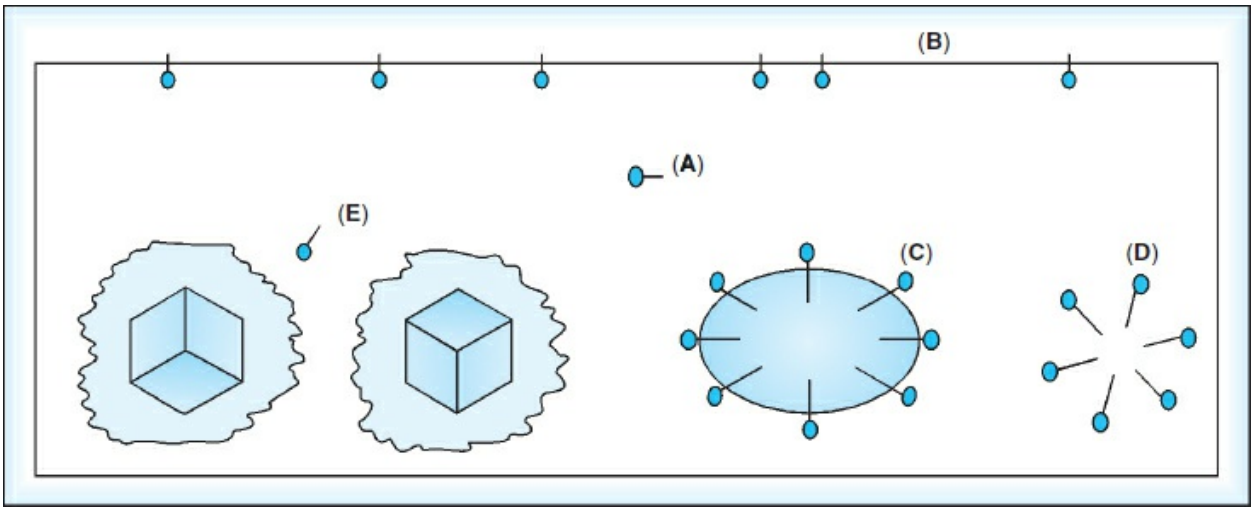



**Fig. 18.2:** Schematic representation of boundary regions encountered in suspensions and emulsions. Key: (A) Surface active agent molecules; (B) Surface active agent oriented at the air-water interface; (C) Surface active agent oriented at the oil-water interface; (D) Micelle; (E) Suspended particles

surrounded by a region of potential interaction

When a beaker containing 50 ml of oil layered on 50 ml of water is examined visually, the interface appears as a sharp discontinuity between the two phases, as shown in (Fig. 18.3A). An interface is actually a region of finite dimension that has composition and properties different from either two phases (Fig. 18.3B) more correctly describes an interface as a region that is a few molecules thick in which a gradation of composition and properties exist. The density does not jump abruptly from 1.0 to 0.9 in moving from the water phase to the oil phase, but rather a gradual transition occurs. The terms *interfacial region* and *interphase* are often used to describe the region labeled 'd' in Fig. 18.3B. Although the physical properties of interfacial regions vary smoothly upon going from one phase to the other, the notion of a mathematical surface that has no thickness, as in Fig. 18.3A, is still useful for modeling interfacial regions and has been used to describe interfacial phenomena.

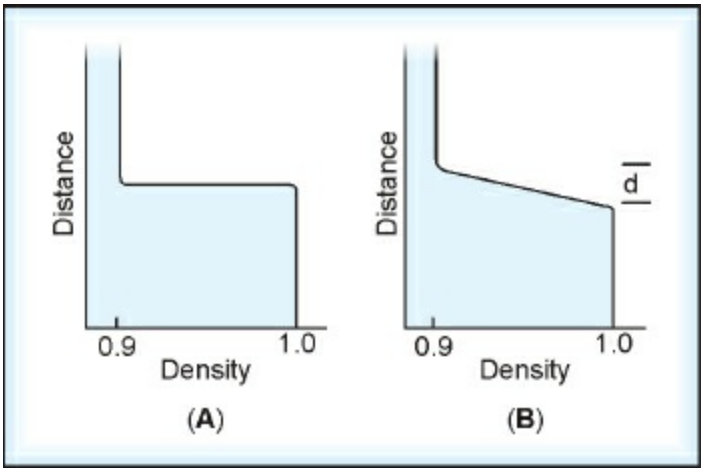



**Fig. 18.3:** Change in density at the oil (density 0.9)/water (density 1.0) interface. Key: (A) Mathematical surface; (B) Interfacial region, 'd'

In addition, the molecules in the interfacial region are not locked into position but are in constant motion. The average interfacial residence time for a molecule of a liquid is believed to be approximately $10^{-6}$ sec. Thus, the interfacial region of suspensions and emulsions is a dynamic, clearly identifiable region between the phases of the system.

When the interfacial region constitutes a large portion of the system—as

when the particle size of the solid phase of a suspension is small or when the globule size of the dispersed phase of an emulsion is small—the overall properties of the system are profoundly influenced by the presence of the interfacial region.

A fundamental thermodynamic equation that describes suspensions and emulsions is as follows:

$$\Delta G = \gamma \Delta A \ldots (1)$$

where, $\Delta G$ is the change in the free energy of the system accompanying a change in interfacial area $\Delta A$, $\gamma$ is the interfacial tension (liquid-liquid for an emulsion or solid-liquid for a suspension), and temperature, pressure, and composition are constant. The term $\Delta G$ represents the work required to increase the area of the interface by an amount equal to A. Since this work is always positive, a system always tends toward that state having the lowest possible interfacial area. This state is thermodynamically stable.

Thus, for suspension, the thermodynamically stable state is a single large particle, whereas, a layer of oil on water, is the thermodynamically stable state for an emulsion. Although all systems tend towards the thermodynamically stable state of minimal interfacial area (which results in dramatic changes in properties), systems may vary considerably in their rates of conversion. If a system undergoes only minor changes during the period of interest, e.g. shelf-life., such a system is viewed as kinetically stable even though it may be unstable with respect to some longer time period. As a consequence, the industrial pharmacist faces the challenging task of preparing a kinetically stable dosage form, i.e. a dosage form whose properties remain satisfactory for an acceptable shelf-life, even though suspensions and emulsions are thermodynamically unstable.

The primary objective of this chapter is to put forth some of the basic theoretic and practical considerations that apply to suspension and emulsion systems, and to relate these principles to formulation methods, evaluation procedures, and manufacturing techniques.

# SUSPENSIONS (SOLID-LIQUID SYSTEM)

Suspensions are heterogeneous systems consisting of two phases. The *continuous* or *external phase* is generally a liquid or semisolid, and the *dispersed or internal* phase is made up of particulate matter that is essentially insoluble in, but dispersed throughout, the continuous phase; the insoluble matter may be intended for physiologic absorption or for internal or external coating functions.

Suspensions can be classified in various ways based on their physical state as suspension, aerosol and foam; based on the size of the dispersed particles as molecular dispersion (size <1.0 nm), colloidal dispersion (0.1 – 0.2 μm) and coarse dispersion (size >0.2 μm); based on behaviour of dispersed phase which may consist of discrete particles (deflocculated suspensions), or it may be a network of particles (flocculated suspensions), resulting from particle-particle interactions.

Suspensions contribute to pharmacy and medicine by supplying insoluble and often distasteful drugs in a form that is pleasant to taste, by providing a suitable form for the application of dermatologic materials to the skin or mucous membranes, and for the parenteral administration of insoluble drugs. Therefore, pharmaceutical suspensions may be classified into three general classes: Oral, topical and parenteral suspensions.

## Oral Suspensions

The solids content of an oral suspension may vary considerably. For example, antibiotic preparations, antacids and radiopaque suspensions contain relatively high amounts of suspended material for oral administration. The vehicle may be a syrup, a sorbitol solution, or a gum-thickened, water-containing artificial sweetener because in addition to ingredients, safety, taste, and mouthfeel are important formulation considerations. In the case of limited shelf life (low chemical stability of the insoluble drug), the dosage form may be prepared as a dry granulation or powder mixture that is reconstituted with water prior to use.

## Topical Suspensions

Historically, the externally applied “shake lotion” is the oldest example of a pharmaceutical suspension. Calamine Lotion USP, as well as other dermatological preparations, is closely associated with the technical development of the pharmaceutical suspension. Because safety and toxicity considerations are most readily dealt with in terms of dermatological acceptability, many useful suspending agents were first introduced in topical formulations. In addition, the protective action and cosmetic properties of topical lotions usually require the use of high concentrations of the dispersed phase, often in excess of 20% (w/v). Therefore, topical lotions represent the best example of suspensions that exhibit low settling rates. Various pharmaceutical vehicles have been used in the preparation of topical lotions, including diluted oil-in-water or water-in-oil emulsion bases, determatological pastes, magmas, and clay suspensions.

## Parenteral Suspensions

The solids content of parenteral suspensions is usually between 0.5 and 5.0% (w/v), except for insoluble forms of penicillin, in which concentrations of the antibiotic may exceed 30% (w/v). These sterile preparations are designed for intramuscular, intradermal, intra-lesional, intra-articular, or subcutaneous administration. The viscosity of a parenteral suspension should be low enough to facilitate injection. Common vehicles for parenteral suspensions include preserved 0.9% (w/v) saline solution or a parenterally-acceptable vegetable oil. The primary factor governing the selection of injectable ingredients is safety. Ophthalmic suspensions that are instilled into the eye must be prepared in a sterile manner. The vehicle employed is essentially isotonic and aqueous in composition.

Traditionally, certain kinds of pharmaceutical suspensions have been given separate designations, such as mucilages, magmas, gels, and sometimes aerosols; also included would be dry powders to which a vehicle is added at the time of dispensing.

Suspensions form an important class of pharmaceutical dosage forms. These disperse systems present many formulations, stability, manufacturing, and packaging challenges. Almost all suspension systems separate on standing. The formulator’s main concern, therefore, is not necessarily to try to eliminate separation, but rather to decrease the rate of settling and to permit easy resuspendability of any settled particulate matter. A satisfactory suspension must remain sufficiently homogeneous for at least the period of time necessary to remove and administer the required dose after shaking its container.

For the most part, only aqueous suspensions are discussed, and little attention is paid to oils or aerosol propellants as suspension vehicles. Also, this discussion is limited to suspensions with particles having diameters greater than 0.2 μm, approximately the lower limit of resolution of optical microscopes. Thus, although suspended particles do not exhibit all the properties of colloids, such as the so called colligative properties, they do have heightened surface properties, that is, whatever surface properties exist are magnified because of the increased surface area.

## THEORETIC CONSIDERATIONS

A knowledge of the theoretic considerations pertaining to suspension technology should ultimately help the formulator to select the ingredients that are most appropriate for the suspension and to use the available mixing and milling equipment to the best advantage. Some understanding of wetting, particle interaction, electrokinetics, aggregation, and sedimentation concepts facilitates the making of good formulatory decisions.

## Wetting

Certain solids are readily wetted by liquid media whereas others are not. In aqueous suspension terminology, solids are said to be either *hydrophilic* (lyophilic or solvent-loving, rarely lyotropic) or *hydrophobic* (lyophobic). Hydrophilic substances are easily wetted by water or other polar, liquids; they may also greatly increase the viscosity of water suspensions. Hydrophobic substances repel water but can usually be wetted by nonpolar liquids; they usually do not alter the viscosity of aqueous dispersions.

A frequently encountered difficulty that is a factor of prime importance in suspension formulation concerns the wetting of the solid phase by the suspension medium. Hydrophilic solids usually can be incorporated into suspensions without the use of a wetting agent, but hydrophobic materials are extremely difficult to disperse and frequently float on the surface of the fluid owing to poor wetting of the particles or presence of tiny air pockets.

By definition, a suspension is essentially an incompatible system, but to exist at all, it requires some degree of compatibility, and good wetting of the suspended material is important in achieving this end.

When a strong affinity exists between a liquid and a solid, the liquid easily forms a film over the surface of the solid. When this affinity is nonexistent or weak, however, the liquid has difficulty displacing the air or other substances surrounding the solid, and there exists an angle of contact between the liquid and the solid.

This *contact angle*, $\theta$, results from an equilibrium involving three interfacial tensions, specifically, those acting at the interfaces between the liquid and vapour phases, at the solid and liquid phases, and at the solid and vapour phases. These tensions are caused by unbalanced intermolecular forces in the various phases.

As can be seen in Fig. 18.4, the contact angle varies from 0 to 180° and is a useful indication of wetting. A low contact angle indicates that adhesive forces between the liquid and the solid predominate and wetting occurs, while a high contact angle indicates that the cohesive forces of the liquid predominate.

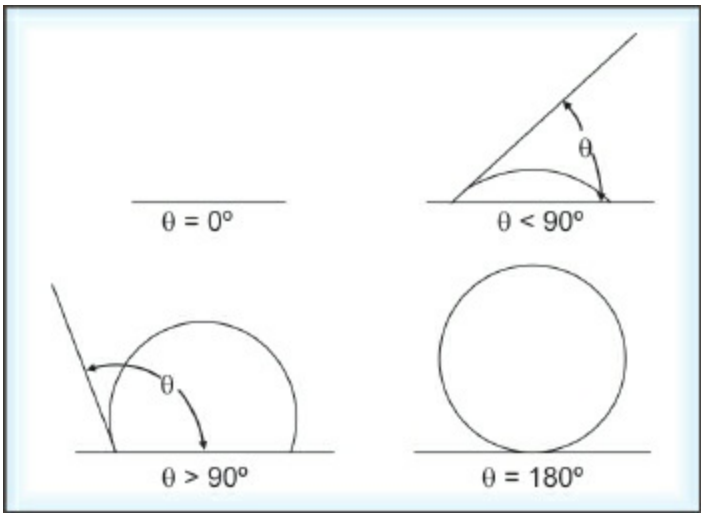



**Fig. 18.4:** The use of the contact angle, $\theta$, to characterize the wetting of a solid by a liquid

The basic equation that applies to wetting is the Young equation, which is based on the change in free energy caused by an increase in the area of a solid that is wetted by a liquid. For a small, reversible change in the position of the liquid on the surface, there is an increase in the liquid-solid interfacial area, $\Delta A$, and a corresponding decrease in the solid-air interface of $\Delta A \cos \theta$. The corresponding free energy change is given by:

$$\Delta G = \gamma_{S/L}\, \Delta A - \gamma_{L/A}\, (\cos\theta)\, \Delta A \quad \ldots (2)$$

As $\Delta A$ approaches zero, the ratio $\Delta G/\Delta A$ approaches zero at equilibrium, so that Eq. (2) reduces to the Young equation:

$$\theta_{S/A} = \theta_{S/L} + \theta_{L/A} \cos\theta \quad \ldots (3)$$

The Young equation states that the contact angle will be <90° if the interaction between the solid and liquid is greater than the interaction between the solid and air, i.e. $\theta_{S/L} > \theta_{S/A}$. Under these conditions, wetting occurs. A general guideline is that solids are readily wetted if their contact angle with the liquid phase is less than 90°. Table 18.1 confirms this rule of thumb, as solids that are known to be easily wetted, such as potassium chloride, sodium chloride, and lactose, have the lowest contact angles. Other materials that are known to be difficult to wet, such as high-density polyethylene and magnesium stearate, have contact angles greater than 90°. Interestingly, the contact angle of chloramphenicol increases from 59 to 125°, indicating a change to a nonwetting surface when the palmitate ester is

formed.

**Table 18.1:** Contact angle of solids of pharmaceutical interest against a saturated aqueous solution of the material

| Material | Contact angle (°) |
|---|---|
| Potassium chloride | 21 |
| Sodium chloride | 28 |
| Lactose | 30 |
| Caffeine | 43 |
| Acetaminophen | 59 |
| Chloramphenical | 59 |
| Phenobarbital | 70 |
| Sulfadiazine | 71 |
| Aspirin | 75 |
| Phenacetin | 78 |
| Hexobarbitol | 88 |
| Polyethylene (high-density) | 100 |
| Salicylic acid | 103 |
| Magnesium stearate | 121 |
| Chloramphenicol palmitate | 125 |

The wetting of a solid during the manufacture of a suspension involves the displacement of air from the solid surface. The first step in the wetting of a powder is *adhesional wetting*, in which the surface of the solid is brought into contact with the liquid surface. This step is the equivalent to going from stage *A* to stage *B* in Fig. 18.5. The particle is then forced below the surface of the liquid as *immersional wetting* occurs (B to *C* in Fig. 18.5). During this step, solid-liquid interface is formed, and solid-air interface is lost. Finally, the liquid spreads over the entire surface of the solid as *spreading wetting* occurs. The work of spreading wetting is equal to the work to form new solid-liquid and liquid-air interfaces minus the loss of the solid-air interface. For the total wetting process, the work required is the sum of the three types

of wetting:

$$W_{total} = -6\gamma L_{/A} \text{ COS}\theta \text{ ... (4)}$$

Therefore, the analysis supports the generalization drawn from the Young equation, as wetting occurs spontaneously, i.e. without any input of work into the system, if the contact angle is <90°.

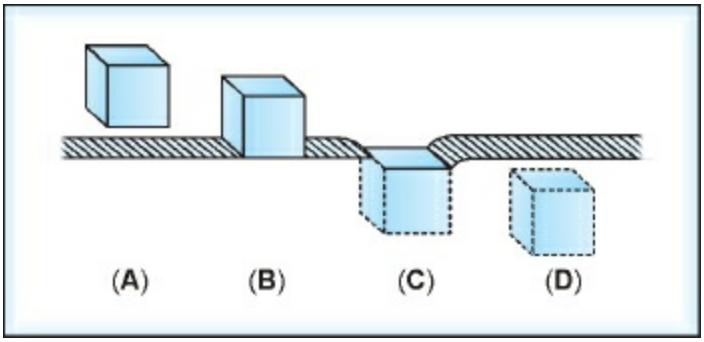



**Fig. 18.5:** The three stages involved in the wetting of a solid: (A → B) Adhesional wetting; (B → C) Immersional wetting; (C → D) Spreading wetting

## *Electric Double Layer*

Formulation of a suspension or emulsion necessitates the use of ionic surface-active agents as the suspending agent or the emulsifying agent. The surface active agents are oriented at the oil-water interface, so that the charged groups form the outside surface. The presence of charge at an interface has profound effects on the nature of the interfacial region (Fig. 18.6) presents a model of the interfacial region of a charged surface and provides a good basis for understanding the behaviour of pharmaceutical suspensions and emulsions.

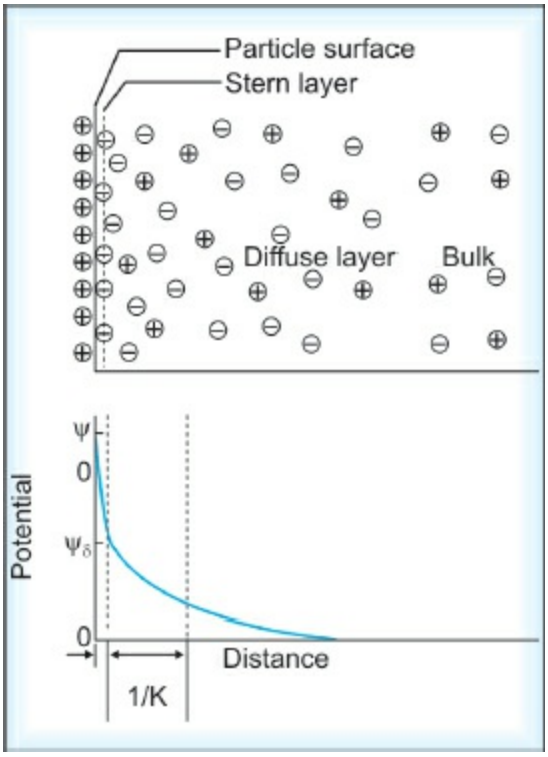



**Fig. 18.6:** Diffuse double-layer model of a positively charged surface in an aqueous medium

A layer of ions of opposite charge is sufficiently held together by the charged surface so that the ions move with the surface **(Stern layer)**. The surface charge is not completely balanced by the Stern layer, and a second region, the ***diffuse layer***, is necessary for complete balance of the surface charge. The diffuse layer contains both anions and cations, but ions that are

of opposite charge to the surface predominate, so that in the neighboring area of any charged particle, there exists an ion atmosphere having a net charge that is opposite to that of the surface. The concept of a diffuse double layer was developed by Gouy and Chapman to describe the transition from points near the charged surface to the electrically neutral bulk solution.

This mobile layer of ions has a definite thickness, which is approximated by the so-called Debye length (1/*k*):

$$1/k = \left(\frac{DkT}{2ne^2Z^2}\right)^{1/2} \quad \ldots (5)$$

where, *D* is the dielectric constant of the medium, *n* is the concentration of ions in the bulk solution phase, e is the electronic charge, *Z* is valence, *k* is Boltzman's constant, and *T* is temperature. Beyond this distance (*1/k*), the net charge density of the ion atmosphere approaches zero, and the electrical potential is reduced considerably below its value at the surface. The major factors affecting the thickness of the double layer are the electrolyte concentration of the solution, *n*, and the valence, z, of counter ions.

The terms lyophobic (hydrophilic) and lyophilic (hydrophobic) were mentioned in the previous section. These terms are sometimes considered synonymous with nonwetting and wetting, respectively. The primary behavioural difference between these two classes of materials is their sensitivity to the presence of electrolytes. Lyophobic materials in suspension are sensitive to the addition of salts, whereas lyophilic materials are not. Addition of large amounts of electrolyte to a solution of a lyophilic material results in the precipitation of lyophilic material, although dilution with the vehicle reverses the precipitation. As distinct from lyophilic materials, dilution with the vehicle does not reverses the precipitation of lyophobic solids. However, in the case of the lyophobic material, one may not observe aggregation, the desired form of matrix formation during sedimentation. The stability of lyophobic colloids is also reduced by lowering the repulsive potential of the electrochemical double layer or by decreasing the degree of hydration.

In addition to the repulsion between particles resulting from the diffuse double layer, the ionic strength and the valence and size of the ions on the surface and in the double layer influence both the total charge and the thickness of the double layer. These factors also influence hydration. Some of these overall interrelationships are illustrated in Fig. 18.7. Note that 

increasing the concentration of ions in the solution decreases the thickness of the diffuse double layer by "swamping," and therefore aggregation is encouraged. Specific adsorption of an ion by the system also neutralizes the surface charge of the particle and allows aggregation. The concentration of electrolyte needed to effect optimal aggregation depends on the balance and type of interacting ion; addition of electrolyte past this point may result in a reversal of charge, which in turn would cause deaggregation and ultimate caking of the system. These ion effects, can be systematized by referring to the *Schulze-Hardy* rule.

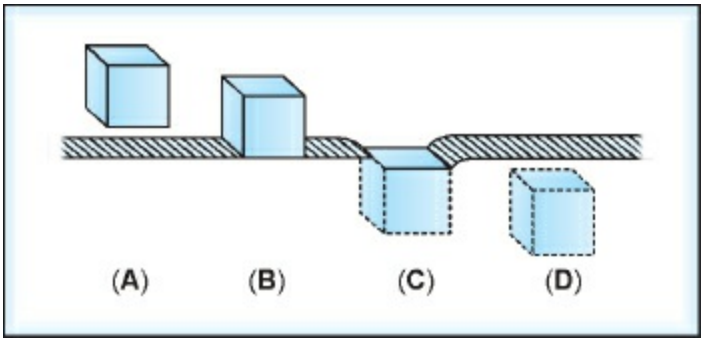



**Fig. 18.7:** Stability of colloidal particles in aqueous suspension depends on hydration and electrostatic charge; these depend on the chemical composition and structure of the substrate at the liquid-solid interface

The *Schulze-Hardy* rule states that the valence of the ions having a charge opposite to that of the hydrophobic particle appears to determine the effectiveness of the electrolyte in aggregating the particles. The aggregating value or efficiency, therefore increases with the valence of the ions. Divalent ions are ten times as effective as monovalent and trivalent are one thousand times as effective as monovalent. It is important to remember that this rule is valid only for systems in which there is no chemical interaction between the aggregating electrolyte and the ions of the double layer of the particle surface. Note also, incidentally, that aggregating forces are of sufficient magnitude to overwhelm the electrostatic repulsion between particles having net charges of the same sign. With respect to actual electrolyte concentrations used, satisfactory aggregation has been found to occur at the following approximate ion concentrations: 25 to 150 mmol/L for monovalent ions, 0.5 to 2.0 mmol/L for divalent ions, and 0.01 to 0.1 mmol/L for trivalent ions. The influence of valence and concentration on the aggregation of a lyophobic particle suspension can be determined experimentally by either measuring the

change in zeta potential or by observing the degree of aggregation in terms of some measurable parameter such as sediment height.

Although it is not pharmaceutically useful to a great extent, the *Hofmeister* or *lyotropic series rule* applies to hydrophilic particles in a manner somewhat analogous to the Schulze-Hardy rule, and takes into account not only the charge but also the ionic size and hydration capability. In order of decreasing aggregating ability, the monovalent cation and anion progressions are respectively $Cs^+$, $Rb^+$, $NH_4^+$, $K^+$, $Na^+$, $Li^+$ and $F^+$, $IO_3^-$, $Cl^-$, $ClO_3$, $Br^-$, $NO_3^-$, $ClO_4^-$, $I^-$, $CN^-$.

## *Interparticle Forces*

Many observed properties of a disperse system reflect the net force of interaction between the particles or globules that the system comprises. The following forces have been identified:

1. *Electrostatic repulsive forces* arise from overlapping of the diffuse double layers of approaching surfaces. These forces depend greatly on the concentration and valence of electrolyte in solution.
2. *Van der Waals attractive forces* arise from the electromagnetic fluctuations in the molecules that make up the surface. These forces are largely independent of the electrolyte.
3. *Repulsive hydration forces* arise from the structuring of water in the interfacial region. These forces are independent of electrolyte concentration. At low electrolyte concentration, the contribution of this force may not be observed, owing to the strong electrostatic repulsive forces. At high electrolyte concentration, however, the diffuse double-layer interactions are weak, and the repulsive hydration forces may determine the interaction of the surfaces.
4. *Born repulsive forces* are of short range and operate over distances of atomic dimensions. They are due to the repulsive effects of atomic orbital overlap.
5. *Adhesive forces* arise when surfaces are in contact. The adhesive forces depend on pH, specific cations, and the crystallographic orientation of the surfaces.
6. *Steric repulsive forces* depend on the size, geometry, and conformation of

molecules that are adsorbed on the surface.

Deryaguin, Landau, Verwey and Overbeek recognized the concept of balance between electrostatic repulsive and van der Waals attractive forces between particles.

Thus, the concept has become known as the *DLVO theory* (Fig. 18.8) which illustrates the repulsive double layer and attractive van der Waals forces at three electrolyte concentrations.

At low electrolyte concentration—the repulsive forces predominate so that the particles remain independent, and the system is considered dispersed.

At high electrolyte concentration—the repulsive forces are greatly reduced, so that the attractive van der Waals forces predominate. These net attractive forces that the particles encounter cause the formation of an aggregate of particles, a process known as *coagulation*. Thus, the DLVO theory explains the fact that the addition of electrolyte to a colloidal system causes coagulation.

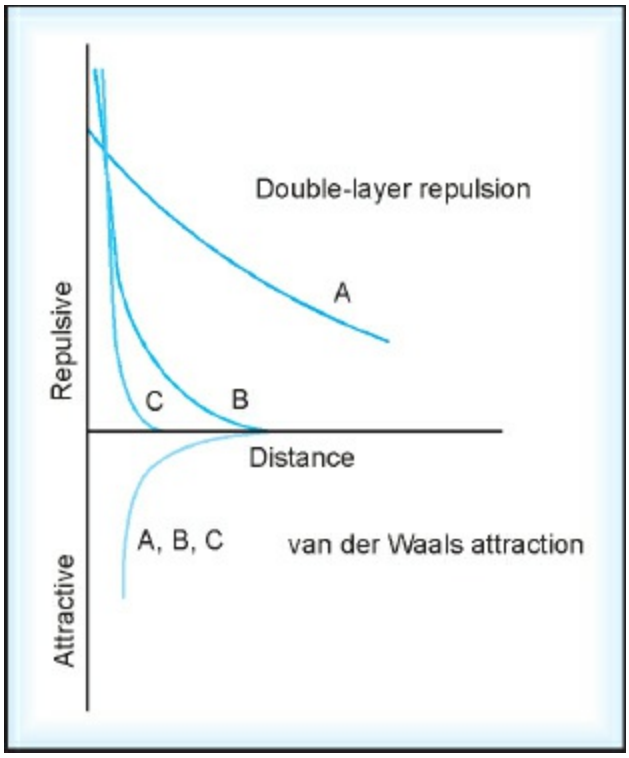



**Fig. 18.8:** Effect of electrolyte concentration on repulsive double-layer forces and attractive van der Waals forces. Key: A. Low electrolyte concentration, B. Intermediate electrolyte concentration, C. High electrolyte concentration

When the repulsive hydration and Born repulsive forces are considered along with the double-layer repulsive and van der Waals attractive forces, the net force diagram shown in Fig. 18.9 is obtained.

As two particles approach each other in an aqueous medium of proper electrolyte concentration, a weak attractive force exists just beyond the range of the double-layer repulsive forces. This attractive region is called the *secondary minimum* and is responsible for the particle interaction termed *flocculation*. Particles therefore experience attraction at significant interparticle distances (10 to 20 nm) and form the fibrous, fluffy, open network of aggregated particles known as floccules, as illustrated by Fig. 18.10. Such a suspension is called *flocculated suspension*. In this suspension type, the structure of aggregates is quite rigid, hence, they settle quickly to form a high sediment height and are easily redispersible because the particles constituting individual aggregates are sufficiently far apart from one another to preclude caking.

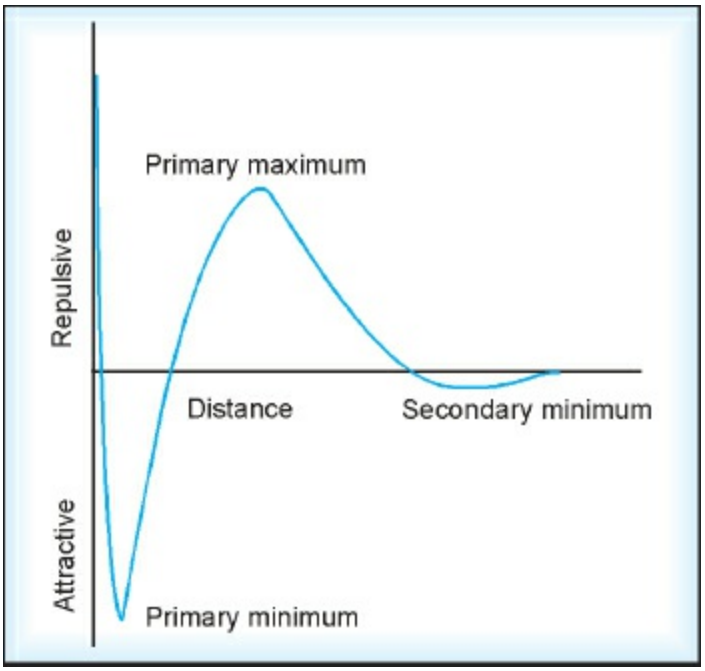



**Fig. 18.9:** Net-potential energy curve for a particle in an electrolyte vehicle

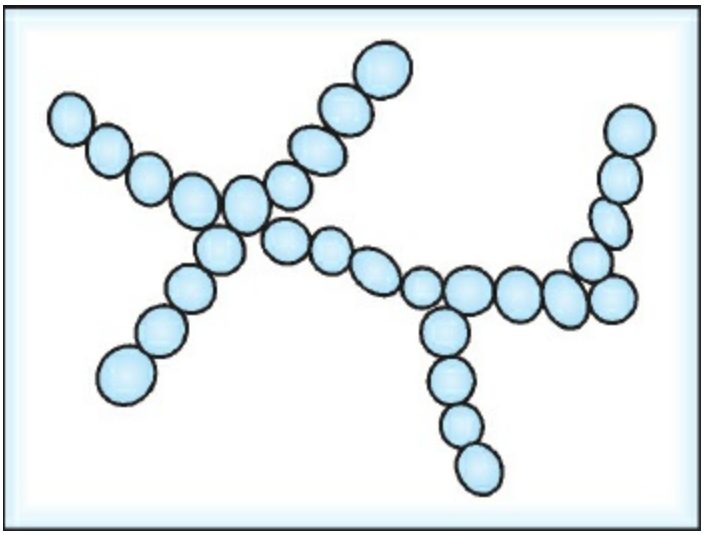


**Fig. 18.10:** An artist's conception of open-network suspension aggregate

The secondary minimum is not observed if the repulsive forces extend further from the surface than attractive forces. Thus, the adjustment of valence and concentration of the electrolyte can induce flocculation.

A repulsive barrier termed the *primary maximum* separates the secondary minimum from the primary minimum. The magnitude of the repulsive forces at the primary maximum determines whether a flocculated system will remain flocculated. If the thermal energy in the system is similar to, or greater than, the repulsive barrier, then the particles in the system are able to move closer together (0.5 to 2.0 nm) and encounter strong attraction due to the *primary minimum*. The strong attraction in the primary minimum gives rise to the particle interaction termed *coagulation*. Other sources of energy, such as centrifugation or compression of the particles due to freezing may force particles into the primary minimum by overcoming the primary maximum and may lead to coagulation.

Closed aggregate, or *coagule* is characterized by a tight packing produced by surface film bonding as shown in Fig. 18.11. These aggregates settle slowly to low sediment heights that approach the sediment density of a dispersed particulate system. Characteristically, sediments composed of closed aggregates are not easily redispersed. The affinity of surface films to each other is responsible for the tenacity of the aggregate not only within an individual aggregate, but also to surrounding aggregates. Upon sedimentation, the aggregates tend to form a single large "film-bonded" aggregate, which is difficult, if not impossible, to redisperse. The surface

films that lead to coagule formation are often surfactants, gases, immiscible liquids, and in the case of non-aqueous suspensions, water.

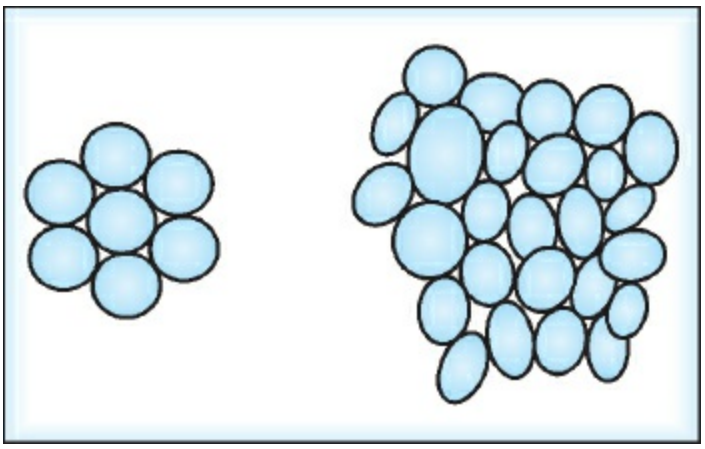


**Fig. 18.11:** An artist's conception of closed suspension aggregate

At low electrolyte concentration, the interaction curve contains a much larger primary maximum than at higher ionic strength conditions, and particle interactions are minimized. Such a suspension is called *dispersed, deflocculated or peptized*. In this suspension type, the individual particles are dispersed as discrete entities as illustrated in Fig. 18.12. This suspension type sediments slowly (as compared with the closed and open aggregate types), attains the lowest possible sediment height, and owing to the closeness of the particle surfaces upon sedimentation, possesses a high potential for *caking*, because of the ease of formation of extensive crystal bridging.

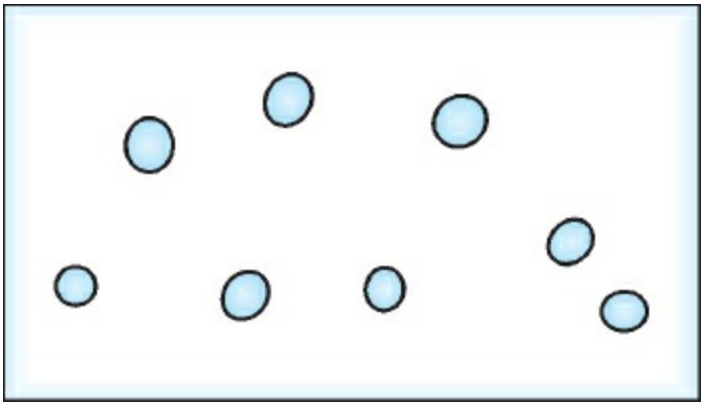


**Fig. 18.12:** An artist's conception of dispersed suspension form

Caking via crystal bridging can be minimized by utilizing the open

network aggregate (floccule) suspension type, as the particles cannot sediment to a close proximity because of the rigidity of the aggregate. From the practical point of view, since fully aggregated suspensions are often unsightly, partial aggregation is often a desired objective, as it leads to resistance to caking and imparts aesthatic qualities to a suspension formulation. If either the concentration or the valence of the background salt is increased, the primary maximum is reduced, so that such particle interactions as flocculation or coagulation occur. Obviously, a pharmaceutical suspension must be redispersible on only mild agitation to ensure dosage uniformity.

## Sedimentation Rates

With regard to actual settling rates, the well-known Stokes relation describes the sedimentation velocity of a particle in suspension:

$$v = \frac{2r^2(d_1 - d_2)g}{9\eta} = \frac{D^2(d_1 - d_2)g}{18\eta} \quad \ldots (6)$$

where, $\upsilon$ = velocity of sedimentation in cm/s; $r$ = particle radius and $D$ = particle diameter in cm; $d_1$ and $d_2$ = density of the particle and the liquid, respectively, in g/ml; $g$ = gravitational constant = 980.7 cm $s^{-2}$; and $\eta$= viscosity of the medium in poise, i.e. g $cm^{-1}$ $s^{-1}$ in cgs units. Note, incidentally, that water at 20° has a viscosity of approximately one centipoise (0.01 poise). The student should be aware of the fact that assumptions were needed to facilitate the derivation of this relationship. Pharmaceutical systems containing less than 2 g of solids per 100 ml generally follow the Stokes equation. Obviously, if the two densities involved are such that a negative velocity results, this is the rate of floatation or creaming.

To illustrate by example a Stokes law calculation, one could find the settling rate of a particle having a radius of 4 qm and a specific gravity of 4 in a medium with a viscosity of 100 cps and a specific gravity of 1.2. Using the Stokes equation (after changing the radius and viscosity to the proper units), a velocity of approximately $1 \times 10^{-4}$ cm/s is found. This means that this suspension packaged in a 10 cm high bottle would settle in $10^5$ s or in a little over one day. Similar “ideal calculations” make it clear that pharmaceutical suspensions are destined to settle, even though one can slow the process, well within the shelf-life times of pharmaceutical products. As the equation indicates, the parameter most powerful in changing the velocity of settling is the particle diameter or radius, as it is a squared term; as technologists, the formulators are most able to control this and the viscosity of the medium.

To handle more concentrated pharmaceutical suspensions (i.e. those containing more than 10 g of solid per 100 ml) requires the development of equations that go beyond the Stokes derivation. Obviously, the mathematics becomes even more complex if one tries to handle heterodispersed particles (i.e. not all particles are the same size), and in addition, irregularly shaped particles. One can, however, bring basic physical and chemical thinking to the kinetic settling and creaming problems. Although exact and completely

useful equations may not be developed, some conceptual value is gained.

T. Higuchi among others, has approached this problem of hindered settling of relatively concentrated suspensions and has obtained an equation with somewhat fewer limitations. By recasting the problem, he considered the settling phenomenon to be equivalent to the movement of an external liquid phase through a bed of internal phase. Fluid flow through packed beds had been previously treated mathematically by Kozeny. Starting with the Kozeny equation, Higuchi developed an equation that somewhat resembles the Stokes equation with respect to the parameters needed to solve it. Specifically, one needs all the factors in Stokes' Law except the particle radius, and in addition, one needs an empiric constant, the volume fraction of the phases (which is a measure of the porosity of the "bed"), and the specific surface area of the particles, i.e. the area in $cm^2/g$.

Such an equation does not take into account the size or size distribution of the particles (an advantage over Stokes), and is of some validity where it can be used, i.e. in more concentrated systems in which the suspended phase forms a sort of bed or where the aggregated structure is relatively firm. The equation has not been subjected to a great deal of experimental verification, but sample calculations show that the Stokes relationship (which does not take into account explicitly the concentration of the internal phase) gives too fast a rate of settling as opposed to actual observations made on more concentrated suspensions. The newer equation shows that the rate of settling decreases rapidly with the concentration of solid phase. Pharmaceutically speaking, the extension of this concept corroborates with the observation that pastes do not settle in the usual sense, although small quantities of liquid components may separate. The new equation neglects, as does Stokes law, the contribution made to suspension stability by the various particle interactions. In summary, it is interesting to observe that suspensions are complex systems from a theoretic standpoint.

Regarding sedimentation and factors affecting it, it should be remembered that as the solids content is increased, viscosity is also increased. Hence, study of the viscosity characteristics of the vehicle alone may not always produce valid observations. Also, the discussions of particle charges, aggregation, zeta potential, and sedimentation rates deal with aqueous systems. The vehicles used for pharmaceutical suspensions are usually aqueous, but may be oleaginous. Examples of oleaginous vehicles include

peanut, sesame, cottonseed, corn, safflower, and castor oils and also fluorocarbon aerosol propellants. Particle-particle and particle-liquid interaction vary greatly depending on the vehicle. Since the electrical potential between two charges is inversely proportional to the dielectric constant of the medium between them, the distribution of ions in the double layer also depends on the dielectric constant of the vehicle. In low dielectric liquids, such as oils, the double layer is many times thicker than in aqueous systems with high dielectric constants. It is, therefore, much more difficult to produce an aggregated particle structure in a low dielectric medium.

## Particle Growth

The size distribution of dispersed systems may increase during aging, owing to four principal mechanisms: Ostwald ripening; polymorphic transformation; crystal habit; and temperature cycling.

### *Ostwa/d Ripening*

Since suspensions are saturated solutions of the particulate substance, small changes in temperature that occur during shelf storage lead to unexpectedly rapid caking via crystal bridging, much in the same way that crystal growth yields can be optimized by alternately warming and cooling a mother crystallization liquor. This process, known as *Ostwald ripening*, is unavoidable in pharmaceutical suspensions of the dispersed type. Suspensions of the dispersed type tend to cake easily, owing to the compact sedimentation that occurs when these suspensions settle.

The basis for Ostwald ripening is found in Eq. (7) and applies to the equilibrium solubility of small particles:

$$\ln \frac{S}{S^0} = \frac{2\gamma V}{rRT} \qquad \ldots (7)$$

where, $S^0$ is the solubility of infinitely large particles, $S$ is the solubility of a small particle of radius $r$, $\gamma$ is the surface tension, and $V$ is the molar volume of the solid.

It is important to distinguish between equilibrium solubility and rate at which a substance dissolves. Dissolution rate is affected by particle size since the surface area of the solid available to the solvent increases with decreasing particle size. Equilibrium solubility at a given temperature, however, is affected by particle size only in the particle size range near the colloidal dimension, i.e. less than 5 μm. For example, the equilibrium solubility at 25°C of calcium sulfate with an average particle size of 2 μm is 2.085 g/L. When the average particle size is reduced to 0.3 μm, the equilibrium solubility increases to 2.476 g/L.

In a practical sense, these values mean that a solution that is saturated with respect to small particles is supersaturated with respect to large particles of the same substance. This condition causes crystal growth in a suspension, as solute diffuses from the saturated layer surrounding small particles to the

saturated layer surrounding the larger particles. Precipitation on the surface of the larger particle occurs as the saturated layer becomes supersaturated with respect to the equilibrium solubility of larger particles. The overall effect is an increase in particle size and a decrease in the number of particles in suspension. The ultimate conclusion of this process is the formation of one large particle that represents the thermodynamically stable state of a suspension as described by Eq. (1).

## Polymorphic Transformation

Polymorphism as applied to crystals specifically refers to the different crystal structures the same chemical compound may have. The difference in the equilibrium solubility of polymorphs provides a driving force for crystal growth in suspension as the particles of the more soluble polymorph go into solution and reprecipitate as the less soluble, i.e. more stable, form. This process is accelerated if the drug powder used to prepare the suspension contains a mixture of polymorphs, or if a seed of the more stable form is introduced. The rate of conversion of a metastable to a stable polymorph may be rapid or slow. When this rate of conversion is very slow, it may be feasible to use the metastable form commercially.

## *Crystal Habit*

Crystal habit may be denned as the outward appearance of an agglomeration of crystals. Although seemingly trivial, crystal habit can be of great importance in suspension redispersibility, sedimentation, physical stability, and appearance. For example, sulfisoxazole can be produced in a single geometric crystal form having relatively similar sizes, but an agglomerate of the crystals can have physical properties vastly different from those of single crystals. Small clumps of sulfisoxazole crystals may exhibit little tendency to disperse because of the tenacity of the clump. These clumps may exhibit retarded dissolution and thus retarded bioavailability rates due to the inability of a dissolution fluid to penetrate to the interior crystal components of the clump.

Traditionally, crystal habit was classified on the basis of the geometry of the agglomerate (needle, prism, plate, etc.), but in reality, most crystal habit morphology is of a non-descript form. The relatively strong, rigid crystalline structure that exists within a crystal is not responsible for the agglomeration of crystals. Rather, weak van der Waals interactions occurring at crystal surfaces hold the agglomerate of crystals in form. Mostly, this occurs as non-geometrically classifiable clumps.

The factors controlling crystal characteristics involve basically either the production of a change in crystal habit (physical shape such as needle, plate, prism) or the production of no change in crystal habit. When there is no change in the crystal habit, the following factors may still be considered: drug

decomposition leading to salting in or out, pH changes with changes in the particle size distribution, and the effect of change in temperature. When there is a change in crystal habit, solvation and polymorphism (presence of one or more crystalline and/or amorphous forms) are of importance. It is also notable that the rate of physiologic absorption can be greatly altered, depending on which crystalline or amorphous forms are administered.

## *Temperature Cycling*

Temperature cycling may lead to crystal growth, as solubility depends on temperature. In most cases, solubility is directly related to temperature, so that a slight rise in temperature leads to an increased equilibrium solubility. A drop in temperature, however slight, results in a supersaturated solution surrounding each particle. Precipitation occurs to relieve the supersaturation, and crystal growth occurs. The temperature effects depend on the magnitude of the change in temperature over a given period of time, the time interval, the effect of temperature on the solubility of the suspended drug, and on recrystallization phenomena.

# FORMULATION COMPONENTS

## Wetting Agents

A frequently helpful pharmaceutical technique for modifying the wetting characteristics of powders involves the use of *surfactants* (sometimes with shearing) to decrease the solid-liquid interfacial tension and contact angle between solid particles and liquid vehicle. Common surfactants used as wetting agents include (1) non-ionic type (polyoxyalkyl ethers, polyoxylakyl phenyl ethers, polyoxy hydrogenated castor oil, sorbitan esters and polyoxy sorbitan esters) and (2) anionic type (docusate sodium and sodium lauryl sulfate). The best range for wetting and spreading by non-ionic surfactants is between a hydrophile-lipophile balance (HLB) value of 7 and 10, although surfactants with values higher than 10 are often used for this purpose.

## Deflocculants and Dispersing Agents

These agents do not appreciably lower surface and interfacial tension; thus, have little tendency to create foam or wet particles. Most deflocculants, however, are not generally considered safe for internal use and as a result the only acceptable dispersant for oral products is lecithin or a lecithin derivative (naturally occurring mixture of phosphatides and phospholipids). Because lecithins vary in water solubility and dispersibility characteristics, proper control of product specifications must be maintained to obtain reproducibility.

A proper surfactant in appropriate concentration improves the dispersion by reducing the interfacial tension. For example, if surfactants with negative charges are adsorbed on the particles, this prevents or minimizes aggregation in the presence of positive ions because of the mutual repulsion of like charges. Examples of such agents include sodium lauryl sulfate and sodium dioctyl sulfosuccinate. In practice, a non-ionic surfactant is usually used to aid the dispersion of insoluble phase. Polyoxyethylene ethers of mixed partial fatty acid esters of sorbitol anhydrides (Tweens), the same compounds without the hydrophilic oxyethylene groups (Spans), higher molecular weight polyethylene glycols (Carbowaxes), and molecular combinations of polyoxyethylene and polyoxypropylene (Pluronics), are frequently used in this manner. One must be careful, especially with the non-ionic surfactants, not to use too high a concentration of the agent. Above the critical micelle concentration, intact micelles are adsorbed to particle surfaces, providing a continuous film for coagule formation, as discussed previously.

## Flocculating Agents

Before aggregating the suspension particles in the open network aggregate, it is important to ensure that the particles are well dispersed in the aqueous phase or other vehicle. Flocculating agents are neutral electrolytes that are capable of reducing the zeta potential of suspended charged particles to zero. The concentration of added electrolyte necessary to produce the flocculated state corresponds to the quantification expressed in the Schulze-Hardy rule previously discussed. Monovalent ions such as, sodium or potassium chloride in small concentrations (0.01–1% w/v), are often sufficient to induce flocculation of weakly charged, water-insoluble, organic nonelectrolytes. In the case of insoluble, highly charged, and polyelectrolyte species, similar concentrations (0.01–1% w/v) of water-soluble divalent or trivalent ions, such as calcium salts, aluminum chloride, sulfates, citrates, and potassium biphosphate, may be required for floc formation, depending on particle charge (positive or negative).

## Suspending Agents

Suspending agents retard settling and agglomeration of the particles by functioning as an energy barrier, which minimizes interparticle attraction and ultimate aggregation. The general choice of suspending agents includes (i) protective colloids, (ii) viscosity-enhancing agents and, (iii) surfactants. Some of the suspending agents used to a large extent in formulation include modified cellulose polymers, proteins such as gelatin, and totally synthetic polymers.

## Protective Colloids

Suspension systems intended for oral, parenteral, ophthalmic, or topical use may not appear elegant because they usually exhibit poor drainage in vials or bottles due to the clusters of particles. These properties may be improved by the addition of protective colloids. Protective colloids differ from surfactants in that they do not reduce interfacial tension. Their solutions differ in viscosity and are used in higher concentrations than surfactants. Protective colloids also differ from other agents in that their effect is due not only to their ability to increase the zeta potential, but also to their formation of a mechanical barrier or sheath around the particles.

**Modified cellulose polymers:** Sodium carboxymethylcellulose (CMC), methylcellulose (Methocel), and hydroxypropylmethylcellulose (Methocel, HG) are widely used in oral, topical, and parenteral dosage forms. Sodium carboxymethylcellulose is an anionic polymer, whereas methylcellulose and hydroxypropylmethylcellulose are non-ionic. Sodium carboxymethylcellulose is used in concentrations of up to 0.5% (w/v) in injectable preparations. In oral dosage forms, they are frequently used in higher concentrations because of the higher solids content of the systems. Sodium carboxymethylcellulose does have some disadvantages: It is incompatible with a number of electrolytes and quaternary ammonium compounds, and forms complexes with certain surfactants. Methylcellulose and hydroxypropylmethylcellulose gel on heating and are affected by electrolytes. One of the more useful of the totally synthetic polymers is polyacrylic acid (Carbopol). It is used mostly in external lotion and gel preparations. After being uniformly dispersed in water, it is neutralized with an organic amine such as triethanolamine or with inorganic alkali to achieve the desired viscosity in a range from pH 6 to 10. The material is extremely sensitive to electrolytes, but is suited for use equally in aqueous and non-aqueous systems. These polymeric agents function primarily as protective colloids and alter the viscosity of the medium.

**Clays:** As a group, the clays (essentially hydrated aluminum and/or magnesium silicates) are also quite useful in suspension formulation. They hydrate further in water to a high degree to form colloidal dispersions having high viscosities. The manner in which the members of the group are prepared, however, has a profound effect on the final product. The clay

should always be added to the water with high shear to effect uniform dispersion and maximum hydration. The pH of aqueous clay dispersions is somewhat alkaline, in the range of pH 8.5 to 9.5; therefore, they also possess some acid-neutralizing capacity. The viscosity of aqueous dispersions of these agents varies, depending on the type and amount of solids dispersed. In general, 5 to 10% (w/v) concentrations of the clay form firm opaque gels. Clays can be formulated in systems in which the pH is between 6 and 11, but they are most stable between pH 9 and 11. Alkaline buffers usually are included to maintain the pH. All dispersions of clays are drastically affected by electrolytes in accordance with the Schulze-Hardy rule. Ethanol also affects these agents by dehydrating the colloid, thereby reducing the viscosity. Clay suspensions and gels are excellent media for mold and bacterial growth, and should therefore be adequately preserved with nonionic antimicrobial preservatives. The paraben esters and benzoates are useful, but cationic “quaternary” preservatives are ineffective. Heat and aging usually increase the viscosity of clay mixtures, but this may become less significant in the more concentrated systems.

## Formulation Adjuvants

Certain aspects of suspension formulation pertain to both aggregate and dispersed suspension systems and are therefore discussed together. Suspension adjuvants must be considered. These agents include the preservative, colour, perfume, and flavour; which may materially affect the characteristics of the suspension system. In general, most colours are used in small quantities and are usually compatible; flavours and perfumes are similarly used and are also usually compatible with the vehicle. To illustrate that one must be on guard against adjuvant interaction, it is noted that Bean and Dempsey showed that the adsorptive power of kaolin suspensions can reduce the activity of some preservatives. Kaolin hurts the quaternary benzalkonium chloride (BAK), but not the non-ionic and less surface-active m-cresol. Similarly, it was shown, that procaine penicillin adsorbed the quaternary BAK and that the supernatants of the BAK systems had less activity than the suspensions; thus, the adsorbent acted as a reservoir.

## SUSPENSION FORMATION

The processes involved in suspension formation are shown in Fig. 18.13. The flocculated state (C) may be reached either directly by wetting and dispersing hydrophobic particles with a suitable flocculating agent, or else by first wetting and dispersing to produce a disperse or deflocculated state (B) with a suitable surfactant and then flocculating with a suitable agent such as a hydrophilic colloid or polyelectrolyte. In contrast to deflocculated or peptized particles, flocculated suspensions (C), which are considered pharmaceutically stable (although colloidally unstable), can always be redispersed with gentle agitation. Addition of too much flocculating agent results in overflocculation and tends to produce agglomerated or coagulated irreversible systems (E). The term plaque (platelike) is used to describe essentially flat agglomerates, whereas the term coagula (clump-like) are reserved for thicker, three-dimensional particle masses. In the absence of a protective colloid, the process of crystal growth is indicated by the arrow connecting (A) to (D).

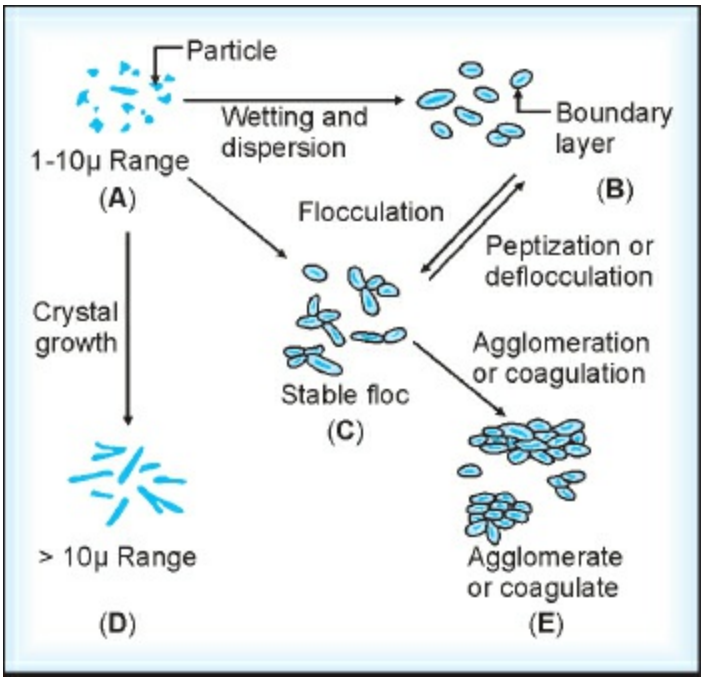



**Fig. 18.13:** Processes involved in suspension formation

## Precipitation Methods

Three precipitation methods are discussed in this section: organic solvent precipitation, precipitation effected by changing the pH of the medium, and double decomposition.

### *Organic Solvent Precipitation*

Water-insoluble drugs can be precipitated by dissolving them in water-miscible organic solvents and then adding the organic phase to distilled water under standard conditions. Examples of organic solvents used include ethanol, methanol, propylene glycol, and polyethylene glycol. Several important considerations are involved when this method is used. Perhaps the most important factor next to particle size control is that the "correct" polymorphic form or hydrate of the crystal be obtained. For example, different forms are obtained when prednisolone is precipitated from aqueous methanol as opposed to aqueous acetone. Besides the influence of the solvent on crystal characteristics, the following additional factors may need to be considered: possible preparation under sterile conditions, inherent solvent entrapment and subsequent toxicity, volume ratios of the organic to the aqueous phase, rate and method of addition of one phase to the other, temperature control (cooling rate and drying conditions), method of drying the precipitate (forced air, vacuum, or freeze drying), and finally, the washing of the precipitate. Where pertinent, sterilant residues should not be overlooked (e.g. ethylene glycol from ethylene oxide gas sterilization procedures).

### *Precipitation by pH*

The method of changing the pH of the medium is perhaps more readily accomplished and does not present the same difficulties associated with organic solvent precipitation. The technique, however, is only applicable to those drugs in which solubility is dependent on the pH value. For example, estradiol suspensions can be prepared by changing the pH of its aqueous solution; estradiol is readily soluble in such alkali as potassium or sodium hydroxide solutions. If a concentrated solution of estradiol is thus prepared and added to a weakly acidic solution of hydrochloric, citric, or acetic acids, under proper conditions of agitation, the estradiol is precipitated in a fine state of subdivision. The type of crystal or polymorphic form depends on

such factors as the concentrations of acid and base and the degree and type of fluid shear imparted to the system.

Insulin suspensions also may be prepared by a pH change method. Insulin has an isoelectric point of approximately pH 5. When it is mixed with a basic protein, such as protamine, it is readily precipitated when the pH is between the isoelectric points of the two components, i.e. pH 6.9 to 7.3. Protamine zinc insulin (PZI) contains an excessive quantity of zinc to retard absorption. According to the British Pharmacopoeia of 1958, a phosphate buffer is added to each individual vial containing the acidified solution of insulin, protamine, and zinc, so that the pH is between 6.9 and 7.3. The preparation is compounded in the final container by mixing the PZI and the buffer in the filling operation. Adreno-corticotropin (ACTH) zinc suspensions are prepared in a similar manner. The precipitate formed in the process is zinc hydroxide or zinc phosphate, on which the ACTH is adsorbed. This combination results in a long-acting preparation when administered. The addition of phosphate salts and organic phosphate to prepare an even longer acting ACTH preparation is also possible.

When either the change in pH or the organic solvent precipitation method is used to prepare a suspension, a degree of supersaturation is brought about suddenly in the batch process to give rise to crystal nucleation and growth, after which the initial supersaturation subsides. Thus, the degree of supersaturation changes throughout the process, and neither the rate of nucleation nor the rate of crystal growth is constant; therefore, the particle size distribution is variable. The degree of supersaturation and rate of nucleation are greatest at the beginning of the process, so that crystals formed initially become the largest because they are exposed to the supersaturated solution for the longest period of time. It appears, therefore, that when less concentrated solutions are used, the particle size distribution is broader than when more concentrated solutions are used.

## *Double Decomposition*

Making suspensions by double decomposition involves only simple chemistry, although some of the aforementioned physical factors also come into play. The reader is referred to standard pharmacy texts to review the preparation of White Lotion (NF XIII), i.e. forming zinc “polysulfide” by mixing zinc sulfate and sulfurated potash solutions.

## Dispersion Methods

When the dispersion method is utilized for suspension preparation, the vehicle must be formulated so that the solid phase is easily wetted and dispersed. The use of surfactants is desirable to ensure uniform wetting of hydrophobic solids. The use of suspending agents, such as the synthetic polymeric polyelectrolytes, natural gums, or clay, may be indicated, depending on the specific application. The actual method of dispersing the solid is one of the more important considerations because particle-size reduction may or may not result from the dispersion process. If particle-size reduction occurs, the particles obtained may have different solubilities if a metastable state is involved, and this may lead to transient supersaturation of the system. A number of dispersion methods are used to prepare suspension products. For the present purposes, there is no need to describe and discuss the comminuting and shearing equipment commercially available because information on such equipment is easily obtained. The reader need only recall that much of what has been and will be discussed with respect to basic suspension technology applies regardless of how the suspension is made.

## Preparative Techniques

The actual preparation of suspensions involves choosing the ingredients (utilizing principles already discussed) and determining the type of manufacturing equipment to be used. Needless to say, each suspension is a separate case and absolute generalization is not possible. If the suspension is made by a dispersion process, it is best to achieve pulverization of the solid by a micronization technique. This involves subjecting the particles to a turbulent air chamber in which they collide with each other and fracture. Particles under 5 μm are readily obtained. Although it is not widely used for this purpose, spray-drying also can be considered a method of comminution to produce a finely divided solid phase.

If the suspension is made by controlled crystallization, a supersaturated solution should be formed and then quickly cooled with rapid stirring. This causes the formation of many nuclei and hence many crystals. It is just the opposite of letting crystals grow large.

At some time during suspension formation, it is likely that shearing will be desired. This homogenization can be accomplished by the conventional stator-rotor colloid mills. Ultrasonic equipment also can be used to effect high-intensity mixing, but usually, this technique is not applied commercially. Of interest, however, is the work of Sheikh, Price, and Gerraughty, who studied the effect of ultrasound on polyethylene spheres in aqueous suspension. The ultrasound reduced the sphere size only when surfactants were added, especially those having high HLBs. When such agents were used as additives, the particles were readily dispersed and hence completely surrounded by liquid. Since ultrasound waves and cavitation shock waves are transmitted to the particles through the liquid medium, a poor suspension would not be as susceptible to size reduction as a better dispersed one. Excessive shearing (or high temperatures) may irreversibly damage polymeric materials such as gums, so that viscosity loss is suffered. Instead of trying to hydrate gums and clays by massive shearing, it is often better, when possible, to give the material the necessary time to hydrate under conditions of mild shearing. An alternate procedure is to mix with, or preferably spray the gum with, a chlorinated hydrocarbon, acetone, or alcohol solution of a wetting agent (e.g. sodium dioctyl sulfosuccinate). About 0.4% (based on the gum weight) of the wetting agent should be added to the gum.

This technique can produce a marked beneficial effect, as wetting of the gum and hence hydration is greatly accelerated.

A final comment is that processing studies in a pilot plant are needed because it is axiomatic that the scale-up operation from laboratory batches to production lots brings with it many troubles and unexpected results.

## Controlled Flocculation

The aim in the formulation of suspensions is to achieve partial or controlled flocculation. The main advantages of the stable floc are as follows. The aggregates tend to break up easily under the agitation of a bottle or vial, or by flow through a small orifice (hypodermic needle and/or syringe) and reform an extended network of particles after the force is removed. Flocculation, therefore, imparts a structure to the suspension with virtually no increase in viscosity. The following examples illustrate how suspensions may be prepared by controlled flocculation procedures:

1. The wetting agent (not more than 0.1–0.2% (w/v) of the final concentration), is dissolved in approximately half the final volume of aqueous vehicle.
2. Microfine particles of the drug at the desired concentration are uniformily spread over the surface of the vehicle and drug is allowed to be wetted undisturbed for as long as 16 h.
3. The wetted slurry is passed through a fine wire sieve (100 mesh size) or wetted slurry is passed through a colloid mill to remove poorly-wetted powder.
4. The slurry concentrate of the drug is agitated gently with an impeller-type mixer.
5. To the slurry, flocculating agent is added till flocculation end point is reached.
6. To determine the endpoint, small samples are transferred to a graduated cylinder, an equal amount of vehicle is added and the cylinders are gently shaken and permitted to stand undisturbed. The sample with the highest ratio of sediment to total suspension volume, exhibiting a clear supernatant and good drainage characteristics is considered to be at the appropriate endpoint.
7. After the flocculation endpoint has been established and verified, the other formulation adjuvants (preservative, colorant, flavour, buffer, etc.) are added, and the slurry is brought to final volume with liquid vehicle.

## Structured Vehicle

Another technique for the preparation of a stable suspension is based on the concept of “structured vehicle” in which the viscosity of the preparation, under static conditions of very low shear, on storage approaches infinity. The vehicle is said to behave like a “false body” that is able to maintain the suspended particles in a state of more or less permanent suspension. Structured vehicles are avoided for the preparation of parenteral suspensions, owing to their high viscosity.

### *Bingham-type Plastic Flow*

Vehicles with Bingham-type plastic rheological flow are characterized by the need to overcome a finite yield stress before flow is initiated. Permanent suspension of most pharmaceutical systems requires yield-stress values of at least 2–5 Pa (20–50 dyn/cm$^2$). Bingham plastic flow is produced by carbomers. Carbomers exhibit a sufficiently high yield value at low solution concentration and low viscosity to produce permanent suspensions.

### *Thixotropic Flow*

Thixotropic flow is defined as a reversible, time-dependent, isothermal gel-sol transition. Thixotropic systems exhibit easy flow at high shear rates and on removing the stress the system is slowly reformed into a structured vehicle. The usual property of thixotropy results from the breakdown and build-up of floccules under stress. The primary advantage of thixotropic flow is that it confers pourability under shear stress and viscosity and sufficiently high yield stress when the shear stress is removed at rest. Pseudoplastic materials (such as hydroxyethylcellulose, hydroxypropyl methyl cellulose or sodium carboxymethyl-cellulose) in combination with a clay (hydrated colloidal magnesium aluminum silicate) or blends and coprecipitates of sodium carboxymethylcellulose and microcrystalline cellulose exhibit thixotropic flow.

### *Emulsion Base*

A wax-type self-emulsifier develops structure or “false body” in suspension systems. A dilute emulsion system is not often considered for suspension purposes because of the potential complexities involved in mixing emulsion

and suspension systems. The drug particles are dispersed in the primary emulsion component prior to dilution with other vehicle components.

## EVALUATION OF SUSPENSION STABILITY

Since stability testing is discussed elsewhere in the pharmaceutical literature, the only emphasis here is on the most pertinent aspects of suspension stability. Techniques for the evaluation of heterogeneous systems generally are complex and are far from being completely satisfactory. Some test methods are so drastic that the stability information is obtained during an evaluation that destroys the system being evaluated. Some methods are somewhat empiric in nature, i.e. the exact basis on which they operate cannot be explicitly defined mathematically. All test procedures suffer some limitations, and the results, therefore, must be cautiously evaluated and interpreted. As the methodology involved in the pertinent stability studies is often somewhat complicated, this section of the chapter is more fully referenced so that further details can be obtained if desired. The purpose here is to point out explicitly one method, and then indicate only the general nature of some of the other approaches taken. Use of evaluation techniques permits the formulator to screen the initial preparations made and also to compare the improved formulations to competitive commercial products. The latter point should not be treated lightly even though it does not deal with absolute standards.

## Sedimentation Volume

Since redispersibility is one of the major considerations in assessing the acceptability of a suspension, and since the sediment formed should be easily dispersed by moderate shaking to yield a homogeneous system, measurement of the sedimentation volume and its ease of redispersion form two of the most common basic evaluation procedures.

The concept of sedimentation volume is simple. In short, it considers the ratio of the ultimate height ($H_u$) of the sediment to the initial height ($H_o$) of the total suspension as the suspension settles in a cylinder under standard conditions.

$$\textit{Sedimentation volume} = H_u/H_o$$

The larger this fraction, the better is the suspendability (Fig 18.14). Methods utilizing the sedimentation volume obtained in a cylinder offer a practical approach to the determination of the physical stability of suspension systems. Particularly good is the fact that the system remains undisturbed. Specifically, it is worth knowing how to use the $H_u$ and $H_o$ concepts. The formulator should obtain the $H_u/H_o$ ratios and plot them as ordinates with time as the abscissae.

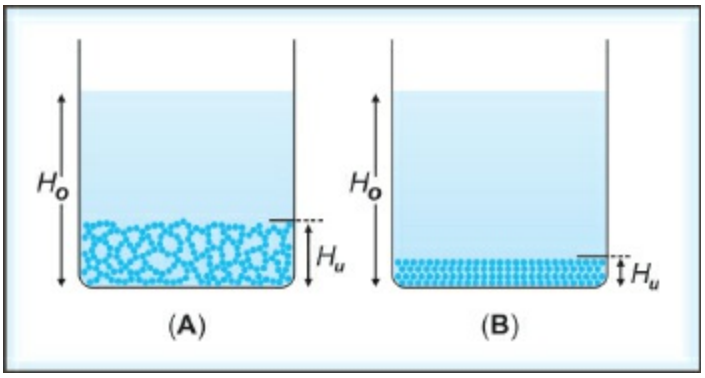



**Fig. 18.14:** Sedimentation of: (A) Flocculated; (B) Deflocculated suspensions

Note that although the conventional $H_u$ is called the "ultimate" height of the sediment, ultimate really means the height at any particular time. The plot just described will at time zero start at 1.0, with the curve then being either horizontal or gradually sloping downward to the right as time goes on. One can compare different formulations and choose the best by observing the

lines, the better formulations obviously producing lines that are more horizontal and/or less steep.

## *Concentrated Suspensions*

Another technique that utilizes essentially the same parameters may be used to evaluate highly concentrated suspensions, which might be difficult to compare because there would be only minimum supernatant liquid. The technique involves diluting the suspension with additional vehicle, i.e. with the total formula with all ingredients except the insoluble phase. As an example, one could dilute 50 ml of a suspension to a volume of 100, 150, or 200 ml. The $H_u$ reading then becomes the volume of sediment in the diluted sample, and $H_o$ equals the original volume of the sample before dilution. The $H_u/H_o$ ratio may in this case be greater than 1. Regardless, the ratio is again plotted against time, and comparisons, between formulas are made as before.

## Degree of Flocculation

One additional concept should also be considered by the formulator. In all the comparisons just mentioned, the screening technique results only in a relative ranking which indicates which preparations are the better ones. It is also useful, however, to consider the possibility of making an absolute evaluation; which may be done as follows:

The degree of flocculation is the ratio of sedimentation volume of the flocculated suspension ($F$) to the sedimentation volume that would be produced in the ultimate dispersed state ($F_\infty$).

$$\text{Degree of flocculation} = F/F_\infty$$

To obtain the completely dispersed suspension form, which represents the least void space for the solid phase and hence the smallest sedimentation volume, electrolytes that promote settling may be added or the preparation may be centrifuged. The $H_u/H_o$ ratio observed is then the lowest figure obtainable. This figure serves as a base line and gives some idea of the degree of aggregation obtained because ratios higher than this minimum represent the existence of the desired aggregated state. In reference to the plots discussed, it is clear that data that produce a line that quickly drops towards this reference point do not represent a good suspension, as any aggregation if there is any at all, is too temporary and infirm.

Another use for $H_u/H_o$ data is possible, and particularly pertinent are the various relationships of Ward and Kammermeyer. In essence, these workers attempted to quantitate settling further using $H_u$ and $H_o$ values. It is known that the ultimate height of the solid phase after settling depends on the concentration of solids and the particle size. These workers found that if $H_o$ and $H_u$ readings (taken on a series of different concentrations of the same solids having a particular average particle size range) are measured in a certain vehicle, the resulting data form a straight line plot if the logarithm of the weight percentage of solids is plotted against the ratio $H_u/H_o$. One can then predict $H_u$ for any given solids concentration by multiplying $H_o$ by the "relative concentration factor", i.e. by $H_u/H_o$.

## Redispersibility

As noted, the evaluation of redispersibility is also important. To help quantitate this parameter to some extent, a mechanical shaking device may be used. It simulates human arm motion during the shaking process and can give reproducible results when used under controlled conditions. It should be remembered, however, that the test conditions are not the same as those encountered under actual use, and further testing should be considered. Nevertheless, the test results are useful and provide guidance during screening procedures.

## Rheologic Methods

In addition to techniques involving sedimentation and redispersibility factors, rheologic methods can also be used to help determine the settling behaviour and arrangement of the vehicle and particle structural features for purposes of comparison.

The majority of rheologic investigations of suspension systems have been done at high shear rates and on systems that must be made uniform before evaluation. For present purposes, the importance of using low shear rates and undisturbed samples cannot be overemphasized. The prime reason for this is the fact that the structure achieved on storage is what should be evaluated. A practical rheologic method involves the use of the Brookfield viscometer mounted on a helipath stand. The T-bar spindle is made to descend slowly into the suspension, and the dial reading on the viscometer is then a measure of the resistance the spindle meets at various levels in a sediment. In this technique, the T-bar is continually changing position and measures undisturbed samples as it advances down into the suspension. This technique also indicates in which level of the suspension the structure is greater, owing to particle agglomeration, because the T-bar descends as it rotates, and the bar is continually entering new and essentially undisturbed material. Data obtained on samples variously aged and stored indicate whether undesired changes are taking place. Thus, using the T-bar spindle and the helipath, the dial reading can be plotted against the number of turns of the spindle. This measurement is made on undisturbed samples of different ages. The results indicate how the particles are settling with time. In a screening study, the better suspensions show a lesser rate of increase of dial reading with spindle turns, i.e. the curve is horizontal for a longer period.

A method combining the use of both rheologic and sedimentation parameters is illustrated by the work of Foernzler, Martin, and Banker, who studied the effect of thixotropy on stability. Although this method does not observe the system under equilibrium conditions and is subject to some challenge, the authors attempted to predict physical stability by a rheologic evaluation of thixotropy. Incidentally, Wood used these workers' data to develop additional correlations. It is important to note that the use of most viscometers and centrifuges in stability studies is not ideal for aggregated systems because their use destroys the structure formed.

## Electrokinetic Techniques

Microelectrophoresis apparatus permits the measurement of the migration velocity of the particles with respect to the surface electric charge or the familiar zeta potential; the latter has units of viscosity times the electrophoretic mobility, or more familiarly, volts. Stanko and DeKay also evaluated suspensions by electrokinetic methods and showed that the zeta potential changes upon the addition of additives and is related to stability. Haines and Martin studied some of the formulation factors that influence the stability of suspensions. They correlated the zeta potential to visually observed caking; zeta potential was again determined by microscopic electrophoresis. It was found that certain zeta potentials produced more stable suspensions because aggregation was controlled and optimized.

## Particle Size Changes

The freeze-thaw cycling technique is particularly applicable to stressing suspensions for stability testing purposes. This treatment promotes particle growth and may indicate the probable future state of affairs after long-term storage at room temperature. Thus, it is of prime importance to be alert for changes in absolute particle size, particle size distribution, and crystal habit. With respect to the latter point, Carless et al. investigated the various crystal forms of cortisone acetate and also noted the acceleration of sulfathiazole crystal growth in suspensions that underwent temperature cycling. Obviously, the physiologic availability and thus the therapeutic effect of the active ingredients may be influenced by such changes. Particle size distributions are sometimes determined by microscopic means. This method of necessity requires dilute suspensions that are counted with the aid of an ocular grid. In some instances, photomicrographs may be taken for permanent records. This method is quite tedious, especially when large numbers of samples are to be evaluated. It is worth noting that certain suspension components, e.g. preservatives or protective colloids, may have a profound effect on the physical performance of the suspension under freeze-thaw conditions. When a low solids content steroid injectable preparation containing sodium carboxymethylcellulose (CMC) and benzyl alcohol, and one containing CMC, methylparaben, and propylparaben, were subjected to freezing and thawing, the former suspension caked badly, while the latter was unaffected. Protective colloids may thus be adversely affected by freezing, thawing, or elevated temperatures; for example, gelatin is sensitive to low temperatures whereas methylcellulose is adversely affected by higher temperatures. Although freeze-thaw cycle studies are useful guides, the best stability information is still obtained from studies conducted at room temperature.

## PACKAGING

Perhaps the final “adjuvant” one should consider is the package. Usually, initial laboratory screening employs conventional graduates or readily available botdes. When final packaging is considered, it should be noted that various types of glass are available. The types vary with respect to their ability to resist water attack, the degree of attack being related to the amount of alkali released from the glass. The USP should be consulted for further details, as it describes both the tests and standards that should be met by containers to be used for packaging parenteral and nonparenteral (oral or topical) products. One point of terminology may be noted: “flint” refers to clear, colour less, brilliant glass. Originally, it contained lead and was also called “lead” or “crystal” glass; however, today in commerce, non-lead, highly colour-free, soda-lime-silica glasses, the most common general-purpose transparent glasses, are also called flint. Parenteral multiple-dose vials may be “flint” (colour-less) or amber, and may be silicone-coated to improve drainage of the suspensions (silicone coating also minimizes the leaching of alkali from the glass). This technique of silicone coating is used widely for suspensions of steroids and combinations of penicillin and dihydrostreptomycin. It is also used in preparations with high solids content, in which formulation modifications cannot measurably improve the drainage of the preparation.

There has been a trend to package suspension systems for oral and topical administration in polyethylene or other plastic containers. Many factors must be considered when a suspension is evaluated in such a container. These factors include loss of flavour and perfume, preservative adsorption, and leaching into the product of substances from the container. Before evaluative procedures are discussed per se, it must be stressed that after the initial stability observations are completed, the determination of the stability of the suspension in the *final package* is an important step of the product development procedure.

## BIOPHARMACEUTICAL CONSIDERATIONS

On a theoretic basis, one would expect the drug bioavailability from a suspension to be equal or somewhat better than that from a tablet during the first hour after administration of the dosage form. This is because the tablet must invariably undergo disintegration before drug dissolution can occur. The suspension, on the other hand, already contains discrete drug particles. Data showing that the suspension drug dosage form is either equally or more bioavailable during the first hour after administration have been documented in the literature.

In suspension, the drug is present in the form of solid particles, which must disperse in the gastrointestinal media and dissolve in them. The rate of drug dissolution, and potentially the drug bioavailability, can be affected by such physical factors as dispersibility, particle size and shape, and polymorphism.

Considering the hydrodynamic conditions generated by the mild agitation of the gastrointestinal musculature, one would expect the suspending agents to influence the efficacy of suspensions with poor dispersion characteristics in the gastric milieu. Antacid suspensions are the case in point. This class of products also demonstrates that the suspension dosage form is far better than the tablet in terms of in vivo efficacy. Using their peristaltic assembly, Simmons and co-authors demonstrated an excellent correlation between the in vitro and in vivo neutralization capacity of a commercial alumina and magnesia oral suspension. Using the in vitro method, which is claimed to simulate the mild agitation in the stomach, they showed that there were significant differences in the neutralization capacity of various commercial antacid suspensions. One product, in particular, failed to disperse through the reaction medium, which can be attributed to the nature of the suspending agent in this product. Similarly, Fordtran and co-authors, and Drake and Hollander, have shown the varying neutralization capacity of numerous commercial antacid suspensions, as illustrated in Table 18.2.

In addition to the suspending agent, the nature of the raw material and the manufacturing process (milling and homogenization) exert a significant effect on the neutralization capacity of antacid suspensions. For example, aluminum hydroxide gel, AHLT-LW (Chattem Chemicals Division, Chattem Inc.), showed superior acid neutralization capacity in comparison to other similar

raw materials when tested under mild agitation conditions. Because of its fluid nature, this raw material is also pumpable during large-scale production of the antacid suspension. The milling operation reduces the size of the suspended antacid particles, thereby making them more reactive with the gastric acid under mild agitation. The aforementioned factors may account for the observed differences in the neutralization capacity of various commercial antacids in Table 18.2, and of those reported elsewhere.

The influence of particle size, vehicle, and additive on the absorption profile of intramuscular procaine penicillin G suspensions is discussed in the extensive work of Buckwalter and Dickinson. Their study showed that larger crystals of drug suspended in water, sesame oil, and peanut oil gave delayed drug absorption with prolonged blood levels; however, when 2% aluminum monostearate was included as a gelling agent in sesame and peanut oils, the micronized drug form exhibited a more prolonged blood level in rabbits. This effect was further confirmed by a study in humans. It might have been due to the incipient depot formation brought about by the gelling agent. Frederick proposed that it was a result of the marked cementing action of the fine particles. As stated in the previous discussions and as demonstrated in this section, the desired form of sedimentation is a state of controlled aggregation. Complete aggregation may produce a formulation with too coarse an appearance, but just the right degree provides a truly elegant final product. In terms of bioavailability, a suspension should be easily dispersed upon shaking, allowing for the removal of a precise drug dose during administration. The suspending agent should permit free and easy drug dispersion in the gastric (or other body) medium under mild agitational conditions.

**Table 18.2:** In vitro neutralization capacity of several commercial antacids

| Product (Manufacturer) | Neutralization capacity (mEq/ml)† | Antacid content (mg/ml) | |
|---|---|---|---|
| | | Aluminum hydroxide | Magnesium hydroxide |
| Maalox TC* (Rorer) | 4.2 | 120 | 60 |
| Mylanta 11** (Stuart) | 4.14 | 80 | 80 |
| Delcid* (Merrell-Dow) | 4.1 | 120 | 133 |
| Gelusil II* (Parke-Davis) | 3.0 | 80 | 80 |
| Aludrox** (Wyeth) | 2.81 | 61.4 | 20.6 |
| Maalox** (Rorer) | 2.58 | 45.4 | 40 |
| Di-Gel** (Plough) | 2.45 | 56.2 | 17.4 |
| Mylanta** (Stuart) | 2.38 | 40 | 40 |
| Silain-Gel** (Robins) | 2.31 | 56.4 | 56.4 |
| Maalox Plus* (Rorer) | 2.3 | 45.4 | 40 |
| Gelusil* (Parke-Davis) | 2.2 | 40 | 40 |
| Amphogel** (Wyeth) | 1.93 | 64 | – |
| Kolantyl Gel** (Merrell-Dow) | 1.69 | 15 | 15 |

† A mEq of antacid is defined as the mEq of HCl that is required to maintain the antacid suspension at pH 3 in vitro for a specific time.

** pH 3 maintained for 2 hours.

* pH 3 maintained for 1 hour

## ILLUSTRATIVE EXAMPLES

This section presents certain formulas that demonstrate some of the principles discussed in this chapter. Additives exert a great influence on the stability and drug bioavailability of formulations. The whole area of suspension product development is a study of the performance characteristics of adjuvants. Skill in finding and matching the everlasting marriage between the suspending characteristics and desired drug bioavailability determines the degree of excellence represented in the final product. It is also the purpose of this section to discuss biopharmaceutical aspects of the suspending agents in contemporary suspensions.

Illustrative formulations are discussed here in these categories (a) suspensions with wetting and aggregating agents, (b) suspensions with low solids content, (c) suspensions with high solids content, (d) antacids (e) biopharmaceutical considerations.

## Suspensions with Wetting and Aggregating Agents

A patent by Macek describes a non-caking aqueous parenteral suspension of cortisone acetate and the wetting principle as follows:

| Ingredients | Percentage in formula |
|---|---|
| Cortisone acetate, USP microfin | 2.5 |
| Polysorbate 80, USP (wetting agent) | 0.4 |
| Carboxymethylcellulose sodium, USP (suspending agent) | 0.5 |
| Sodium chloride, USP (for isotonicity) | 0.9 |
| Benzyl alcohol, NF (preservative) | 0.9 |
| Water for injection, USP, q.s. to make | 100.0 |

On the industrial scale, all the ingredients except cortisone acetate can be dissolved in water for injection using suitable mixing equipment. Cortisone acetate can now be dispersed in the solution. The entire dispersion may be passed through a colloid mill.

The principles of aggregation and wetting may also be observed by making the preparations labeled A through F in the table below.

Comparison of the preparations A and B demonstrates the value of a wetting agent. It is easy to disperse sulfamerazine in B, but it settles on standing and forms a cake because of its deaggregated nature. In C, the inclusion of the cationic aluminum (which should be added last) yields an aggregated suspension. Although the aggregates settle rapidly, they form a high volume sediment that does not cake and is easily redispersed.

Preparation D is similar to C except that it contains carboxymethylcellulose, which acts as a protective colloid and a viscosity builder. Therefore, this aggregated suspension does not cake on standing and settles more slowly than C. From a physical stability point of view, D is the best suspension in this series. Suspension E cakes on standing because the anionic diphosphate does not cause aggregation as was evident in C, in which the cationic aluminum acted as an aggregating agent. Preparation F illustrates a gross incompatibility of a negatively-charged cellulose derivative with the

positively-charged aluminum ions.

For additional examples illustrating the phenomena of aggregation and wetting in aerosol dispersion systems, the reader is referred to Chapter 21, “Pharmaceutical Aerosols”.

## Suspensions with Low Solids Content

Table 18.3 illustrates the components that are required to prepare a model parenteral suspension. This route of administration limits the formulator to a rather narrow range of additives. The samples are best prepared by making a concentrate of the dispersant in a volume equal to 10% of the final volume, thoroughly mixing in the active ingredient with the help of a colloid mill or other device, and adding the remaining components to a solution of the preservative(s). The latter should be prepared using about 80% of the final total volume. This solution is then added to the portion containing the active ingredient, and sufficient purified water is added to bring it to the final volume. Note that the preservatives are added in a slightly excess amount to compensate for their binding to polysorbate. This binding has been shown to inactivate the preservative in direct proportion to the amount bound.

| | Percentage in preparations | | | | | |
|---|---|---|---|---|---|---|
| **Ingredients** | **A** | **B** | **C** | **D** | **E** | **F** |
| Sulfamerazine, USP | 2.0 | 2.0 | 2.0 | 2.0 | 2.0 | 2.0 |
| Docusate sodium,* USP | — | 0.2 | 0.2 | 0.15 | 0.2 | — |
| Aluminum chloride hexahydrate,† USP | — | — | 0.1 | 0.1 | — | 0.25 |
| Carboxymethylcellulose sodium (7MP),‡ USP | — | — | — | 0.02 | — | 0.15 |
| Potassium biphosphate, NF | — | — | — | — | 0.1 | — |
| Purified water, USP q.s. to make | 100 | 100 | 100 | 100 | 100 | 100 |

The following observations are made with respect to Table 18.3:

A — No dispersion or very little wetting of solid; this may depend on the recrystallization solvent (acetone versus dimethylformamide).
B — Good dispersion, rapid settling, caking.
C — Good dispersion, rapid settling, severe caking, poor redispersibility, deaggregation.
D — Good dispersion, rapid settling, slight aggregation.
E — Good dispersion, slow settling, moderate aggregation.
F — Good dispersion, slow settling, fine aggregation.
G — Good dispersion, slow settling, aggregation.
H — Good dispersion, slow settling, coarse aggregation.

**Table 18.3:** Low solids content suspensions

| | Concentration in mg/ml | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Sample** | **A** | **B** | **C** | **D** | **E** | **F** | **G** | **H** |
| Steroid* | 25 | 25 | 25 | 25 | 25 | 25 | 25 | 25 |
| Polysorbate 80† (dispersant) | — | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1:0 | 1.0 |
| Sodium citrate (buffer) | — | — | — | 10.0 | — | — | — | — |
| Sodium chloride (for isotonicity) | 9.0 | 9.0 | 9.0 | — | 9.0 | 9.0 | 9.0 | 9.0 |
| Benzyl alcohol (preservative) | — | — | 9.0 | 9.0 | 9.0 | — | 9.0 | — |
| Chlorobutanol (preservative) | — | — | — | — | 5.0 | 5.0 | — | — |
| Methylparaben (preservative) | — | — | — | — | — | — | 1.8 | 1.8 |
| Propylparaben (preservative) | — | — | — | — | — | — | 0.2 | 0.2 |
| Purified water q.s. to make 1.00 ml | | | | | | | | |

*Cortisone acetate or prednisolone acetate.

†Tween 60 or Tween 40 could also be used.

It is important to note that protective colloids, such as polyethylene glycol 4000, carboxymethylcellulose sodium, and methylcellulose all modify these characteristics. Sorbitol or dextrose can be included to adjust density.

## Suspensions with High Solids Content

Table 18.4 illustrates a different set of sample preparations. Again, it should be recognized that additives often markedly affect the properties of the formulation.

The samples are best prepared as follows. The lecithin is added to the penicillin G, then the remaining vehicle containing the other components is added. The product is first passed through a 40-mesh screen, then through a colloid mill, so that the procaine penicillin G is uniformly coated with lecithin. The products must be placed in silicone-treated vials or cylinders for proper study because of the poor drainage from the walls of containers that are not so treated.

The samples exhibit characteristics that are not as readily distinguishable as those seen with the low solids content suspensions. Phase separation is not the primary criterion for evaluating the physical performance of the product as it is with the low solids suspensions.

The following is observed if the products are shaken after standing for about one week at room temperature. Samples A, B, and E are difficult to redisperse; samples C, D, F, and G are easily redispersed. If samples are stored for longer periods of time at room temperature, samples A to D show colour formation, while E to G do not show colour formation (or show it only slightly), owing to the antioxidant effect. Since the product is normally refrigerated during storage, the antioxidant is an added safeguard against deterioration. A good test of the influence of preservatives on the products involves drawing 5 ml of the preparations into a hypodermic syringe fitted with a 22-gauge needle, and trying to eject the contents of the syringe. Samples A, B, and E are difficult to eject, whereas the aggregated samples C, D, F, and G, because of their structure, are more easily emptied, and as indicated before, are generally redispersible.

**Table 18.4:** High solids content suspensions

| Sample | Concentration in mg/ml | | | | | | |
|---|---|---|---|---|---|---|---|
| | A | B | C | D | E | F | G |
| Penicillin G procaine,* USP | 200,000† | 200,000 | 200,000 | 200,000 | 200,000 | 200,000 | 200,000 |
| Procaine hydrochloride, USP | 20.0 | 20.0 | 20.0 | 20.0 | 20.0 | 20.0 | 20.0 |
| Sodium citrate, USP (buffer) | 20.0 | 20.0 | 20.0 | 20.0 | 20.0 | 20.0 | 20.0 |
| Lecithin (protective) colloid) | 2.0 | 2.0 | 2.0 | 2.0 | 2.0 | 2.0 | 2.0 |
| Sodium formaldehyde sulfoxylate, USP (antioxidant) | — | — | — | — | 2.5 | 2.5 | 2.5 |
| Benzyl alcohol, NF (preservative) | — | 5.0 | — | — | 5.0 | — | — |
| Butylparaben, NF (preservative) | — | — | 0.15 | — | — | 0.15 | — |
| Methylparaben/propyl-paraben, NF (preservatives) | — | — | — | 1.0/0.1 | — | — | 1.0/0.1 |
| Purified water q.s. to make 1.0 ml | | | | | | | |

*Micronized, sterile.

†Based on 1000 units/mg.

## Antacid Suspensions

Antacids constitute a single class of drugs available in both suspension and tablet forms; consumers prefer the suspension form. This is due to the fact that in vivo effectiveness of a well-formulated antacid suspension is superior to that of its tablet counterpart (detailed discussion is in the next section). Therefore, antacid formulations merit special treatment.

Aqueous aluminum hydroxide (but not magnesium hydroxide) antacid suspensions tend to thicken or gel during their shelf-life. This gelling accelerates during storage under warm conditions (30 to 40°C). Dramatic thickening is observed in the case of high-potency antacids containing large amounts of aluminum hydroxide gel. A patent by Afford teaches how to circumvent this problem by the addition of a hexitol (sorbitol or mannitol) in concentrations from 0.5 to 7%, depending on the concentration of aluminum hydroxide in the suspension. This gelling can also be prevented by the addition of 0.1 to 0.5% (w/v) potassium or sodium citrate, the former of which is preferred because of consumer demand for low-sodium antacids, particularly with those of higher potency. The gel-preventing action of the citrate ions may be analogous to the mechanism of action of monobasic potassium phosphate on the positively-charged bismuth subnitrate suspension. Aluminum hydroxide particles have an excess of positive charge because of the surrounding $Al^{3+}$ ions. With the addition of potassium citrate to the aluminum hydroxide gel-type antacids, the apparent zeta potential may be decreased to a point at which the system exhibits a maximum aggregation with a resultant thinning effect. The following formula demonstrates a hexitol-stabilized antacid system.

| Ingredients | Percentage in formula |
|---|---|
| Aluminum hydroxide gel, AHLT-LW | 36.000 |
| Sorbitol, NF, or mannitol, USP | 7.000 |
| Methylparaben, NF | 0.200 |
| Propylparaben, NF | 0.020 |
| Saccharin, NF | 0.050 |
| Peppermint oil, NF | 0.005 |

| | |
|---|---|
| Alcohol, USP | 1.000 |
| Purified water, USP q.s. to make | 100.000 |

The formula can be prepared by dissolving methylparaben, propylparaben, saccharin, and peppermint oil in the alcohol and transferring the solution, with agitation, to a vessel containing nearly one-half of the volume of purified water. The aluminum hydroxide gel, AHLT-LW, is added and then dispersed using a high-speed propeller or other high-speed disperser. To demonstrate the effect of sorbitol or mannitol, the preparation should be made with and without it. Each product is then stored in bottles at ambient temperature and at 40°C. The addition of sucrose, dextrose, propylene glycol, glycerin, or polyethylene glycol 400 in place of sorbitol or mannitol does not prevent gelling.

Aluminum hydroxide has a constipating effect. Therefore, it is normally combined with the laxative effect of magnesium hydroxide in commercial antacid formulations.

These formulas can be manufactured by dissolving the methylparaben, propylparaben, saccharin, and peppermint oil in alcohol and then transferring the mixture, with agitation, to a vessel containing nearly one-third of the volume of purified water in which either the potassium citrate or the sorbitol solution has been dissolved. The aluminum hydroxide gel, is added along with the magnesium hydroxide paste, and the mixture is agitated with a high-speed disperser. Alternately, the entire product can be passed through a colloid mill. The final volume is made up with purified water and a neutral gum-type suspending agent may be included to reduce the separation.

The taste of an antacid must be considered for consumer acceptance. Potassium citrate or sorbitol solution are included to prevent gelling; however, potassium citrate has its own unpleasant taste, while the sorbitol has a cool sweet taste that is pleasant. The parabens at the concentrations given previously impart a numbing aftertaste to the tongue, especially in the presence of the peppermint flavour. The paraben concentrations may be reduced to some extent provided the following factors are considered. The pH of the antacid product is around 8, and the pKa of the parabens is approximately 8. Thus, 50% of the parabens are in the inactive ionized form. The concentrations can be somewhat reduced, however. This can be

accomplished by either including an oxidizing-type preservative with a decomposing half-life of about two weeks or by pasteurizing the final bottled product.

# EMULSIONS (LIQUID-LIQUID SYSTEM)

A precise definition of the term *emulsion* depends on the observer's point-of-view. The physical chemist defines an emulsion as a thermodynamically unstable mixture of two essentially immiscible liquids. For the product development technologist, it is more useful to regard an emulsion as an intimate mixture of two immiscible liquids that exhibits an acceptable shelf life near room temperature.

| **Ingredients** | **Percentage in formula** | |
|---|---|---|
| | **A** | **B** |
| Aluminum hydroxide gel, AHLT-LW | 23.330 | 28.750 |
| Magnesium hydroxide (Hydro-magma) paste | 13.110 | 16.400 |
| Sorbitol solution, (70%) USP | — | 10.000 |
| Potassium citrate, USP | 0.600 | — |
| Methylparaben, NF | 0.200 | 0.200 |
| Propylparaben, NF | 0.020 | 0.020 |
| Saccharin, NF | 0.100 | 0.050 |
| Peppermint oil, NF (or other flavour) | 0.005 | 0.005 |
| Alcohol, USP | 1.000 | 1.000 |
| Purified water, USP q.s. to make | 100.000 | 100.000 |

Hydro-magma is a trademark of Merck and Co. Inc., 126 East Lincoln Avenue, Rahway, NJ 07065.

When two immiscible liquids are mechanically agitated, both phases initially tend to form droplets. When the agitation is stopped, the droplets quickly coalesce, and the two liquids separate. The lifetime of the droplets is materially increased if an *emulsifier* is added to the two immiscible liquids. Usually, only one phase persists in droplet form for a prolonged period of time. This phase is called the *internal (disperse* or *discontinuous) phase*, and

is surrounded by an external (continuous) phase. An assembly of close-packed monodisperse spherical droplets as the internal phase can occupy no more than approximately 74% of the total volume of an emulsion. It is evident, however, that the internal phase can exceed 74% if the spherical particles are not mono-disperse (as in most emulsions). A further increase in the ratio of internal and external phases can result if the internal phase is assumed to consist of polyhedral particles rather than spheres.

An emulsifier functions and is operationally defined as a stabilizer of the droplet form (globules) of the internal phase. On the basis of their structure, emulsifiers (wetting agents or surfactants) may be described as molecules comprising both hydrophilic (oleophobic) and hydrophobic (oleophilic) portions. For this reason, misgroup of compounds is frequently called *amphiphilic* (i.e. water- and oil-loving).

It is almost universally accepted that the term emulsion should be limited to *liquid-inliquid* systems. Emulsions are normally formed by “mixing” two immiscible liquids. If necessary, the two phases are heated to ensure that they are liquids during emulsification. The most common types of pharmaceutical or cosmetic emulsions include water as one of the phases and an oil or lipid as the other. If the oil droplets are dispersed in a continuous aqueous phase, the emulsion is termed *oil-in-water (o/w);* if the oil is the continuous phase, the emulsion is of the *water-in-oil* type (*w/o*). It has been observed that o/w emulsions occasionally change into w/o, emulsions and vice versa. This change of *emulsion type* is called inversion.

Since approximately 1978, two additional types of emulsions, classified as *multiple emulsions*, received the attention of surface chemists. It is entirely feasible to prepare a multiple emulsion with the characteristics of *oil-in-water-in-oil* (o/w/o) or of *water-in-oil-inwater (w/o/w)* emulsions. Such emulsions also can invert; however, during inversion they usually form “simple” emulsions. Thus, a w/o/w emulsion normally yields an o/w emulsion.

The particle size of the disperse phase determines the appearance of an emulsion. The radius of the emulsified droplets in an opaque, usually white, emulsion ranges from 0.25 to 10 μm. It is fairly well established that dispersed particles having a diameter of less than ¼th the wave length of visible light, i.e. less than approximately 120 nm, do not refract light, and therefore appear transparent to the eye. Dispersions of a liquid to such small

particle sizes yield *microemulsions* or *micellar emulsions*. Often, these terms are erroneously used interchangeably because such emulsions appear transparent to the human eye in daylight. In a microemulsion, disperse globules having a radius below the range of 10 to 75 nm are present.

The production of a transparent dispersion of an oil by micellization does not result in the formation of droplets, but in the inclusion of the lipid into micelles, which may, but need not, possess spherical shapes. In terms of size, micelles have dimensions ranging from about 5 to 20 nm. To the practicing technologist, transparent emulsions, solubilized oils, micellar emulsions, and microemulsions are one and the same because they appear clear. However, solubilization in any form represents an entirely different phenomenon, from that of emulsification.

## APPLICATION AND UTILITY

Emulsions are sometimes difficult to prepare and require special processing techniques. To warrant this type of effort and to exist as useful dosage forms, emulsions must possess desirable attributes and cause a minimum of associated problems. The “mixing” of immiscible liquids for various purposes has been met by the emulsification process for centuries. Today, emulsions continue to have a variety of cosmetic and pharmaceutical applications. The latter may be further be classified by the route of administration, i.e. topical, oral, or paren-teral. In principle, cosmetic applications and topical pharmaceutical applications are similar and together form one of the most important groups of emulsions.

i. Patient acceptance undoubtedly is the most important reason why emulsions are popular oral and topical dosage forms. Many medicinal agents have an objectionable taste or texture, and can be made more palatable for oral administration when formulated into emulsions. As a result, mineral oil-based laxatives, oil-soluble vitamins, and high-fat nutritive preparations are commonly administered as o/w emulsions.

ii. The utility of orally-administered emulsions resides in their efficacy, i.e. absorption or bioavailability of the drug. It has been demonstrated that some drugs are more readily absorbed when they are administered orally in the form of emulsions. It has even been reported that normally unabsorbable macromolecules, such as insulin and heparin, are absorbed when they are incorporated into emulsions.

iii. Patient acceptance is also important in topically-applied emulsions. Emulsions possess a certain degree of elegance and are easily washed off whenever desired. In addition, the formulator can control the viscosity, appearance, and degree of greasiness of cosmetic or dermatologic emulsions.

iv. With regard to emulsion type, o/w emulsions are most useful as water-washable drug bases and for general cosmetic purposes. W/o emulsions are employed more widely for the treatment of dry skin and emollient applications. The utility of topical emulsions depends on their ability to “penetrate”. This much abused term has entirely different meanings to the layman and to the technologist. To the former, rapid “penetration” is desirable and refers to the disappearance of the product or of oiliness

from the skin during injunction. It is generally believed that this process of penetration into the skin is facilitated if the emulsion is thixotropic, i.e. if it becomes less viscous during shearing. To the technologist, penetration of the vehicle is of secondary importance; instead, rapid and efficient penetration of the drug moiety to the site that needs to be treated is desired.

v. Emulsions have been used for the intravenous administration of lipid nutrients, which is facilitated by emulsification and probably would be impossible unless the lipid were in the form of an emulsion. Such emulsions of the o/w type require most rigorous control of the emulsifying agent and/or particle size.

Some other pharmaceutical and clinical applications of emulsions include the following:

vi. Radiopaque emulsions have been used as diagnostic agents in X-ray examinations.

vii. W/o emulsions have been employed to disperse water-soluble antigenic materials in mineral oil for intramuscular depot injection. The presence of emulsifiers in injectable drugs that are relatively insoluble in water (or serum) may help to lower the tendency of the drug to crystallize and cause thrombophlebitis.

viii. Emulsification of perfluorinated hydrocarbons is required to make them useful as oxygen carriers in blood replacements.

ix. Emulsions also possess an important cost advantage over single-phase preparations. Most lipids and solvents for lipids that are intended for application to or into the human body are relatively costly. As a result, dilution with a safe and inexpensive diluent, such as water, is highly desirable from an economic point-of-view as long as efficacy or performance is not impaired.

## THEORY OF EMULSIFICATION

When oil and water are mixed and agitated, droplets of varying sizes are produced. A tension exists at the interface because the two immiscible phases tend to have different attractive forces for a molecule at the interface. A molecule of phase A is attracted into phase A and repelled by phase B. In general, the greater the degree of immiscibility, the greater is the interfacial tension. For example, liquid hydrocarbons, such as those found in mineral oil, exhibit an interfacial tension against water of approximately 50 dynes/cm, whereas a more polar vegetable oil, such as olive oil, exhibits a value of 23 dynes/cm. The interfacial tension at a liquid interface is defined as the work required to create 1 $cm^2$ of the new interface.

A fine dispersion of oil and water necessitates a large area of interfacial contact, and its production requires an amount of work equal to the product of interfacial tension; and the area change. Thermodynamically speaking, this work is the interfacial free energy imparted to the system. A high interfacial free energy favours a reduction of interfacial area, first by causing droplets to assume a spherical shape (minimum surface area for a given volume) and then by causing them to coalesce (with a resultant decrease in the number of droplets). This is the reason for including the words “thermodynamically unstable” in the classic definition of opaque emulsions.

## DROPLET STABILIZATION

Two conceptual alternatives exist for creating opaque, i.e. milky-appearing, emulsions. Such dispersions can be formed and stabilized by lowering the interfacial tension and/or by preventing the coalescing of droplets. According to classic emulsion theory, emulsifying agents are capable of performing both objectives. The materials commonly used as emulsifying agents can be divided into three categories: surface-active, hydrophilic colloids, and finely-divided solids. They reduce the interfacial tension, and act as barriers to droplet coalescence since they are adsorbed at the interface, or more precisely, on the surface of the suspended droplets. Emulsifying agents assist in the formation of emulsions by three mechanisms:

1. Reduction of interfacial tension—thermodynamic stabilization.
2. Interfacial film formation—mechanical barrier to coalescence.
3. Electrical repulsion—electrical barrier to approach of particles.

## Reduction of Interfacial Tension

The adsorption of a surfactant lowers the interfacial tension between two liquids. A reduction in attractive forces of dispersed liquid for its own molecules lowers the interfacial free energy of the system and prevents coalescence or phase separation. Even though the reduction of interfacial tension lowers the interfacial free energy produced on dispersion, it is the role of emulsifying agents as interfacial barriers that is most important. This can be seen clearly when one considers that many polymers and finely-divided solids, not efficient in reducing interfacial tension, form excellent interfacial barriers, act to prevent coalescence, and are useful as emulsifying agents.

## Interfacial Film Formation

It could be considered as an extended interfacial tension theory, in which the adsorbed emulsifier at the interface surrounds the dispersed droplets forming a coherent monomolecular or multimolecular film, which prevents the coalescence, as the droplets approach each other. The stability of the emulsions depends on the characteristics of the film formed at the interface which in turn depends upon the type of emulsifier.

### *Surface Active Agents-monomo/ecu/ar Film Formation*

The formation of films by a surfactant on the surface of water or oil droplets has been studied in great detail. It is reasonable to expect an amphiphilic molecule to align itself at a water-oil interface in the most energetically—favourable position—oleophilic portion in the oil phase and hydrophilic portion in the aqueous phase. It is also well-established that the surface-active agents tend to concentrate at interfaces and that emulsifiers are adsorbed at oil-water interfaces as monomolecular films (Fig. 18.15). These monomolecular films formed at the interface depend upon the nature, characteristics and concentration and combination of surfactant(s).

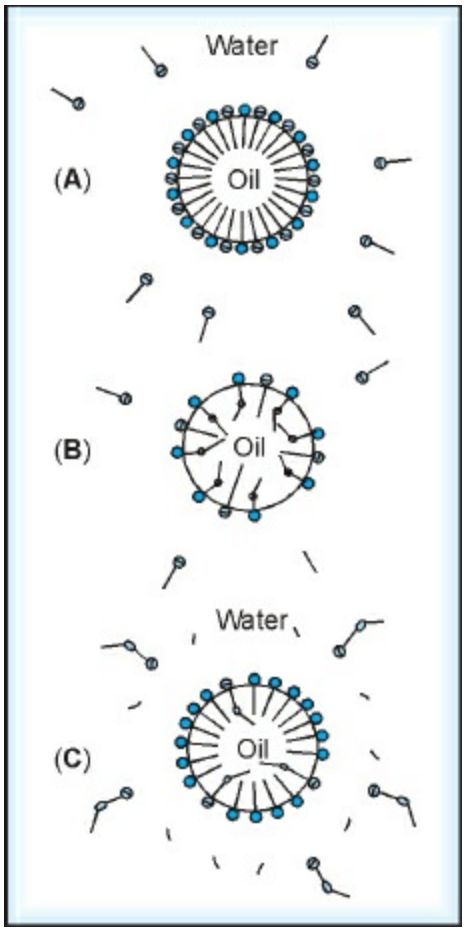



**Fig. 18.15:** Schematic representation of the relationship between mixed film formation, mechanical strength, and the stability of emulsions: (A) ⊖ Cetyl sulfate Na; • cholesterol; closely packed condensed complex; excellent emulsion; (B) ⊖ Cetyl sulfate Na; • oleyl alcohol; no closely packed condensed complex; poor emulsion; (C) O Cetyl alcohol; ⊖ sodium oleate; fairly closely packed monolayer; negligible complex formation—rather poor emulsion

**Gaseous films:** In gaseous films, the adsorbed surfactant molecules separate, do not adhere to each other laterally, and move freely around the interface. One example of a gaseous film is that formed by anionic surfactant, sodium dodecyl sulfate. The charged sulfate head groups repel one another in the aqueous solution as the droplet covered with the film moves closer to another. When the film is strongly anchored to the dispersed phase droplet, the emulsion is stable. If the monolayer film is loosely fixed, the adsorbed molecules move away from the interface and coalescence occurs.

**Condensed films:** If the concentration of the emulsifier is high enough, it forms a rigid film between the immiscible phases, which acts as a mechanical bar to both adhesion and coalescence of the emulsion droplets. The molecules of the long straight-chain fatty acids, such as palmitic and stearic acids, are more tightly packed and the film steeply rises from the compression. The hydrocarbon chains are adjacent to and in cohesive contact with one another. As the chains interlock, the molecules do not freely move in the interface leading to a stable emulsion.

**Expanded films:** The films formed by oleic acid are more expanded films than those of palmitic and stearic acids. The hydrocarbon chains in oleic acid are less cohesive and less orderly packed in the liquid than those in stearic acid. The unsaturated double bond is polar and has a greater affinity for water. The presence of branched and bent-shaped hydrocarbon chains, bulky head groups and multiple polar groups causes lateral cohesion to be reduced and expanded films to form. Non-ionic surfactants produce the same interfacial films in a similar fashion as these mentioned above. As expected, there is no charge repulsion contribution, however, the polar polyoxyethylene groups of the surfactants are hydrated and bulky, causing steric hindrance among droplets and preventing coalescence.

**Interfacial complex condensed films:** Measurements of the area occupied by a single molecule of surface-active agent at the interface of emulsion droplets have shown that in stable emulsions, the molecules of surface-active agents are in fact closely packed and form a tough interfacial film. To improve stability, the combinations of surfactants are often used rather than a single surfactant. Combination of a water-soluble surfactant that produces a gaseous film, and an oil-soluble auxiliary surfactant produces a stable interfacial complex condensed film. This film is flexible, highly viscous, coherent, elastic, and resistant to rupture since the molecules are efficiently packed between each other.

This concept is illustrated by the classic study of Schulman and Cockbain. They found that o/w emulsions stabilized by mixtures of sodium cetyl sulfate and cholesterol, which were known from other experiments to form rigid and tightly-packed films, were extremely stable (scheme A, Fig. 18.15). When oleyl alcohol was substituted for cholesterol, however, the steric effect of a double bond (which produces a kink in the carbon chain) resulted in the formation of a poor interfacial complex, and emulsion stability

was proportionately low (scheme B, Fig. 18.15). On the other hand, if this last system was changed somewhat by using sodium oleate and cetyl alcohol, it was possible to obtain a rather poor emulsion, but one that was more stable than that of the previous case (scheme C, Fig. 18.15). Thus, a tightly-packed emulsifier film contributes to the stability of the emulsion. This steric argument by Schulman and Cockbain has been used to explain the well-known fact that mixed emulsifiers are often more effective than single emulsifiers. The ability of the mixture of emulsifiers to pack more tightly contributes to the strength of the film, and hence, to the stability of the emulsion. Most emulsifiers probably form fairly dense gel structures at the interface and produce a stable interfacial film.

**Lamellar liquid crystalline films:** Recent studies have helped to clarify further the nature of these interfacial films. Stable emulsions are now believed to comprise liquid crystalline layers on the interface of emulsified droplets with the continuous phase. In their pioneering studies, Friberg and co-workers were able to show by optical (polarized light) and electron microscopy and low-angle X-ray diffractometry that mixed emulsifiers can interact with water to form three-dimensional association structures. The classic concept of emulsions as two-phase systems with a monomolecular layer of emulsifier at the interface must be revised. Emulsions should instead be viewed as three-component systems comprising oil, “water”, and lamellar liquid crystals, the latter consisting of consecutive layers of water-emulsifier-oil-water (Fig. 18.16).

Interlamellar layers representing the internal and external phases of an emulsion have recently been identified by freeze fracture micrography of o/w creams. In addition, it has recently been learned that emulsion droplets can be surrounded by liquid crystals of a closed lamellar type in appropriately prepared emulsions.

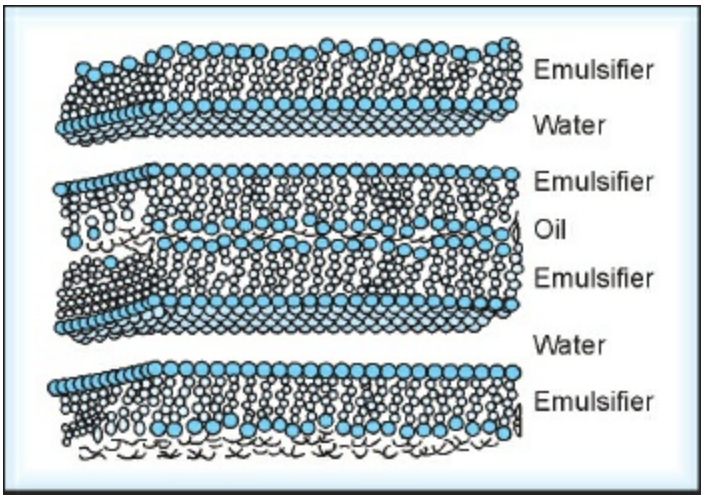



**Fig. 18.16:** A lamellar liquid crystal consists of consecutive layers of water-emulsifier-oil-emulsifier-water

## *Hydrophilic Colloids-multimolecular Film Formation*

Hydrophilic colloids such as polysaccharides and proteins do not lower appreciably the interfacial tension but form a multimolecular film at the oil-water interface (Fig. 18.17). The multimolecular films are strong and elastic and give mechanical protection to coalescence. An additional effect of these hydrophilic colloids is the electrostatic charge repulsion due to the carboxylic acid groups of polysaccharides and amino acid groups of proteins. Hydrophilic colloids are used to stabilize emulsions and form o/w-type emulsions.

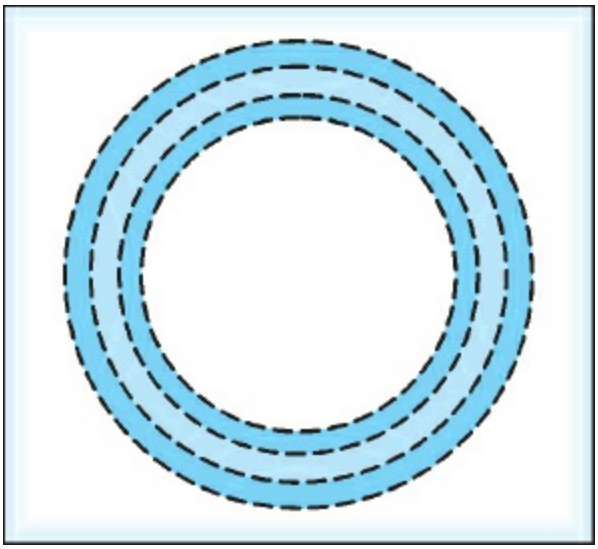


**Fig. 18.17:** Multimolecular film formed by hydrophilic colloids

## *Finely Divided So/ids-so/id Particle Film Formation*

Finely divided solid particles are lodged at the interface, adhere strongly to each other, forming a stable film at the surface. They form stable emulsions by preferentially wetted by one of the phases (Fig. 18.18). When wetted by water, the contact angle is less than 90°, and o/w-type emulsions are formed, while when wetted by oil, w/o-type emulsions are formed.

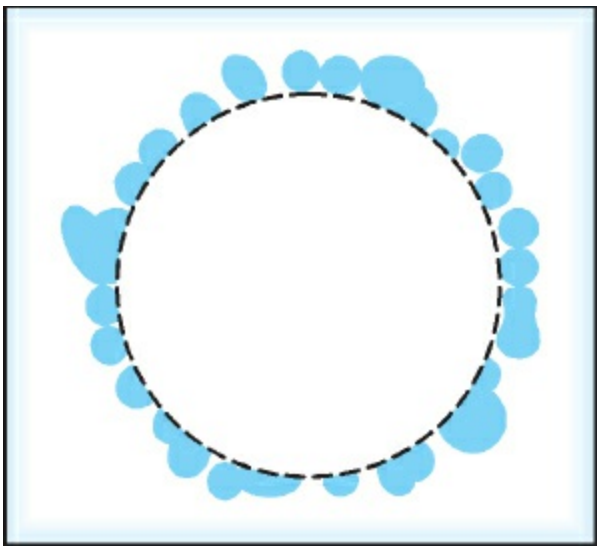


**Fig. 18.18:** Adsorption of finely divided solid particles on liquid droplets

## Electrical Repulsion

It has just been described how interfacial films or lamellar liquid crystals significantly alter the rates of coalescence of droplets by acting as barriers. In addition, the same or similar film can produce repulsive electrical forces between approaching droplets. Such repulsion is due to an electrical double layer, which may arise from electrically-charged groups oriented on the surface of emulsified globules. To simplify, let us consider the case of an o/w emulsion stabilized by a sodium soap. Not only are the molecules of this surfactant concentrated in the interface, but because of their polar nature, they are oriented as well (Fig. 18.19). The hydrocarbon tail is dissolved in the oil droplet, while the ionic heads are facing the continuous aqueous phase. As a result, the surface of the droplet is studded with charged groups, in this case negatively-charged carboxylate groups. This produces a surface charge on the droplet, while cations of opposite sign are oriented near the surface, producing what is known as the (diffuse) double-layer of charge.

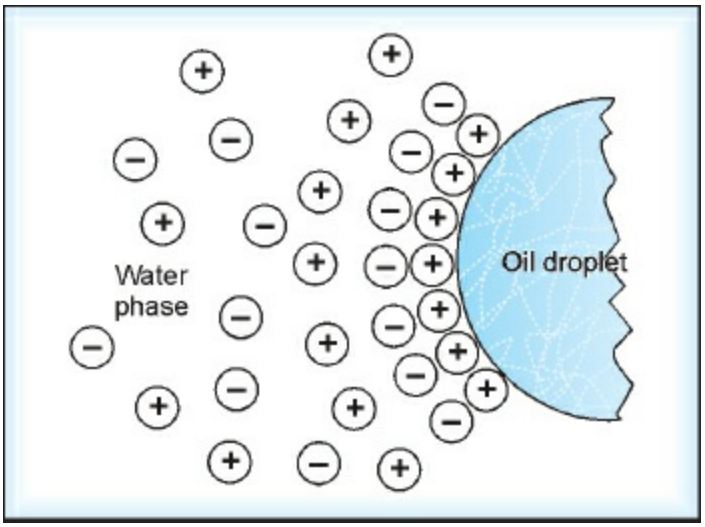



**Fig. 18.19:** Idealized representation of the electrical double-layer at an oil-water interface

The potential produced by the double layer creates a repulsive effect between the oil droplets and thus hinders coalescence. Although the repulsive electrical potential at the emulsion interface can be calculated, it cannot be measured directly for comparison with theory. The related quantity, however, zeta potential, can be determined. The zeta potential for a surfactant-

stabilized emulsion compares favourably with the calculated double-layer potential. In addition, the change in zeta potential parallels rather satisfactorily the change in double-layer potential as the electrolyte is added. These and related data on the magnitude of the potential at the interface can be used to calculated the total repulsion between oil droplets as a function of the distance between them.

## EMULSION TYPE

Only o/w and w/o emulsions have achieved commercial and practical importance. To understand the various factors that determine whether an o/w or a w/o emulsion will be produced, one must again think in terms of two critical features (1) droplet formation and (2) formation of an interfacial barrier. The phase volume ratio, i.e. the relative amount of oil and water, determines the relative number of droplets formed initially and hence the probability of collision; the greater the number of droplets, the greater is the chance for collision. Thus, normally, the phase present in greater amount becomes the external phase.

To predict the type of emulsion formed under a given set of conditions, the interaction of various parameters must be estimated. This estimation is nearly impossible, and only a few generalized and somewhat empiric rules can be given.

1. If the amphiphile is essentially water-soluble (e.g. potassium soap or polyoxyethylene alkyl ether with more than 5 ethylene oxide units), it will usually favour o/w emulsification; if the surfactant is primarily soluble in the lipid portion (calcium soap, polyoxyethylene alkyl ether with less than 5 ethylene oxide units), it may yield w/o emulsions if the other conditions are favourable.
2. The polar portions of emulsifier molecules are generally better barriers to coalescence than their hydrocarbon counterparts. It is, therefore, possible to make o/w emulsions with relatively high internal phase volumes. On the other hand, w/o emulsions (in which the barrier is of hydrocarbon nature) are limited in this regard and invert easily if the amount of water present is significant. For example, a water-mineral oil-sorbitan monooleate system, ordinarily expected to favour w/o emulsion formation because of the lack of ethylene oxide units, does so only if the amount of water present constitutes less than 40% by volume. At higher amounts of water, only o/w emulsions form.
3. Even at 20% and 30% water, w/o emulsions form only if the water is added to the oil with mixing. The addition of both phases together, followed by mixing, favours o/w emulsions at all concentrations above 10% water.

4. Finally, the type of emulsion formed is influenced to some extent by the viscosity of each phase. An increase in the viscosity of a phase aids in making that phase the external phase.

Despite these complications, one can expect a predominantly water-soluble emulsifier to form o/w emulsions, whereas the reverse is true of primarily oil-soluble surfactants—**Bancroft's Rule**. Occasionally, it is desirable to determine the type of emulsion formed. Methods for this purpose are shown in Table 18.5.

**Table 18.5:** Methods for the determination of emulsion type

| Test | Observation | Comments |
|---|---|---|
| Dilution test | Emulsion can be diluted only with external phase | Useful for liquid emulsions only |
| Dye test | Water-soluble solid dye tints only o/w emulsions and reverse Microscopic observation usually helpful | May fail if ionic emulsifiers are present |
| $CoCl_2$/filter paper | Filter paper impregnated with $CoCl_2$ and dried (blue) changes to pink when o/w emulsion is added | May fail if emulsion is unstable or breaks in the presence of electrolyte |
| Fluorescence | Since oils fluoresce under UV light, o/w emulsions exhibit dot pattern, w/o emulsions fluoresce throughout | Not always applicable |
| Conductivity | Electric current is conducted by o/w emulsions, owing to the presence of ionic species in water | Fails in non-ionic o/w emulsions |

## Microemulsions

Operationally, microemulsions may be defined as dispersions of insoluble liquids in a second liquid that appear clear and homogeneous to the naked eye. Microemulsions are frequently called *solubilized* systems because on a macroscopic basis they seem to behave as true solutions. Careful examination of these complex systems has shown that clear emulsions can exist in several differentiable forms. Microemulsions should not be confused, however, with solutions formed by cosolvency, e.g. the clear system *consisting* of water, benzene, and ethanol.

Blending of a small amount of oil with water results in a two-phase system because “water and oil do not mix”. If the same small amount of oil is added to an aqueous solution of a suitable surfactant in the micellar state, the oil may preferentially dissolve in the interior of the micelle because of its hydrophobic character. This type of micellar microemulsion, has also been called an o/w micellar solution. Similarly, w/o solubilization—especially that by a nonionic surfactant has recently been attributed to the existence of swollen micelles. In these systems, sometimes called reverse micellar solutions, water molecules are found in the polar central portion of a surfactant micelle, the non-polar portion of which is in contact with the continuous lipid phase. A third type of microemulsion (usually of the w/o type) is formed by ionic surfactants (e.g. sodium stearate) in the presence of cosurfactants (e.g. pentanol or dioxyethylene dodecyl ether) with hydrocarbons (e.g. hexadecane) and water. The pseudoternary phase representing the existence of various emulsions and micellar system is shown in Fig. 18.20.

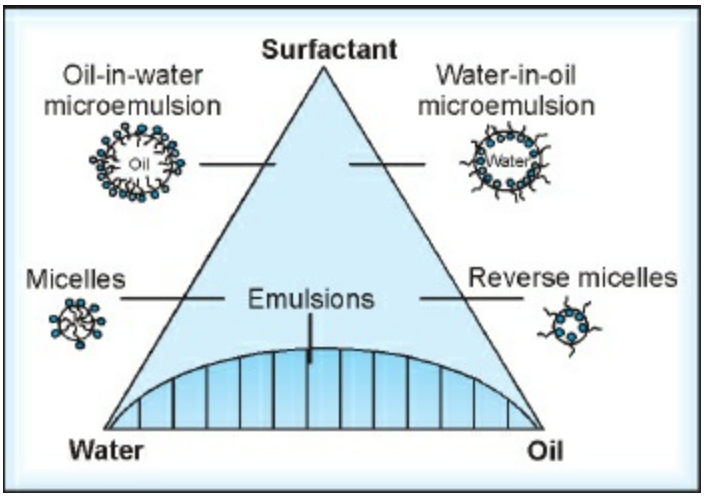



**Fig. 18.20:** Pseudo-ternary phase diagram illustrating the existence of emulsion and micellar system

In general, microemulsions or solubilized systems are believed to be thermodynamically stable. Transparent or clear emulsion in which a water-insoluble oil or drug is “dissolved” in an aqueous surfactant system play an important role in drug administration.

## FORMULATION COMPONENTS

It is difficult to designate a general approach or a set of rules for selecting the components and their amounts to yield a desired emulsion. The ingredients of any pharmaceutical or cosmetic emulsion must conform to various requirements. There are situations in which certain oils, emulsifiers, and other ingredients must be avoided or used exclusively. Usually, however, ingredient selection is made on the basis of the experience and personal tastes of the formulator and by trial and error. Formulators are cautioned to establish the safety and regulatory acceptance of emulsion ingredients for a particular application.

## Lipid Phase

The materials making up the oil portion of an emulsion and their relative amounts are determined primarily by the ultimate use of the product. For pharmaceutical and cosmetic products, the oil phase, unless it is the active ingredient, may include a wide variety of lipids of natural or synthetic origin. The consistency of these lipids may range from mobile liquids to fairly hard solids. Some of the lipids useful for pharmaceutical or cosmetic emulsions are listed in Table 18.6.

A drug in an emulsion-type of dosage form distributes itself between the oil phase and the aqueous phase in accordance with its oil/water partition coefficient. The drug's absorption by the gastrointestinal tract or the skin can be expected to depend on its solubility in the oil phase. In principle, the less soluble an active ingredient is in the non-volatile portion of the vehicle, the more readily it penetrates into and through a barrier. On the other hand, a finite solubility of the active ingredient in the vehicle is necessary to ensure its presence in a fine state of subdivision. It is generally accepted that the release of a medicinal agent from a dosage form is a function of the solubilities of the agent in the base and in the body membrane. The key point is that the drug must not be so soluble preferentially in the base that it prevents penetration or transfer.

**Table 18.6:** Ingredients for oil phase of emulsions

| Class | Identity | Consistency |
|---|---|---|
| Hydrocarbon | Mineral oils | Fluids of varying viscosity |
| Hydrocarbon | Petrolatum | Semisolid |
| Hydrocarbon | Polyethylene waxes | Solids |
| Hydrocarbon | Microcrystalline waxes | Solids |
| Ester | Vegetable oils | Fluids of varying viscosity |
| Ester | Animal fats | Fluids or solids |
| Ester | Lanolin | Semisolid |

| | | |
|---|---|---|
| Ester | Synthetics (e.g. i-propylmyristate) | Fluids |
| Alcohols | Long-chain (natural and synthetic) | Fluids or solids |
| Fatty acids | Long-chain (natural and synthetic) | Fluids or solids |
| Ethers | Polyoxypropylenes | Fluids of varying viscosity |
| Silicones | Substituted | Fluids of varying viscosity |
| Mixed | Plant waxes (e.g. Candelilla) | Solid |
| Mixed | Animal waxes (e.g. Bees) | Solid |

A final consideration in the selection of a lipid component for a topical preparation is its “feel”. Emulsions normally leave a residue of the oily components on the skin after the water has evaporated. Therefore, the tactile characteristics of the combined oil phase are of great importance in determining consumer acceptance of an emulsion.

**Phase ratio:** The ratio of the internal phase to the external phase is frequently determined by the solubility of the active ingredient, which must be present at a pharmacologically effective level. If this is not the primary consideration, the phase ratio is normally determined by the desired consistency. As a rule of thumb, it can be assumed that fluid emulsions result from low levels of the internal phase, whereas heavier emulsions are the result of higher percentages of the internal phase. Also, a high internal phase ratio normally requires a high level of emulsifying agent; this point affects the decision concerning the phase ratio.

## Emulsifying Agents

It is customary to differentiate three broad classes of emulsifying agents: the surfactants, the hydrophilic colloids, and the finely divided solids. Although hydrophilic colloids and finely divided solids can be used as the only emulsifier, their greatest utility is in the form of auxiliary emulsifiers; accordingly, they are discussed under this heading. A particular class of emulsifier is selected primarily on the basis of required "shelf-life" stability, the type of emulsion desired, and emulsifier cost.

### *Surface Active Agents or Surfactants*

Substances having both hydrophilic and hydrophobic regions in their molecular structures are called surface active agents or surfactants (Fig. 18.21). These materials are soluble in both water and oil. Upon addition of the surfactants into the dispersed system, the hydrophilic (polar) and hydrophobic (nonpolar) groups orient themselves in a monomolecular layer facing the polar (i.e. water) and non-polar (i.e. oils) solvents, respectively. Surfactants diffuse from the solution onto the interface where adsorption and accumulation take place. The interfacial tension must be lowered for the interface to expand and if the interfacial tension is decreased sufficiently, the dispersed system will readily be emulsified.

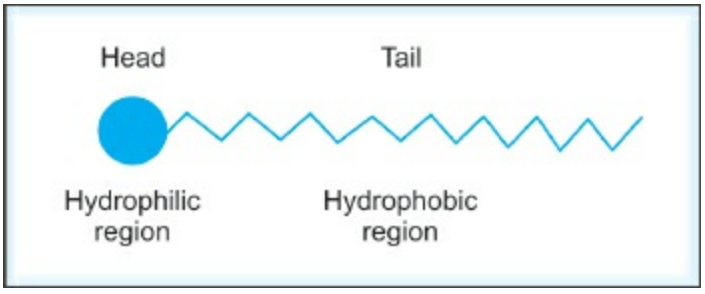



**Fig. 18.21:** Representative structure of surface active agent

As the concentration of surfactant in an aqueous solution is increased, the interfacial tension is appreciably lowered. Further addition leads to a saturated level at the surface where the surfactant molecules are closely packed. Beyond saturation the excess surfactant moves into the bulk and form micelles within the aqueous solution and there is no longer a change in the surface tension.

The number of surfactants available for the formation of emulsions is so huge that even a cursory description is impossible. Surfactants are classified into four main categories depending on the nature of the charge carried by the hydrophilic part of the surfactant: anionic, cationic, non-ionic, and ampholytic surfactants (Table 18.7).

**Table 18.7:** Classification of surfactants for pharmaceutical emulsions

| Surfactant | Typical representative | Utility | chemical structure |
|---|---|---|---|
| **Anionic surfactants** | | | |
| *Soaps* | Sodium oleate | T | $CH_3(CH_2)_7CH=CH(CH_2)_7COO^-Na^+$ |
| | Sodium palmitate | T | $CH_3(CH_2)_{14}COO^-Na^+$ |
| *Sulphates* | Sodium lauryl sulfate | TO | $CH_3(CH_2)_{11}SO_4^-Na^+$ |
| *Ether sulphates* | Sodium laureth sulphate | | $C_{12}H_{25}(O\text{-}CH_2\text{-}CH_2)_3\text{-}O\text{-}SO_3\text{-}Na^+$ |
| *Benzene sulphonate* | 4-benzyl dodecane sulphonate sodium | T | $C_3H_7CH(C_8H_{17})\text{-}C_6H_4\text{-}SO_3\text{-}Na+$ |
| *Hemiesters* | Sodium dioctyl sulfosuccinate | TO | $CH_3(CH_2)_7O_2CCH_2CH(SO_3Na)CO_2(CH_2)_7CH_3$ |
| *Sarcosides* | Lauryl sarcosinate | TO | |
| **Cationic surfactants** | | | |
| *Amines* | Tetradecyl methyl amine | T | $C_{14}H_{29}NHCH_3$ |
| *Quaternary ammoniums* | Bezalkonium chloride | TO | C8H17 (to C18H37) $N^+(CH3)2\ Cl^-\text{–}CH_2\text{–}C_6H_5$ |
| | Hexadecyltrimethyl ammonium chloride | TO | $CH_3(CH_2)_{15}\text{–}N^+(CH_3)_3\ Cl^-$ |
| **Anionic surfactants** | | | |
| | Cetrimide | TO | $C_{16}H_{33}\text{-}N^+(CH_3)_3\ Br^-$ |
| *Miscellaneous* | Cetylpyridinium chloride | T | $CH_3(CH_2)_{14}CH_2\ N^+C_5H_5Cl^-$ |
| **Nonionic surfactants** | | | |
| *Ethoxylated alcohols* | Tridecanol | TO | $O=C(NH\text{-}(O\text{-}CH_2\text{-}CH_2)_n\text{-}O\text{-}CO\text{-}R_1)(NH\text{-}(O\text{-}CH_2\text{-}CH_2)_n\text{-}O\text{-}CO\text{-}R_2)$ |
| *Ethoxylated amides* | Di-acyl ethoxy urea | | |
| *Fatty acid esters (Span 60)* | Sorbitan monostearate | TO | |
| | Polyoxyethylene sorbitan monolaurate (Polysorbate 20 or Tween 20) | TO | |
| | Glycerol trimer | TO | |
| **Ampholytic surfactants** | | | |
| *Ammonium phosphates* | Lecithin | TOP | |
| *Amino propionic acids* | | TO | |
| *Quaternary* | Betain | TO | $R\text{–}N^+(CH_3)_2\text{–}CH_2\text{–}COO^-$ |
| *compounds* | sulfobetain | TO | $R\text{–}N^+(CH_3)_2\text{–}CH_2\text{–}SO_3^-$ |

T = some representatives useful in topicals.

O = some representatives useful in oral preparations or ingested drugs.

P = some representatives useful in parenterals.

*Anionic surfactants* are negatively charged in an aqueous solution (i.e. COO), and widely used because of their cost and performance. Sodium lauryl sulfate, the main component of which is sodium dodecyl sulfate, is highly soluble in water and commonly used to form oil-in-water (o/w) emulsions. Reacting an alkali hydroxide with a fatty acid (e.g. oleic acid) can produce alkali metal soaps (e.g. sodium oleate). Careful attention must be paid to the pH of the dispersion medium and the presence of multivalent metals. Alkali earth metal soaps (e.g. calcium oleate) produce stable water-in-oil (w/o) emulsions because of their low water solubility and are produced by reacting oleic acid with calcium hydroxide. Triethanolamine stearate produces stable o/w emulsions *in-situ* by reacting triethanolamine in aqueous solution with melted stearic acid at approximately 65°C (e.g. vanishing cream).

*Cationic surfactants* are positively charged in an aqueous solution (e.g. quaternary ammonium and pyridinium), and expensive. Because of their bactericidal action, they are widely used for other applications such as preservatives, sterilizing contaminated surfaces, and emulsions.

*Non-ionic surfactants* consist of a $(CH_2CH_2O)_nOH$ or OH as the hydrophilic group and exhibit a variety of hydrophile-lipophile balances (HLB) which stabilize o/w or w/o emulsions. Unlike anionic and cationic surfactants, non-ionic surfactants are useful for oral and parenteral formulations because of their low irritation and toxicity. Based on their neutral nature, they are much less sensitive to changes in the pH of the medium and the presence of electrolytes. The best use of non-ionic surfactants is to produce an equally balanced HLB of two non-ionic surfactants: one hydrophilic and one hydrophobic. Sorbitan esters (Spans) are the products of the esterification of a sorbitan with a fatty acid. Their hydrophilicity comes from the hydroxyl groups of the saturated cyclic ring. They are not soluble in water and used for w/o-type emulsions. Polysorbates (Tweens), on the other hand, are soluble in water since a number of ethylene oxides are adducted by the hydroxyl groups of the sorbitan esters. They are hence used as emulsifying agents for o/w emulsions. In general, both sorbitan esters and polysorbates are used in conjunction to produce a wide range of emulsions. Fatty alcohol polyoxyethylene ethers are condensation products of fatty alcohols with polyethylene glycol, while fatty acid polyoxyethylene

esters are esterification products of fatty acids with polyethylene glycol. They are soluble in water and used in conjunction with auxiliary emulsifying agents (e.g. cetyl and stearyl alcohols) to give o/w emulsions.

*Ampholytic surfactants* possess both cationic and anionic groups in the same molecule and are dependent on the pH of the medium. Lecithin is used for parenteral emulsions.

## *Hydrophilic-lipophilic Balance (HLB) Concept*

Emulsion technologists have for many years selected emulsifiers from an intuitive knowledge of their hydrophilic, lipophilic behaviour and of the type of emulsion produced with a given lipid or aqueous phase. This approach is most readily illustrated with non-ionic surfactants, but the principles involved can be extrapolated to any type of emulsifier or combination of emulsifiers. It is apparent that the choice of specific emulsifiers by this method, although practical, is empiric and tedious. To systematize the hydrophilic/lipophilic approach to emulsifier selection, Griffin in 1947 developed the (still somewhat empiric) system of the HLB of surfactants. The HLB value of an emulsifier can be determined experimentally or can be computed as long as the structural formula of the surfactant is known.

The HLB values of the surfactants based on polyhydric alcohol fatty acid esters may be estimated by:

$$HLB = 20\left(1 - \frac{S}{A}\right) \quad \text{... (8)}$$

where, $S$ is the saponification number of the ester and $A$ is the acid number of the fatty acid. If one cannot obtain the saponification numbers (e.g. beeswax and lanolin derivatives), their *HLB* values may be calculated by:

$$HLB = (E + P)/5 \text{ ... (9)}$$

where, $E$ is the weight percent of oxyethylene chains in the surfactant and $P$ is the weight percent of the polyhydric alcohol groups (e.g. glycerol or sorbitol) in the material. If the hydrophilic region is polyoxyethylene, the *HLB* value is calculated by:

$$HLB = E/5 \text{ ... (10)}$$

Another useful means of finding the *HLB* of an unknown emulsifier is that proposed by Davies, which permits calculation of the *HLB* value by

algebraically adding the values assigned to a particular atomic grouping within the molecule of the emulsifier.

$$HLB \text{ of surfactant} = \Sigma(\text{hydrophilic groups}) - \Sigma(\text{hydrophobic groups}) + 7 \quad \text{... (11)}$$

Occasionally, it will be found that a single emulsifier can yield the desired type of emulsion at the desired viscosity. More often, however, especially in the case of o/w emulsions, stable emulsions can be prepared readily by utilizing a combination of a lipophilic and a hydrophilic surfactant. Such combinations appear to produce mixed interfacial phases of high surface coverage as well as of sufficient viscosity to prevent creaming and promote stability.

It is not easy to match the required *HLB* value of the oil or the oil mixture with that of a single surfactant to form the most stable emulsion. The appropriate combination of surfactants should be chosen. The *HLB* value of the mixture of surfactants *A* ($HLB_A$) and *B* ($HLB_B$) is calculated by:

$$HLB_{mxurs} = f_{AX} HLB_A + (1 - f_A) \times HLB_B \quad \text{... (12)}$$

where, $f_A$ is the weight fraction of surfactant *A* in the mixture.

In general, molecules that are oil-soluble or oil-dispersible have low *HLB* values; those that are water-soluble have high *HLB* values. The simplified classification of *HLB* values based on their dispersibility in water is included in Table 18.8, whereas classification based on their application is included in Table 18.9.

**Table 18.8:** *HLB* values of selected emulsifiers

| Chemical designation | HLB | Water dispersibility |
|---|---|---|
| Ethylene glycol distearate | 1.5 | No dispersion |
| Sorbitan tristearate | 2.1 | |
| Propylene glycol monostearate | 3.4 | |
| Sorbitan sesquioleate | 3.7 | |
| Glyceryl monostearate, non-self-emulsifying | 3.8 | Poor dispersion |
| Propylene glycol monolaurate | 4.5 | |
| Sorbitan monostearate | 4.7 | |
| Diethylene glycol monostearate | 4.7 | |
| Glyceryl monostearate, self-emulsifying | 5.5 | |
| Diethylene glycol monolaurate | 6.1 | Milky dispersion (not stable) |
| Sorbitan monopalmitate | 6.7 | |
| Sucrose dioleate | 7.1 | |
| Polythylene glycol (200) monooleate | 8.0 | |
| Sorbitan monolaurate | 8.6 | |
| Polyoxyethylene (4) lauryl ether | 9.5 | Milky dispersion (stable) |
| Polyoxyethylene (4) sorbitan monostearate | 9.6 | |
| Polyoxyethylene (6) cetyl ether | 10.3 | |
| Polyoxyethylene (20) sorbitan tristearate | 10.5 | Translucent to clear dispersion |
| Polyxyethylene glycol (400) monooleate | 11.4 | |
| Polyoxyethylene glycol (400) monostearate | 11.6 | |
| Polyoxyethylene (9) nonyl phenol | 13.0 | |
| Polyethylene glycol (400) monolaurate | 13.1 | Clear solution |
| Polyethylene glycol (4) sorbitan monolaurate | 13.3 | |
| Polyethylene glycol (20) sorbitan monooleate | 15.0 | |
| Polyethylene glycol (20) oleyl ether | 15.4 | |
| Polyethylene glycol (20) sorbitan monopalmitate | 15.6 | |
| Polyethylene glycol (20) cetyl ether | 15.7 | |
| Polyethylene glycol (40) stearate | 16.9 | |
| Sodium oleate | 18.0 | |
| Polyethylene glycol (100) stearate | 18.8 | |
| Potassium oleate | 20.0 | |
| Sodium lauryl sulfate | Approx. 40 | |

**Table 18.9:** Application of surfactants based on HLB value

| HLB | Application |
|---|---|
| 1–3 | Antifoaming agents |
| 3–6 | w/o emulsifying agents |
| 7–9 | Wetting agents |
| 8–16 | o/w emulsifying agents |
| 13–15 | Detergents |
| 15–18 | Solubilizing agents |

The *HLB* required for emulsifying a particular oil in water can be determined by trial and error, i.e. by preparing appropriate emulsions with emulsifiers having a range of *HLB* values and then determining that *HLB* value that yields the "best emulsion". Although the numbers have been derived empirically, they are useful starting points for the preparation of a variety of emulsions. A list of required *HLB* values for lipids that are of interest in pharmaceutical preparations is shown in Table 18.10. The knowledge of the required *HLB* permits the selection of an emulsifier or a combination of emulsifiers that will produce the required *HLB*.

## *Required HLB*

*HLB* values of combinations may be determined by taking weighted averages of the individual surfactant *HLB* values. For example, Table 18.10 indicates that a w/o emulsion of lanolin requires an *HLB* of about 8. Thus, a 68:32 mixture of sorbitan monostearate (*HLB* 4.7) and polyoxyethylene (20) sorbitan monooleate (*HLB* 15.0) could be used to yield an emulsifier exhibiting an average *HLB* value of about 8.0. In fact, almost any *HLB* can be obtained by appropriate blending of emulsifiers, with the additional advantage, in most cases, of greater efficiency at lower concentrations. However, optimal emulsion stability and desirable rheologic characteristics cannot be achieved by random blending of emulsifiers. Shinoda and co-workers point out that the blending of two emulsifiers of very high and very low *HLB* to achieve an intermediate *HLB* can result in an unstable emulsion unless an emulsifier of medium *HLB* is included. Therefore, a 65:35 blend of sorbitan tristearate and polyoxyethylene (100) stearate, which has an *HLB* of about 8, could be expected to be inferior for the indicated emulsification to the blend of sorbitan monostearate and polyoxyethylene (20) monooleate.

**Table 18.10:** HLB values required by commonly used lipids

| | **O/w emulsion (Fluid)** | **W/o emulsion (Fluid)** |
|---|---|---|
| Cetyl alcohol | 15 | — |
| Stearyl alcohol | 14 | — |
| Stearic acid | 15 | — |
| Lanolin, anhydrous | 10 | 8 |

| | | |
|---|---|---|
| Mineral oil, light and heavy | 12 | — |
| Cottonseed oil | 10 | 5 |
| Petrolatum | 12 | 5 |
| Beeswax | 12 | 4 |
| Paraffin wax | 11 | 4 |

## *Phase Inversion Temperature (PIT)*

The practical importance of emulsification temperature on emulsion stability has been known to formulators for many years. As a rule, maximum particle size reduction occurs at or near the PIT. At that temperature, surfactants that are normally water-soluble may actually become soluble in the oil phase. As the emulsion cools, emulsifiers migrate, e.g. by changing their location from the internal to the external phase of the emulsions. How this alters emulsion formation, particle size, and stability has not been rigorously studied; however, some of Lin's data in Table 18.11 illustrate these points. In this simple system of one single emulsifier at *HLB* 9.9, complete placement of the surfactant into the oil phase seems most advantageous regardless of the temperature. However, the emulsification will be successful regardless of the location of the emulsifier as long as the temperature exceeds the PIT during emulsification.

**Table 18.11:** Effect of surfactant lotion and emulsification temperature on droplet size

| | Average droplet size (μm) at emulsification temp. | | | |
|---|---|---|---|---|
| Percentage of emulsifier in oil phase* | 30°C | 40°C | 50°C | 60°C |
| 0 | 15.0 | 13.0 | 11.0 | 0.2 |
| 40 | 13.0 | 10.0 | 9.0 | 0.2 |
| 80 | 2.0 | 2.0 | 1.5 | 0.2 |
| 100 | 1.5 | 1.5 | 0.9 | 0.2 |

* The emulsion consists of 30% mineral oil, 5% polyoxyethylene (5) oleyl ether, and 65% water. The emulsifier was added to one or both phases before emulsification at the indicated temperatures.

## *Determination of Surfactant Amount*

This discussion of the selection of emulsifiers would be incomplete without a brief examination of how one can determine that surfactant mixture of which the least amount is required for optimal stability of an emulsion. Often, this

goal can be achieved by determining the amount of water that can be solubilized in a given oil-plus-surfactant(s) mixture under carefully—controlled temperature and stirring conditions. For this purpose, 10 g of the lipid/surfactant mixture is weighed into a 68 ml capacity square glass vial. After equilibration at a temperature at which this (not always homogeneous) system is fluid, water is added in 0.10 ml increments. The mixture is shaken and allowed to stand at the equilibration temperature until all air bubbles have escaped. The addition of water (in 0.10 ml increments) is continued until the system remains permanently turbid. If the initial lipid/surfactant mixture is not clear, it will usually become clear upon the addition of water and then become cloudy again upon continued addition of water. This second *cloudpoint* is the end of the titration. As a rule, the most stable o/w emulsion with the finest particle size results at that surfactant/oil ratio that can tolerate the largest quantity of water and still remain clear.

## *HLB System Shortcomings*

Emulsion specialists generally agree that the *HLB* system is useful and that it may be used judiciously and with caution. It is a dictum of the *HLB* concept that the *HLB* value is critical, but this is not always the case. There is no assurance that a stable emulsion prepared from one chemical class of emulsifiers at a particular *HLB* can be duplicated by another class of emulsifiers exhibiting the same *HLB*. Thus, marked differences in emulsion type, viscosity, and time for phase separation were noted when polyoxyethylene ether derivatives were compared to polyoxyethylene ester-type surfactants having the same *HLB* and concentration. Additional complications arise from the observations that the *HLB* required for a particular emulsion to some extent depends on the phase ratio and the salt content.

- Several improvements in the classical *HLB* system for the selection of emulsifiers have been proposed. The phenol index, developed by Marszall, makes it possible to determine the “effective” *HLB* of non-ionics as a function of their concentration and in the presence of additives such as alcohols and glycols.
- Griffin originally proposed that emulsifiers with *HLB* values ranging from 3.5 to 6.0 should be used for w/o emulsions, but Ford and Furmidge showed that correlation between emulsion type and *HLB* is far from

perfect. Stable mineral oil-in-water emulsions have been obtained with a combination of nonionic ethers having an *HLB* value as low as 1.9. *HLB* may be one consideration in the preparation of a stable emulsion; another is the solubility of the emulsifier's lipid chain in the oil phase.

- Most of our current knowledge of the selection of emulsifiers is based on the *HLB* concept and is applied to the commonly used nonionic surfactants. Recently, the optimization of the stability of a parenteral, ultrasonically—emulsified nutrient oil preparation stabilized with various phospholipids and a nonionic has also been explained on the basis of *HLB*. A somewhat different interpretation of the emulsifying efficacy of phospholipids, which does not involve *HLB*, was offered by Rydhag. She reported that the best soya bean oil-in-water emulsions can be obtained with mixed (commercial) phospholipids containing the largest amounts of negatively-charged phospholipids (i.e. phosphatidyl inositol, phosphatidic acid, and phosphatidyl serine). Pure phosphatidyl choline (or its combination with phosphatidyl ethanolamine) yields the least stable emulsion. These findings are best explained by the phase-forming ability of these emulsifiers as postulated by Friberg.

## Auxiliary Emulsifiers

### *Hydrophilic Colloids*

Polymers that are water-sensitive (swellable or soluble) have some utility as primary emulsifiers; however, their major use is as auxiliary emulsifiers and as thickening agents. Natural and synthetic clays of the smectite or amphibole groups are commonly used for building the viscosity of emulsions or for suspending solids, such as pigments, in makeup preparations. A large variety of natural and synthetic clays is available, and the selection of a useful clay is occasionally difficult. The most commonly used clays, bentonites, are derived from montmorillonite, a typical smectite clay. These swell in the presence of water but raise the viscosity of aqueous media only at pH 6 or higher. Clays derived from the amphibole group, such as attapulgite, thicken not by swelling but primarily because of particle anisotropy, which interferes with the formation of a compact sediment.

The naturally-occurring gums and synthetic hydrophilic polymers listed in Table 18.12 are useful as emulsifiers and as emulsion stabilizers. Most natural hydrocolloids are polysaccharides, and their chemistry is extremely complex. These gums exhibit some type of incompatibility or instability depending on the presence of various cations, on pH, or on a second hydrophilic polymer. Some of the most useful synthetic hydrocolloids are ethers derived from cellulose. Among the completely synthetic group of polymers, the carboxyl vinyl polymers deserve special mention. Their outstanding characteristic is their ability to impart a yield value to aqueous systems. These materials are also included in Table 18.12.

**Table 18.12:** Organic hydrocolloids useful in emulsion technology

| Source | Name | Comment |
|---|---|---|
| Tree exudates | Gum Arabic (Acacia) | Essentially neutral polysaccharide |
| | Gum ghatti | Essentially neutral polysaccharide |
| | Karaya | Essentially neutral polysaccharide |

| | | |
|---|---|---|
| | Tragacanth | Essentially neutral polysaccharide |
| Sea weeds | Agar, Carrageenan | Sulfated polysaccharide |
| | Alginates | Acidic polysaccharide |
| Seed extracts | Locust bean | Essentially neutral polysaccharide |
| | Guar | Essentially neutral polysaccharide |
| | Quince seed | Essentially neutral polysaccharide |
| Synthetic (fermentation) | Xanthan gum | Essentially neutral polysaccharide |
| Cellulose | Methyl-, hydroxyethyl-hydroxypropyl ether | Neutral polysaccharide |
| | Carboxymethyl ether | Anionic polysaccharide |
| Collagen | Gelatin | Amphoteric protein |
| Synthetic | Polyoxethylene polymer | Neutral |
| | Carboxyvinyl polymer (cross-linked) | Anionic |

The water-sensitive hydrocolloids generally favour o/w emulsions because they form excellent hydrophilic barriers. Their use is warranted whenever it is desired to increase the viscosity of an emulsion without a corresponding increase in the lipid portion of the emulsion. Proteins, as a group, are effective not only as primary emulsifiers but also as auxiliary emulsifiers. They are particularly useful in oral dosage forms.

## *Finely Divided Solids*

Finely-divided solids have been shown to be good emulsifiers, especially in combination with surfactants and/or macromolecules that increase viscosity. Included are polar inorganic solids, such as heavy metal hydroxides, certain non-swelling clays, and pigments. Even nonpolar solids, e.g. carbon or 

glyceryl-tristearate, can be used. Polar solids tend to be wetted by water to a greater extent than by the oil phase, whereas the reverse is true for non-polar solids. In the absence of surfactants, w/o emulsions are favoured by the presence of non-polar solids, presumably because the wetting by oil facilitates the coalescence of oil droplets during the initial steps of emulsification. An analogous interpretation may be given for the tendency of polar solids to favour water as the external phase.

In the presence of wetting agents, i.e. when such solids are used as auxiliary emulsifiers, their behaviour is controlled by the so-called Young equation. For example, barium sulfate in the presence of sodium laurate (at pH 12) favours o/w emulsions, whereas barium sulfate coated with sodium dodecyl sulfate favours w/o emulsions. In view of the limited utility of such solids as primary emulsifiers, or even as auxiliary emulsifiers, they are not of major interest to the formulator.

## Viscosity Modifiers

Once the desired emulsion and emulsifiers have been chosen, a consistency that provides the desired stability and yet has the appropriate flow characteristics must be attained. It has already been mentioned that the viscosity of an emulsion can be altered by manipulating the composition of the lipid phase, by variations in the phase ratio and the surfactants and by the addition of gums. It is well-known that the creaming of fluid emulsions depends on their rheologic character as well as on the surface characteristics of the interfacial film. The use of gums, clays, and synthetic polymers in the continuous phase of emulsions is a powerful tool for enhancing an emulsion's stability. As was pointed out earlier in this Chapter, the sedimentation or creaming rate of suspended spherical particles is inversely proportional to the viscosity in accordance with Stoke's law. When all other variables are held constant, an increase in viscosity generally minimizes creaming, rising, or sedimentation. Since emulsions should flow or spread, and since higher viscosity favours stability, thixotropy in an emulsion is desirable. In the case of emulsions that also contain suspended solids, the presence of a yield value assures at least reasonable stability.

It is routinely observed that the building up of viscosity in a freshly prepared emulsion requires some time. It is recommended, therefore, that a newly-formulated emulsion be allowed to rest undisturbed for 24 to 48 h before it is determined whether its rheologic properties correspond to those that are required. The viscosity of emulsions responds to changes in composition in accordance with the following generalizations.

1. There is a linear relationship between emulsion viscosity and the viscosity of the continuous phase: (The use of gums and clays for o/w emulsions has already been noted.) In the case of w/o emulsions, the addition of polyvalent metal soaps or the use of high melting waxes and resins in the oil phase can be used to increase viscosity. Emulsion viscosity is not very sensitive to viscosity changes of the internal phase.
2. The greater the volume of the internal phase, the greater is the apparent viscosity.
3. To control emulsion viscosity, three interacting effects must be balanced by the formulator (1) the viscosity of o/w and w/o emulsions can be increased by reducing the particle size of the dispersed phase, (2)

emulsion stability is improved by a reduction in particle size and (3) flocculation, or clumping, which tends to structure the internal phase, can be a stabilizing effect, but it increases viscosity.

4. As a rule, the viscosity of emulsions increases upon aging.

## Antimicrobial Preservative

Emulsions often contain a number of ingredients, such as carbohydrates, proteins, sterols, and phosphatides, all of which readily support the growth of a variety of microorganisms. Even in the absence of any of the aforementioned natural ingredients, the mere presence of a mixture of lipid and water in intimate contact frequently allows microorganisms to establish themselves. As a result, the inclusion of a preservative is a necessary part of the formulation process. Several points must be kept in mind in selecting a preservative. Microbial contamination may occur during the development or production of an emulsion or during its use. Frequently, the microbial contamination can arise from the use of impure raw materials or from poor sanitation during preparation. Alternately, contamination may be the result of invasion by an opportunistic microorganism. Finally, the consumer may actually inoculate the product during use. It is commonly believed that a preservative or a system of preservatives can protect the emulsion against all of these possibilities. The formulation of a self-sterilizing emulsion is exceedingly difficult without the use of potent antimicrobial agents, most of which have been or are in the process of being reviewed by governmental regulatory agencies. Prevention of contamination is recommended, and certain cardinal rules must be observed. The most important one is the use of uncontaminated raw materials, including water. A second precaution is meticulous housekeeping and careful cleaning of equipment (with live steam). Once a microbiologically uncontaminated product has been prepared, a relatively mild antimicrobial agent suffices to protect the product against chance contamination by microorganisms. It is also desirable that the preservative system be effective against invasion by a variety of pathogenic organisms and be adequate to protect the product during use by the consumer.

As is true of most ingredients of a formulation, the preservative system must first meet the general criteria of low toxicity, stability to heat and storage, chemical compatibility, reasonable cost, and acceptable taste, odour, and colour. Efficacy against a variety of organisms is required since fungi, yeasts, and bacteria are common contaminants. The more important groups of preservatives and some popular examples used in emulsions are shown in Table 18.13. The activity of the antimicrobial agents listed in the table varies widely and depends on the microorganism involved.

The concentration of preservative required in an emulsion depends to a large extent on its ability to interact with microorganisms. Since microorganisms can reside in the water or the lipid phase or both, the preservative, regardless of its water-oil partition coefficient, should be available at an effective level in both phases. It is almost inconceivable that a single preservative could distribute itself at effective concentrations between the phases, regardless of their compositions. It is therefore customary to include a preservative that is soluble in the water phase and one that is primarily soluble in the oil phase. The esters of p-hydroxybenzoic acid are particularly good examples because the methyl ester is water-soluble, whereas the propyl and higher esters exhibit almost no water solubility. The distribution of a preservative between the lipid and the aqueous phase of emulsions can be determined by procedures commonly employed for evaluating distribution coefficients.

**Table 18.13:** Some typical preservatives used in pharmaceutical and cosmetic emulsions

| Type | Example | Characteristics and utility |
|---|---|---|
| Acids and acids derivatives | Benzoic acid and salts<br>Propionic acid and salts<br>Dehydroacetic acid | Antifungal agent |
| Alcohols | Chlorobutanol<br>Phenoxy-2-ethanol | Eye preparations<br>Synergist |
| Aldehydes | Formaldehyde<br>Glutaraldehyde | Broad spectrum |
| Formaldehyde donors | Hexamethylenetetramine<br>Mono (and di–) methylol dimethyl hydantoin | |
| Phenolics | Phenol<br>Cresol<br>*a*-Chlorometaxylenol<br>Methyl *p*-hydroxybenzoate<br>Propyl *p*-hydroxybenzoate<br>Butyl *p*-hydroxybenzoate | Broad spectrum |
| Quaternaries | Chlorhexidine salt<br>Benzal konium chloride<br>Cetyl trimethyl ammonium bromide | |
| Mercurials | Phenylmercuric acetate<br>Sodium ethylmercurithiosalicylate | |
| Miscellaneous | 6-Acetoxy-2, 4-dimethyl-*m*-dioxane | |
| | 2, 4, 4′ -Trichloro-2′ hydroxy-diphenylether | Primarily against gram-positive bacteria |
| | (1-(3-Chloroallyl)-3, 5, 7 triazo-1 -azoniadamantane chloride | Broad spectrum and synergist |
| | Imidizolidinyl urea compound | Synergist |
| | Bromo-2-nitropropanediol-1, 3<br>5-Bromo-5-nitrol-1, 3-dioxane<br>2-Thiophyridine N-oxide (and salts)<br>2-Methyl-4-isothiazolin-3-one and 5-chloro derivative | Broad spectrum |

Complex problems arise whenever the preservative interacts with one of the emulsion ingredients. Such interactions may inactivate the preservative. The interaction with emulsion ingredients of various alkyl hydroxybenzoates (the most widely used preservatives in emulsions) has been studied for many years and is illustrated by the data in Table 18.14. As a rule, the so-called bound preservatives are not readily available to exert antimicrobial activity. It is apparent that the phenolic preservatives are especially susceptible to interaction with compounds containing polyoxyethylene groups.

**Table 18.14:** Interaction of parabens with emulsion ingredients

| | Methyl-p-hydroxy-benzoate | | Propyl-p-hydroxy-benzoate | |
|---|---|---|---|---|
| **Macromolecular compounds** | **Free (%)** | **Bound (%)** | **Free (%)** | **Bound (%)** |
| Gelatin | 92 | 8 | 89 | 11 |
| Methyl cellulose | 91 | 9 | 87 | 13 |
| Polyethyleneglycol 4000 | 84 | 16 | 81 | 19 |
| Polyvinylpyrrolidone | 78 | 22 | 64 | 36 |
| Polyoxyethylene monostearate | 55 | 45 | 16 | 84 |
| Polyoxyethylene sorbitan monolaurate | 43 | 57 | 14 | 86 |
| Polyoxyethylene sorbitan monooleate | 43 | 57 | 10 | 90 |

To compensate for the loss of preservative by interactions, an amount equal to the complexed material may be added. It has also been found that the addition of various alcohols seems to activate p-hydroxybenzoate esters in the presence of nonionics; propylene glycol appears to be especially useful.

Several other factors can alter the ability of preservatives to protect a product against microbial contamination. The pH is known to exert a major influence on the ability of acidic or phenolic preservatives to interfere with microbial growth. These agents are almost completely inactivated by converting them into anions. Other factors include the phase ratio, the degree of aeration during preparation, and especially the presence of flavors and perfumes, some of which have antimicrobial properties. Combinations of preservatives are often used, since they have been shown to increase the effectiveness of preservative action, either by an enhancement of the spectrum of activity or by some synergistic behaviour.

Although much has been written on the subject of preservation and on the utility of a particular preservative or preservative combination, the selection of a preservative for actual use in a specific emulsion is somewhat empiric. Formulators are cautioned not to depend on chemically-determined availability of preservatives to establish the microbiologic cleanliness of a product. Instead, rigorous microbiologic examination of the final composition is required to determine whether an emulsion is properly preserved.

## Antioxidant

Many organic compounds are subject to autoxidation upon exposure to air, and emulsified lipids are particularly sensitive to attack. Many drugs commonly incorporated into emulsions are subject to autoxidation and resulting decomposition.

Upon autoxidation, unsaturated oils, such as vegetable oils, give rise to rancidity with the resultant unpleasant odour, appearance, and taste. On the other hand, mineral oil and related saturated hydrocarbons are subject to oxidative degradation only under rare circumstances.

Autoxidation is a free-radical chain oxidation reaction. It can be inhibited, therefore, by the absence of oxygen, by a free radical chain breaker, or by a reducing agent. Materials that are useful as antioxidants by one or more of these three mechanisms are listed in Table 18.15. The choice of a particular antioxidant depends on its safety, acceptability for a particular use, and its efficacy. Antioxidants are commonly used at concentrations ranging from 0.001 to 0.1% (w/v). Butylated hydroxyanisole (BHA), butylated hydroxytoluene (BHT), L-tocopherol, and the alkyl gallates are particularly popular in pharmaceuticals and cosmetics. BHT and BHA have a pronounced odour and should be used at low concentrations.

**Table 18.15:** List of antioxidants

| | |
|---|---|
| Gallic acid | L-Tocopherol |
| Propyl gallate | Butylated hydroxytoluene |
| Ascorbic acid | Butylated hydroxyanisol |
| Ascorbyl palmitate | 4-Hydroxymethyl-2,6-di-tert-butylphenol |
| Sulfites | |

Alkyl gallates have a bitter taste, whereas L-tocopherol is well-suited for edible or oral preparations, such as those containing vitamin A. Almost all antioxidants are subject to discolouration in the presence of light, trace metals, and alkaline solutions. Combinations of two or more antioxidants have been shown to produce synergistic effects. In some cases, compounds

completely devoid of antioxidant activity by themselves enhance the effectiveness of certain antioxidants. For example, alkyl gallates, BHT, and BHA are much more effective in the presence of citric, tartaric, or phosphoric acids. A related way of enhancing the activity of phenolic antioxidants involves the use of a small amount of a sequestrant for heavy metal, calcium, and magnesium ions.

## EMULSION FORMATION

The spontaneous formation of an emulsion is a relatively rare occurrence. Instead, emulsion preparation by the commonly employed dispersion method requires a sequence of processes for breaking up the internal phase into droplets and for stabilizing them in the external phase. The complete process must be designed in such a way that these two steps are carried out before the internal phase can coalesce. Usually, the break-up of the internal phase (by physical means) is fairly rapid; however, it is believed that the stabilization step and rate of coalescence are time- and temperature-dependent.

The application of energy in the form of heat, mechanical agitation, ultrasonic vibration, or electricity is required to reduce the internal phase into small droplets. The amount of work input depends on the length of time during which energy is supplied; thus, timing (scheduling of work input) becomes another important physical parameter.

## Condensation Method

Vapourization is an effective way of breaking almost all the bonds between the molecules of a liquid. It is possible, therefore, to prepare emulsions by passing the vapour of a liquid into an external phase that contains suitable emulsifying agents. This process of emulsification, called the condensation method, is relatively slow, is limited to the preparation of dilute emulsions of materials having a relatively low vapour pressure, and is therefore primarily of theoretical importance.

The more practical emulsification is affected by changes in temperature in a number of ways. The interactions are complex, and it is almost impossible to predict whether a raise in temperature will favour emulsification or coalescence. An increase in temperature decreases the interfacial tension as well as viscosity. One would therefore predict—and this is usually true—that emulsification is favored by an increase in temperature. At the same time, however, an increase in temperature raises the kinetic energy of droplets and thereby facilitates their coalescence. This type of instability is normally observed when emulsions are stored at elevated temperatures for long periods of time. Changes in temperature alter the distribution coefficients of emulsifiers between the two phases and cause emulsifier migration. The distribution of the emulsifier as a function of temperature cannot be correlated directly with either emulsion formation or stability, since changes in surface tension and viscosity occur simultaneously.

## Phase Inversion Method

The most important influence that temperature has on an emulsion is probably inversion. Almost 50 years ago it was observed that w/o emulsions of benzene in water that were stabilized with sodium stearate invert to o/w emulsions upon heating and reform w/o emulsions upon cooling. The temperature at which inversion occurs depends on emulsifier concentration and is called *phase inversion temperature (PIT)*. This type of inversion can occur during the formation of emulsions, since they are generally prepared at relatively high temperatures and are then allowed to cool to room temperature. Emulsions formed by a phase inversion technique are generally considered quite stable and are believed to contain a finely dispersed internal phase. The PIT is generally considered to be the temperature at which the *hydrophilic* and lipophilic properties of the *emulsifier* are in *balance* and is therefore also called the *HLB* temperature.

Shinoda's description of the processes at or near the PIT is almost universally accepted today. An o/w emulsion stabilized by a nonionic polyoxyethylene-derived surfactant contains oil-swollen micelles of the surfactant as well as emulsified oil. When the temperature is raised, the water-solubility of the surfactant decreases; as a result, the micelles are broken, and the size of emulsified oil droplets begins to increase. A continued rise in the temperature causes separation into an oil phase, a surfactant phase, and water. It is near this temperature that the now water-insoluble surfactant begins to form a w/o emulsion containing both water-swollen micelles and emulsified water droplets in a continuous oil phase.

## Low-energy Emulsification

The classic process of emulsification just described requires considerable expenditure of energy during both the heating and cooling cycles of emulsion formation. The principle of low-energy emulsification has been formalized by Lin in recent years, although some of his suggestions may have been previously used by other practitioners. In low-energy emulsification, all of the internal phase, but only a portion of the external phase, is heated. After emulsification of the heated portions, the remainder of the external phase is added to the emulsion concentrate, or the preformed concentrate is blended into the continuous phase. In those emulsions in which a phase inversion temperature exists, the emulsion concentrate is preferably prepared above the PIT, which results in emulsions having extremely small droplet size. As in the case of the classic emulsification technique, the temperature for the preparation of the emulsion concentrate is critical. It is important to effect in situ neutralization of acidic emulsifying components during the emulsion step. By careful control of the variables (emulsification temperature, mixing intensity, the amount of the external phase employed during emulsification, and the method of bleeding), it is reportedly possible to produce emulsions with smaller and more uniform particle size than those resulting from the conventional process.

### *Mechanical Equipment for Emulsification*

Almost all methods used for breaking up the internal phase into droplets depend on “brute force” and require some sort of agitation. When a liquid jet of one liquid is introduced under pressure into a second liquid, the initially cylindric jet stream is broken up into droplets. The factors that enter into the breakup of a liquid jet include the diameter of the nozzle, the speed with which the liquid is injected, the density and viscosity of the injected liquid, and of course, the interracial tension between the two liquids. A similar break-up into droplets occurs when a liquid is allowed to flow into a second liquid that is agitated vigorously. Once the initial breakup into droplets has occurred, the droplets continue to be subject to additional forces due to turbulence, which cause deformation of the droplet and further breakdown into smaller droplets. Various types of equipment are available to affect droplet break-up and emulsification either in the laboratory or in production. Regardless of size and minor variations, such equipment can be divided into

four broad categories (1) mechanical stirrers, (2) homogenizers, (3) ultrasonifiers and (4) colloid mills.

During the formulation of an emulsion, the mechanical requirements of preparation, and particularly the problems associated with scale-up to production-size equipment, must be considered. The most important factor involved in the preparation of an emulsion is the degree of shear and turbulence required to produce a given dispersion of liquid droplets. The amount of agitation required depends on the total volume of liquid to be mixed, the viscosity of the system, and the interfacial tension at the oil-water interface. The latter two factors are determined by the emulsion type, the phase ratio, and the type and concentration of emulsifiers. For this reason, no single method of dispersion can be used for all emulsions, and conversion from one method to another is difficult.

Multiple emulsions of the w/o/w type are normally prepared by first forming a w/o emulsion with the aid of a low *HLB* emulsifier. This w/o emulsion is then slowly incorporated into an aqueous phase that contains an emulsifier of a significantly higher *HLB*, i.e. approximately 12 to 14.

## *Mechanical Stirrers*

An emulsion may be stirred by means of various impellers mounted on shafts, which are placed directly into the system to be emulsified. Simple top-entering propeller mixers are adequate for routine development work in the laboratory and for production purposes, if the viscosity of the emulsion is low. If more vigorous agitation is required, or if the preparation has moderate viscosity, turbine type mixers are employed both in the laboratory and in production. Other mixers, provided with paddle blades, counter rotating blades, or planetary action blades, are available for special requirements. The degree of agitation is controlled by the speed of impeller rotation, but the patterns of liquid flow and resultant efficiency of mixing are controlled by the type of impeller, its position in the container, the presence of baffles, and the general shape of the container (Chapter 1). Despite the variation in flow behaviour and efficient mixing that can be produced, the use of stirrers for the formation of emulsions is often limited when vigorous agitation of viscous systems is required, when extremely fine droplets are needed, or when foaming at high shear rates must be avoided.

## *Homogenizers*

In a homogenizer, the dispersion of two liquids is achieved by forcing their mixture through a small inlet orifice at high pressures. Homogenizers of varying designs are useful for handling either liquids or pastes, since the rate of throughput is little affected by viscosity. It must be remembered, however, that homogenization raises the temperature of the emulsion, and subsequent cooling may be required. The use of a homogenizer is warranted whenever a reasonably monodisperse emulsion of low particle size is required. Another useful piece of equipment, which combines mixing with some homogenizing action, is the high-pressure homogenizer described in Chapter 2.

## *Ultrasonifiers*

The use of ultrasonic energy to produce pharmaceutical emulsions has been demonstrated, and many laboratory-size models are available. These transduced piezoelectric devices have limited output and are relatively expensive. They are useful for the laboratory preparation of fluid emulsions of moderate viscosity and extremely low particle size. Commercial equipment is based on the principle of the Pohlman liquid and are described in Chapter 2.

## *Colloid Mills*

Homogenizers and ultrasonic equipment depend on sudden changes in pressure to effect the dispersion of liquids. By contrast, colloid mills operate on the principle of high shear, which is normally generated between the rotor and the stator of the mill (Chapter 1). Colloid mills are used primarily for the comminution of solids and for the dispersion of suspensions containing poorly-wetted solids but are also useful for the preparation of relatively viscous emulsions.

## Spontaneous Emulsification

Spontaneous emulsification occurs when an emulsion is formed without the application of any external agitation. Emulsifiable concentrates and microemulsions are typical examples. The former are blends of the internal phase with emulsifiers, which "bloom" when they are added to the external phase. Microemulsions commonly form spontaneously, but not all spontaneous emulsions are transparent. The phenomenon of spontaneous emulsification is observable when a drop of oil is placed on an aqueous solution of an emulsifier, in which case the interface becomes extremely unstable and results in the formation of fine droplets. Spontaneous emulsification evidently is not practiced commercially.

In general, the considerations applicable to opaque emulsions are also pertinent to the preparation of clear emulsions. The amount of internal phase in clear emulsions or in solubilized systems is generally lower than that in opaque emulsions. Most emulsion technologists have found that an increase in the surfactant concentration(s) reduces the opacity of all types of emulsions, and if carried further, can result in solubilization.

Cosmetic and pharmaceutical microemulsions usually do not employ the cosolvents required for the more classic microemulsions of theoretic interest. Instead, modern commercial solubilized systems are frequently based on non-ionic emulsifiers, which result in the formation of micellar solutions. If the solubilizate is a drug, the drug available for absorption through the skin may not equal the total amount of drug in the product. That portion of the drug that is incorporated into the interior of micelles may not be available for absorption unless the micelles exhibit instability. The actual location of the drug is described by a micellar distribution coefficient defined as:

$$K_m = \frac{S_m}{S_w} \qquad \text{... (13)}$$

where, $S_m$ is the solubility of the active ingredient in the micellar phase and $S_w$ is its solubility in water. The value for $K_m$ is established by classic solubility determinations. The value for $S_m$, e.g. in the case of dexamethasone, increases linearly with increasing concentrations of a non-ionic surfactant, such as polyoxyethylene (20) stearyl alcohol. Comparable results are obtained by equilibrium (or differential) dialysis through

membranes that are impermeable to surfactant micelles but permeable to the free steroid, which also yields a slightly different distribution coefficient, $K'_m$. In the case under discussion, $K_m$ is large (≈ 400) and almost identical to $K'_m$.

During the processing of solubilizates, these “clear” emulsions are frequently opaque at high temperatures (in view of the PIT). Clearing of the emulsion occurs as the preparation cools, and solubilized systems (in the absence of cosolvents) usually remain clear to temperatures close to the freezing point of one of the major components. In the case of non-ionic o/w solubilizates, it is sometimes helpful to include a small amount of a surfactant with a high *HLB* (e.g. alkyl sulfate) to raise the PIT if it is too close to the expected storage temperature.

Whenever it is desirable to “dissolve” a small amount of a flavour or fragrance in an essentially aqueous system, the formulator should be guided by some important practical rules: (1) Surfactants in the *HLB* range of 15 to 18 are ideal solubilizing agents for this purpose (*see* Tables 15.1 and 17.4), (2) It sometimes helps to add a small amount of a surfactant with an HLB in the range of 8 to 12. 3. As a rule, 3 to 5 times as much surfactant as oil is required to effect solubilization. The oil should always be blended with the surfactant before addition to the (warmed) aqueous phase. This is particularly important for the oil-soluble vitamins. The incorporation of a flavour or fragrance into solubilizing systems generally reduces the organoleptic impact of the oil, since as described above, part of it is entrapped into micelles. To enhance the flavour or fragrance intensity, it is sometimes necessary to use a cosolvent, such as ethanol.

## PRODUCTION ASPECTS

In routine production, it is customary to prepare emulsions by a batch process using kettles, agitators, and related equipment. However, it is possible to design combinations of equipment that permit continuous manufacturing of emulsions. The selection of commercial equipment for the production of emulsions is based in part on the production capacity and the power requirements for various types of apparatus.

**Foaming during agitation:** During the agitation or transfer of an emulsion, foam may be formed. Foaming occurs because the water-soluble surfactant required for emulsification generally also reduces the surface tension at the air-water interface. To minimize foaming, emulsification may be carried out in closed systems (with a minimum of free air space) and/or under vacuum. In addition, mechanical stirring, particularly during the cooling of a freshly prepared emulsion, can be regulated to cause the air to rise to the top. If these precautions should fail to eliminate or reduce foaming, it is sometimes necessary to add foam depressants (antifoams); however, their use should be avoided, if at all possible, since they represent a chemical source of incompatibility. Sometimes the use of ethyl alcohol accelerates the coalescence of foam on the surface of emulsions. On the other hand, the most effective defoamers are long-chain alcohols and commercially available silicone derivatives, both of which are generally believed to spread over the air-water interface as insoluble films.

**Chemical stability:** Chemical inertness is an absolute and almost obvious requirement for emulsion ingredients. For example, it would be futile to utilize soap as an emulsifier in a system having a final pH of less than 5. Similarly, one would not use an easily hydrolyzed ester in an emulsion that is either acidic or alkaline. Some lipids are subject to chemical changes due to oxidation (rancidity); in general it is simpler to avoid their use than to depend on antioxidants to ensure their stability. It is important, therefore, that the chemical nature of all emulsion constituents be understood before a selection for a given preparation is made.

Unfortunately, predictions of hydrolytic stability made by classic chemical or pharmaceutical procedures may on occasion be unreliable, as a result of micellar catalysis. This type of catalysis can be observed whenever the reactive species are present on or near the micellar surface. Under these

conditions, hydrolytic (and substitution) reactions can be accelerated. The decomposition of drugs via micellar catalysis has not been studied extensively, but the hydrolysis of alkyl sulfates is a simple and particularly important example of micellar catalysis. The hydrolysis of dodecyl sulfates has been shown to depend not only on the pH of the medium, but also on the presence of a variety of electrolytes and on the concentration of the surfactant; in addition, it is subject to micellar catalysis.

**Safety:** Safety and toxicologic clearance of components of pharmaceutical and cosmetic emulsions are absolute requirements. It is essential, therefore, for the formulator to depend heavily on toxicologic information from suppliers or in the scientific literature, as well as on regulatory activities by governmental agencies. Despite these almost obvious limitations, the formulator has an enormous choice of emulsion ingredients, which differ in their cost and ability to yield the desired product.

## ADDITIONAL RECOMMENDATIONS

In the laboratory development of emulsions, it is common practice to prepare an oil phase containing all the oil-soluble ingredients and to heat it at about 5 to 10°C above the melting point of the highest melting ingredient. The aqueous phase is normally heated to the same temperature, and then the two phases are mixed. A laboratory beaker containing a hot emulsion cools fairly rapidly to room temperature, but a production tank filled with hundreds of gallons of hot material cools more slowly unless external means of cooling are employed. This is one reason that the simple transfer, of a laboratory process to production requires extensive studies of the cooling and agitation schedule:

- It is advisable to utilize jacketed equipment for the large-scale preparation of emulsions, so that the heating and cooling cycles can be carefully controlled.
- In the preparation of anionic or cationic o/w emulsions, it is customary to add the oil phase to the water phase, although some technologists prefer the inversion technique, i.e. addition of the water phase to the oil phase.
- In the case of non-ionic emulsions, which exhibit a PIT, the inversion technique is not required since temperature alone can be used to control this stage of emulsification.
- If soap is used as the emulsifier, it is usually prepared in situ by combining the alkali with the water phase and the fatty acid with the oil phase.
- Oil-soluble emulsifiers are commonly added to the lipid phase, whereas the water-soluble emulsifiers are dissolved in the aqueous phase. Occasionally, it may prove advantageous to include even the water-soluble emulsifier in the oil phase.
- In the preparation of w/o emulsions, it is almost always necessary to add the water phase slowly to the oil/emulsifier blend.
- To avoid losses, volatile flavours or perfumes are preferably added at the lowest temperature at which incorporation into the emulsion is possible (usually 55 to 45°C).
- If a gum is employed, it should be completely hydrated or dissolved in the aqueous phase before the emulsification step. If a heat-sensitive gum

is used, it may be necessary to incorporate the gum solution after the emulsion has been formed. The use of two different organic gums can cause incompatibility.

- It is also noted that anionic and cationic emulsifiers in about equimolar quantities rarely yield satisfactory emulsions.
- Not unexpectedly, emulsions designed for parenteral administration can be prepared with only a limited number of emulsifiers (*see* Table 18.7).
- It is recommended that parenteral emulsions, especially those designed for intravenous injection, be homogenized until a satisfactory particle size is achieved.
- Since the use of conventional preservatives is contraindicated, such preparations require sterilization at high temperature but must still yield acceptable emulsions after this heating/cooling cycle.
- Whenever an emulsion is formed at elevated temperatures, the loss of water due to evaporation must be made up. This is done best by adjusting to "final weight" with water when the emulsion reaches about 35°C.

**Practical examples:** A few selected formulations that have been published in the literature are presented below. They have not been prepared by the author nor have they been screened for stability or safety. They are cited here merely to illustrate the utility of various emulsifiers and to indicate some practical means for forming various emulsions. Only the first one is discussed in detail.

## 1. Oral Emulsion (o/w)

| | | % w/w |
|---|---|---|
| (A) | Cottonseed oil, winterized | 460.0 g |
| | Sulfadiazine | 200.0 g |
| | Sorbitan monostearate | 84.0 g |
| (B) | Polyoxyethylene (20) sorbitan monostearate | 36.0 g |
| | Sodium benzoate | 2.0 g |
| | Sweetener | q.s. |
| | Water, potable | 1000.0 g |
| (C) | Flavour oil | q.s. |

*Procedure*

1. Heat (A) to 50°C and pass through a colloid mill.
2. Add (A) at 50°C to (B) at 65°C and stir while cooling to 45°C.
3. Add (C) and continue to stir until the room temperature is reached.

**Discussion and critique:** Sulfadiazine is essentially water-insoluble. A suspension or emulsion is required to yield a fluid oral dosage form. To maintain sulfadiazine in suspension, the viscosity of the final product must be reasonably high. This could be achieved by the use of gums or by developing an emulsion high in internal phase. The choice of a cottonseed oil o/w emulsion is probably arbitrary except for the fact that o/w preparations are quite palatable.

Table 18.8 shows that an HLB of about 10 is required to yield a fluid emulsion of cottonseed oil. Although a single emulsifier, such as polyoxyethylene (4) sorbitan monostearate (HLB 9.6), might be satisfactory, the use of mixed emulsifiers is generally preferred. In view of their “safety” and availability, a blend of sorbitan monostearate (*HLB* 4.7) and polyoxyethylene (20) sorbitan monostearate (*HLB* 14.9) seems promising. The ratio required to yield an *HLB* of 10.0 is computed from $a \times 4.7 + b \times 14.9 = 10$, where a and b are the weight fraction of each of the two emulsifiers and where $a + b = 1$. It is found that a blend of 48% of the lipophilic and 52% of the hydrophilic emulsifier yields the desired *HLB*. In

fact, the formula calls for 70% of the hydrophobic emulsifier, equivalent to an *HLB* of the blend of 8.2. This must be attributed to the presence of the sulfadiazine and the need for a high viscosity emulsion. The ratio of emulsifiers: oil (about 1:4) is high and again points towards an unpredictable effect of sulfadiazine.

To develop a smooth product, it is necessary to reduce the particle size of sulfadiazine. For this purpose, the blend of the drug, the oil, and the oleophilic emulsifier is warmed slightly and passed through a colloid mill. The emulsion itself is formed by adding the drug suspension to the aqueous phase, but in contrast to usual emulsion technique, the two phases are blended at different temperatures. Presumably, heating of the drug suspension to 65°C would materially lower its viscosity and cause excessive settling of the drug particles unless specialized stirring equipment were employed.

The addition of the flavouring oils at a lower temperature prevents loss due to volatility. The following additional points are pertinent: Since sulfadiazine has a broad antimicrobial spectrum, the absence of a preservative in the oil phase is not surprising. Nevertheless, the emulsion should be protected against moulds and fungi; the use of sodium benzoate for this purpose is of doubtful merit. The absence of an antioxidant in this emulsion suggests the presence of additives in the cottonseed oil.

## 2. Medicated Ointment Base (w/o)

| | % w/w |
|---|---|
| (A) Mineral oil, USP | |
| (125/135 Saybolt units at 100°F) | 25.0 |
| Microcrystalline wax | |
| (MP 170–180°F) | 10.0 |
| Cetyl alcohol | 5.0 |
| Mixed lanolin alcohols (high in cholesterol) | 10.0 |
| Sorbitan sesquioleate | 3.0 |
| Propyl p-hydroxybenzoate | 0.1 |
| (B) Glycerine | 3.0 |
| Methyl p-hydroxybenzoate | 0.1 |
| Deionized water | 43.8 |
| (C) Medicament | q.s. |

*Procedure*

1. Heat (A) and (B) separately to 75°C.
2. Add (B) to (A) and stir while cooling to 45°C.
3. Add (C), mix, and package.

**Discussion and critique:** This emulsion illustrates a conventional approach (the addition of solid oil-soluble components) for increasing the viscosity of the external phase in a w/o emulsion. Lanolin and its derivatives are generally useful for formulating w/o emulsions and may assist in effecting dissolution of the medicament. The addition of preservatives to both phases is noteworthy.

## 3. Emollient Cream (o/w)

| | | % w/w |
|---|---|---|
| (A) | Acetylated lanolin | 2.00 |
| | Stearic acid | 2.00 |
| | Petrolatum | 4.00 |
| | Lanolin absorption base | 6.00 |
| | Glyceryl monostearate, pure | 12.00 |
| | Propyl-p-hydroxybenzoate | 0.15 |
| (B) | Water | 62.90 |
| | Glycerol | 10.00 |
| | Sodium lauryl sulfate | 0.50 |
| | Methyl p-hydroxybenzoate | 0.15 |
| | Perfume | 0.30 |

*Procedure*

1. Heat (A) and (B) separately to 75°C.
2. Add (A) to (B) and stir until temperature reaches 45°C.
3. Add perfume and continue to stir gently until room temperature is reached.

**Discussion and critique:** This cosmetic preparation illustrates the use of relatively high concentrations of glyceryl monostearate to create a viscous o/w emulsion. Lanolin absorption bases are blends of lanolin (and its derivatives) with hydrocarbons and a small amount of emulsifier. They are normally employed in the formulation of w/o emulsions. Evidently, the method of emulsification and the use of the extremely hydrophilic sodium lauryl sulfate permit the formation of an o/w emulsion. Particularly striking is the use of free stearic acid as a means of bodying the lipid portion of this cream. The preparation's viscosity would be reduced significantly if the stearic acid were neutralized, e.g. with triethanolamine.

## 4. Makeup Cream (o/w)

| | | % w/w |
|---|---|---|
| (A) | Magnesium aluminum silicate | 2.60 |
| | Sodium carboxymethyl cellulose | 0.40 |
| | Water, distilled | 42.40 |
| (B) | Dispersing agent | 0.30 |
| | Propylene glycol | 5.00 |
| | Water, distilled | 12.30 |
| (C) | Talc | 18.50 |
| | Kaolin | 1.30 |
| | Titanium dioxide | 3.70 |
| | Iron oxides | 1.50 |
| (D) | Isopropyl myristate | 5.00 |
| | Stearyl alcohol | 2.00 |
| | Lanolin absorption base | 2.00 |
| | Sorbitan monolaurate | 0.75 |
| | Polyoxyethylene (20) sorbitan monolaurate | 2.25 |
| | Preservative | q.s. |
| | Perfume | q.s. |

*Procedure*

1. Blend solids of (A) and add to water at 80°C; stir until smooth.
2. Micropulverize (C) and add to (B); pass through colloid mill to yield smooth paste.
3. Add (B) + (C) to (A) and heat to 60 to 65°C.
4. Heat (D) to 70°C and add to (A) + (B) + (C) blend; stir until temperature reaches 45°C; add perfume and mix until cool.

**Discussion and critique:** This composition illustrates the use of an organic gum and an inorganic clay to achieve viscosity and a yield point in order to reduce settling of heavy pigments. The need for a pigment dispersing

agent (a complex alkylaryl sulfonate) can generally be avoided by blending the micro-pulverized pigments (C) with the warm oil phase (D) and passing this blend through a colloid mill. The mixture of (C) + (D) can then be added to the aqueous phase containing the gums, humectants, and a trace of surfactant (about 0.1% sodium lauryl sulfate).

## 5. Fat Emulsions for Parenteral Nutrition (o/w)

| | | % w/w |
|---|---|---|
| (A) | Cottonseed oil, winterized | 15.0 |
| | Polyethylene glycol 200 monopalmitate, purified | 1.2 |
| | Tartaric acid ester of cotton seed fatty acid monoglyceride | 0.3 |
| (B) | Polyoxyethylene—poly oxypropylene block polymer | 0.3 |
| | Isotonic glucose solution | 83.2 |

*Procedure*

1. Blend and heat (A) and (B) separately to 45°C.
2. Circulate (B) in a two-stage homogenizer at 3000 psi.
3. Add (A) to (B) at 40 to 45°C and raise the pressure to 4000 psi.
4. Continue homogenization until the emulsion has been cycled five times.
5. Package and autoclave at 121°C for 17 min.

**Discussion and critique:** This emulsion is unusual since it could be expected to show signs of hydrolysis of the lipid and a drop in pH upon storage. Better stability could be achieved through use of a system of Pluronic F68 and L-α-phosphatidyl choline as described by Guay and Bissailon.

| | |
|---|---|
| Soybean oil | 10.0% |
| Dextrose | 4.0% |
| Emulsifier (blend of L-α-phosphatidyl choline and Pluronic F-68 at HLB 10) | 3.0% |
| Water | q.s.100.0% |

The aqueous phase (at 95°C) was added to the oil phase (also at 95°C), and the blend was mechanically stirred for 45 min. Two hundred grams of this mixture was then sonicated in a 400 ml beaker—using a model W140D (20 kHz) sonifier (Branson Sonic Power Co. in an ice bath at 16 kHz for 12

min. The emulsion was then autoclaved at 121°C and 15 psi for 15 min. The resulting product exhibited good pH stability and resistance to creaming and coalescence during centrifugation.

## 6. High Internal Phase Emulsion (w/o)

| | | % w/w |
|---|---|---|
| (A) | Glyceryl monoisostearate | 2.5 |
| | Isoparaffin ($C_{10}$—$C_{12}$) | 5.0 |
| | Mineral oil (light) | 5.0 |
| | Microcrystalline wax | 0.3 |
| | Acetylated lanolin | 1.0 |
| | Propyl p-hydroxybenzoate | 0.1 |
| (B) | Monosodium glutamate | 3.0 |
| | Methyl p-hydroxybenzoate | 0.2 |
| | Water | 82.7 |
| | Perfume | 0.2 |

*Procedure*

1. Heat (A) and (B) separately to about 60°C.
2. Add the aqueous phase to the oil phase and homogenize.
3. Add perfume at about 40 to 45°C.

**Discussion and critique:** This low viscosity skin cream illustrates the concept that exceptionally high internal phase ratio products can exhibit reasonable stability. Glyceryl monoisostearate probably has an HLB of about 3.5 to 4.0 and is a good w/o emulsifier. The key to the stability of this preparation is the use of an amino acid salt as an emulsion stabilizer. The practical feasibility of this approach is documented in European Patent #0,009,404 of 4/2/80, but no theoretic basis for this concept has been provided.

## 7. Mouthwash

| | | % w/w |
|---|---|---|
| 1. | Cetylpyridinium chloride | 1.00 g |
| 2. | Citric acid, USP | 1.00 g |
| 3. | Sweetener (sodium saccharin) | 0.40 g |
| 4. | Flavour oils (peppermint, eucalyptus, and clove oils) | 1.50 ml |
| 5. | Polyoxyethylene (20) sorbitan monostearate | 3.00 g |
| 6. | Alcohol, USP | 100.00 ml |
| 7. | Sorbitol solution (70%) | 200.00 g |
| 8. | Water, potable | q.s. 1000.00 ml |

*Procedure*

1. Dissolve components 1, 2 and 3 in a sufficient amount of the water and add component 6.
2. Mix components 4 and 5 and add blend slowly to the hydroalcoholic solution while stirring.
3. Add the remaining ingredients (7 and 8).

**Discussion and critique:** The level of surfactant in this preparation is surprisingly low, as is the level of alcohol. It would appear that the level of sorbitol (14%) in this product contributes to the solubilization efficacy of this mouthwash. Cetylpyridinium chloride, which is included primarily as an antimicrobial agent, also contributes to the solubilizing power of the system. This product illustrates the use of alcohol in the presence of a solubilizer to increase flavour impact (*see* under previous section “Clear Emulsions”).

## STABILITY OF EMULSIONS

Becher has indicated that the physical properties of an emulsion and its stability cannot be considered separately. Accordingly, this section is concerned with the more important physical properties of emulsions, their changes under external influences, and their relationship to emulsion stability.

## Emulsion Stability

It has already been noted that on purely thermodynamic grounds, emulsions are physically unstable. A reduction of the interfacial area by coalescence reduces the system's energy, and this process is thermodynamically favoured. For this reason, Garrett defined a stable emulsion as the one that "would maintain the same number of sizes of particles of the dispersed phase per unit volume of weight of the continuous phase. The total interfacial energy must be invariant with time to conform to this definition".

Thermodynamic stability of emulsions differs from stability as defined by the formulator or the consumer on the basis of entirely subjective judgments. Acceptable stability in a pharmaceutical dosage form does not require thermodynamic stability. If an emulsion creams up (rises) or creams down (sediments), it may still be pharmaceutically acceptable as long as it can be reconstituted by a modest amount of shaking. Similar considerations apply to cosmetic emulsions; however, in the latter, creaming is usually unacceptable because any unsightly separation makes the product cosmetically inelegant. It is important, therefore, to remember that the standard of stability depends to a large extent on the observer, since subjective observations or opinions by themselves do not suffice to define such a parameter as acceptable stability. Stability should be defined in the sense given to it by Garrett, i.e. on a purely objective basis. Shelf-life is a useful term to describe the subjective evaluation of stability.

A product's shelf life may be directly related to its "kinetic" stability. Kinetic stability means that the physicochemical properties of an emulsion do not change appreciably during a reasonably long period of time. On the other hand, "thermodynamic" stability of the type commonly postulated for solubilized systems or microemulsions is generally temperature-dependent. Thus, after the temperature of a solubilized product has been disturbed, it will eventually return to its original (and in this particular case clear or transparent) state when the temperature is returned to "normal." Thermodynamics does not and cannot predict how quickly the original (clear) state is restored.

### *Symptoms of Instability*

As soon as an emulsion has been prepared, time- and temperature-dependent

processes occur to effect its separation. During storage, an emulsion's instability is evidenced by creaming, reversible aggregation (flocculation), and/or irreversible aggregation (coalescence) (Fig. 18.22).

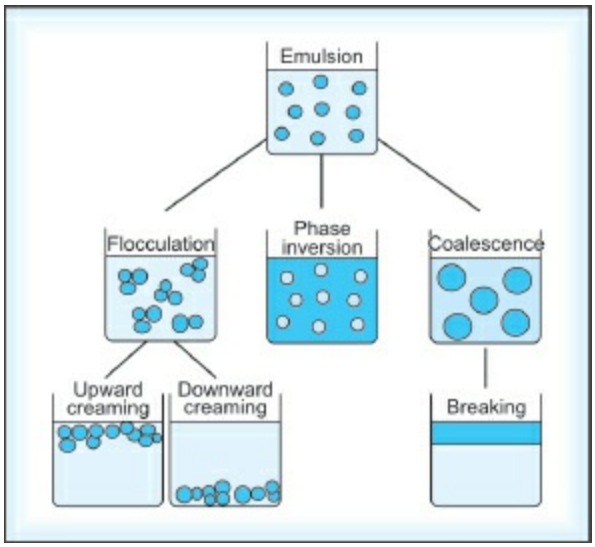



**Fig. 18.22:** Symptoms of instability problems of emulsions

## *Flocculation*

Flocculation of the dispersed phase may take place before, during, or after creaming. *It is best described as reversible aggregation of droplets of the internal phase in the form of threedimensional clusters*. Flocculation is influenced by the charges on the surface of the emulsified globules. In the absence of a protective (mechanical) barrier at the interface, e.g. if an insufficient amount of emulsifier is present, emulsion droplets aggregate and coalesce rapidly. Flocculation of emulsion droplets can occur only when the mechanical or electrical barrier is sufficient to prevent droplet coalescence. In other words, flocculation differs from coalescence primarily by the fact that the interfacial film and the individual droplets remain intact. The reversibility of this type of aggregation depends on the strength of the interaction between particles, as determined by the chemical nature of the emulsifier, the phase volume ratio, and the concentration of dissolved substances, especially electrolytes and ionic emulsifiers. For example, a 2% hexadecane-in-water emulsion stabilized with 0.09% Aerosol OT, a negatively charged surfactant, remains deaggregated in the form of single droplets, presumably as a result of repulsion between negatively charged oil droplets. An increase of the ionic

strength with electrolytes or an increase of the emulsifier concentration tends to promote flocculation. Although electrolytes are commonly used for demulsification, a modest level of electrolyte is frequently helpful in stabilizing emulsions. A typical example is the observation by Vold and Groot that sodium chloride reduces oil separation in a Nujol/water emulsion exposed to ultracentrifugation.

A high internal phase volume, i.e. tight packing of the dispersed phase, tends to promote flocculation. However, gum acacia can produce a flocculated o/w emulsion with as little as 0.002% of orange oil. Thus, it is probably safe to say that most practical o/w and w/o emulsion systems exist in a flocculated state.

Flocculation, emulsion viscosity, and shear thinning may be closely related. The viscosity of an emulsion depends to a large extent on flocculation, which restricts the movement of particles and can produce a fairly rigid network. Agitation of an emulsion breaks the particle-particle interactions with a resulting drop of viscosity, i.e. shear thinning.

## *Creaming*

Under the influence of gravity, suspended particles or droplets tend to rise or sediment depending on the differences in specific gravities between the phases. If creaming takes place without any aggregation, the emulsion can be reconstituted by shaking or mixing. The Stokes equation is most useful in gaining an understanding of the process of creaming, even though Stokes made a number of unrealistic assumptions: The equation is based on spherical particles that have essentially the same size and are separated by a distance that makes the movement of one particle independent of that of another. *In fact, creaming involves the movement of a number of heterodisperse droplets, and their movements interfere with each other and may cause droplet deformation*. If flocculation takes place, the criterion of sphericity is lost, and complex corrections for these variations must be made before Stokes law can be applied quantitatively to the behaviour of emulsions.

Despite its defects, Stokes equation is qualitatively applicable to emulsions. It shows that the rate of creaming is a function of the square of the radius of the droplet. Thus, larger particles cream much more rapidly than smaller particles. It is also apparent that the formation of larger aggregates by coalescence and/or by flocculation will accelerate creaming. The reverse is

also true, i.e. the smaller the particle size of an emulsion, the less likely it is to cream. Stokes equation predicts that no creaming is possible if the specific gravities of the two phases are equal. Adjusting the specific gravity of the dispersed phase is frequently a practical means of achieving improved emulsion stability. Finally, Stokes law shows that the rate of creaming is inversely proportional to the viscosity; this is the reason for the well known fact that increased viscosity of the external phase is associated with improved shelf-life.

## *Coalescence*

*Coalescence is a growth process during which the emulsified particles join to form larger particles*. Any evidence for the formation of larger droplets by merger of smaller droplets suggests that the emulsion will eventually separate completely or *Break*. The major factor which prevents coalescence in flocculated and unflocculated emulsions is the mechanical strength of the interfacial barrier. This is particularly true in o/w systems containing nonionic surfactants and in w/o emulsion-systems in which electrical effects are negligible. Thus, it is widely recognized that good shelf-life and absence of coalescence can be achieved by the formation of a thick interfacial film from macromolecules or from particulate solids. This is the reason a variety of natural gums and proteins are useful as auxiliary emulsifiers when used at low levels, but can even be used as primary emulsifiers at higher concentrations.

## Assessment of Emulsion Shelf-life

The final acceptance of an emulsion depends on stability, appearance, and functionality of the packaged product. The most obvious problems facing the formulator are: (1) What is acceptable emulsion shelf-life? and (2) What are the predictive indicators of shelf-life? The formulator requires unequivocal and quantitative answers to these questions.

As is true with most dosage forms, the container used for packaging an emulsion may be expected to be a source of incompatibility. Possible problems include interaction of ingredients with the container, extraction of material from the container, and loss of water and volatile ingredients through the container or the closure. For this reason, whatever the nature of the container, final evaluation of the product must be conducted in the container that will be used commercially.

No quick and sensitive methods for determining potential instability in an emulsion are available to the formulator. Instead, he is forced to wait for interminable periods at ambient conditions before signs of poor shelf-life become clearly apparent in an emulsion. To speed up his stability program, the formulator commonly places the emulsion under some sort of stress. Alternately, he may seek a test or parameter that is more sensitive for the detection of instability than mere macroscopic observations. Both approaches may be faulty. The first one may eliminate many good emulsions because excessive artificial stress has been applied. An accelerated aging test should speed up only the processes involved in instability under “normal” storage conditions. If the stress is excessive, abnormal processes may come into play. The second one may eliminate only those emulsions that are extremely poor unless the parameter correlates well with shelf life. It is therefore essential to use sound judgment and great care in setting up a meaningful stability program for a given emulsion.

### *Stress Conditions*

Stress conditions normally employed for evaluating the stability of emulsions include (1) aging and temperature, (2) centrifugation and (3) agitation.

### *Aging and Temperature*

It is routine to determine the shelf-life of all types of preparations by storing

them for varying periods of time at temperatures that are higher than those normally encountered. The Arrhenius equation, which predicts that a 10°C increase in the temperature doubles the rate of most chemical reactions, is not applicable to emulsions. The Arrhenius equation is based on the concept that the same chemical reactions take place at all temperatures albeit at different rates. It is generally recognized that in the case of emulsions, changes in temperature bring into play new reactions. It is important therefore, for the formulator to realize that exposure to unrealistically high temperatures may produce meaningless results. It is clearly established that many emulsions may be perfectly stable at 40 or 45°C, but cannot tolerate temperatures in excess of 55 or 60° even for a few hours. The varied effects of temperature changes on emulsion parameters have been discussed before: viscosity, partitioning of emulsifiers, inversion at the phase inversion temperature, and crystallization of certain lipids. In view of these problems, shelf-life cannot be predicted by studying emulsions at temperatures in excess of 50°C even for relatively short periods of time, unless there is some reason to believe that the preparation will be exposed to such a high temperature in normal handling, such as in sterilization of parenteral emulsions.

A particularly useful means of evaluating shelf-life is cycling between two temperatures. Again, extremes should be avoided, and cycling should be conducted between 4 and 45°C. This type of cycling approaches realistic shelf conditions, but places the emulsion under enough stress to alter various emulsion parameters.

The normal effect of aging an emulsion at elevated temperature is acceleration of the rate of coalescence or creaming, and this is usually coupled with changes in viscosity. Most emulsions become thinner at elevated temperature and thicken when allowed to come to room temperature. This thickening can be excessive if the emulsion is not agitated during the cooling cycle; sometimes, the low viscosity can be “frozen” into the emulsion if it is chilled rapidly. In view of these variations, the formulator must evaluate each emulsion separately and on the basis of his own personal experience. Freezing can damage an emulsion more than heating, since the solubility of emulsifiers, both in the lipid and aqueous phases, is more sensitive to freezing than to modest warming. In addition, the formation of ice crystals develops pressure that can deform the spherical shape of emulsion droplets.

## *Centrifugation*

It is commonly accepted that shelf life under normal storage conditions can be predicted rapidly by observing the separation of the dispersed phase due to either creaming or coalescence when the emulsion is exposed to centrifugation. Stokes law shows that creaming is a function of gravity, and an increase in gravity therefore accelerates separation. Becher indicates that centrifugation at 3,750 rpm in a 10 cm radius centrifuge for a period of 5 h is equivalent to the effect of gravity for about one year. The modest speed suggested by Becher is probably reasonable. On the other hand, ultracentrifugation at extremely high speeds (approximately 25,000 rpm or more) can be expected to cause effects that are not observed during normal aging of an emulsion. Ultracentrifugation of emulsions creates three layers: a top layer of coagulated oil, an intermediate layer of uncoagulated emulsion, and an essentially pure aqueous layer. Rapid formation of a clear oily layer is the first clue to “abnormal” phenomena taking place during ultracentrifugation. Groot and Void showed how the rate of oil formation in a Nujol: water: sodium dodecyl sulfate (50:50:0.2%) emulsion depends on the rate of centrifugation. Separation was extremely rapid at 56,000 rpm, somewhat slower at about 40,000 rpm, and no oil was separated after *2Vi* h of centrifugation at approximately 11,000 rpm. These findings suggest that the force of ultracentrifugation does not cause oil separation until it is high enough to break or rupture the absorbed layer of emulsifier that surrounds each droplet. It is concluded that centrifugation, if used judiciously, is an extremely useful tool for evaluating and predicting the shelf-life of emulsions.

## *Agitation*

It is a paradigm of emulsion science that the droplets in an emulsion exhibit Brownian movement. In fact, it is believed that no coalescence of droplets takes place unless droplets impinge upon each other owing to their Brownian movement. Simple mechanical agitation can contribute to the energy with which two droplets impinge upon each other. It is rarely appreciated how useful the evaluation of an emulsion by agitation at or near room temperature can be. It was already noted that excessive shaking of an emulsion or excessive homogenization may interfere with the formation of an emulsion. As a corollary, agitation can also break emulsions. A typical case, well

known to all, is the manufacture of butter from milk. Some clear microemulsions become cloudy upon short agitation in a blender due to coalescence of particles. Similarly, conventional emulsions may deteriorate from gentle rocking on a reciprocating shaker. This is related, in part to the impingement of droplets and in part to reduction of viscosity of a normally thixotropic system.

## *Chemical Parameters*

The need for chemical stability of the components of emulsions has already been noted. A typical problem encountered in the presence of polyethylene glycols or derivatives of polyethylene glycol is their propensity towards autoxidation. This phenomenon can cause the formation of undesirable odours, of acidic components, and of all types of oxidative byproducts. The instability of non-ionic esters leading to hydrolytic degradation may result in changes in the dielectric constant of the emulsion. This phenomenon parallels the observations of physical instability and has been attributed to the formation of stearic acid from, for example, polysorbate 80.

## *Physical Parameters*

The most useful parameters commonly measured to assess the effect of stress conditions on emulsions include (1) phase separation, (2) viscosity, (3) electrophoretic properties and (4) particle size analysis and particle count.

## *Phase Separation*

The rate and extent of phase separation after aging of an emulsion may be observed visually or by measuring the volume of the separated phase. It is important to differentiate between creaming and coalescence, since the means of correcting these defects are different.

Relatively little quantitative information is available concerning oil separation in practical systems in the absence of centrifugation. A study of mineral oil-water emulsions stabilized with either polyoxyethylene sorbitan monooleate or sodium lauryl sulfate showed that the amount of coalescence observed at room temperature depends on the concentration of emulsifier. At low levels, i.e. below 0.1%, visible coalescence of the oil phase occurs after only one month’s storage. When the concentration of emulsifier is raised to 2 or 5%, the amount of visible coalescence is negligible even after two years

storage. Two additional points are noteworthy: The two emulsifiers perform similarly, and coalescence can be observed quite early.

A particularly simple means of determining phase separation due to creaming or coalescence is apparently so trivial that it has evidently not been described in the literature. It involves withdrawing small specimens of the emulsion from the top and the bottom of the preparation after some period of storage and comparing the composition of the two samples by appropriate analysis of water content, oil content, or any suitable constituent.

## *Viscosity*

Although the viscosity of an emulsion is an essential performance criterion, its use for shelf-life studies is not concerned with the absolute values of viscosity, but with changes in viscosity during aging. The number of instruments available for the determination of consistency/viscosity is overwhelming. Since emulsions are generally non-Newtonian and since the instrument should have universal utility, it is best to avoid capillary and falling-sphere viscometers. Viscometers of the cone-plate type are particularly useful for emulsions, but instruments utilizing coaxial cylinders are the easiest to use. In the case of fairly viscous materials, the use of a penetrometer is often helpful in detecting changes of viscosity with age.

As a rule, globules in freshly prepared w/o emulsions flocculate quite rapidly. Consequently, the viscosity drops quickly and continues to drop for some time (5 to 15 days at room temperature), and then remains relatively constant. o/w emulsions behave quite differently. In this case, globule flocculation causes an immediate increase in viscosity. After this initial change, almost all emulsions show changes in consistency with time, which follow a linear relationship when plotted on a log-log scale. The complete absence of a slope (no change in viscosity with age) is believed to be ideal, although most acceptable systems exhibit modest increases of viscosity between 0.04 and 400 days (Fig. 18.23). Other emulsions exhibit much more drastic and sudden non-linear increases in viscosity after two or three months' aging; such a behaviour is frequently followed by a drop in viscosity, which is probably associated with phase separation.

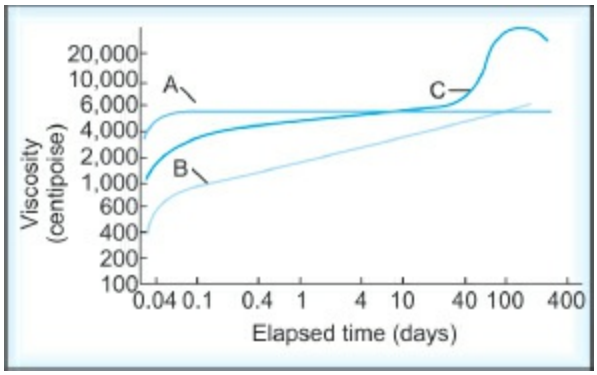



**Fig. 18.23:** Simplified aging curves of emulsions. Ideal shelf-life: (A) typical shelf-life, (B) and questionable shelf-life (C)

A practical approach for the detection of creaming or sedimentation, before it becomes visibly apparent, utilizes the helipath attachment of the Brookfield viscometer. As a result of emulsion separation, the descending rotating spindle meets varying resistance at different levels and registers fluctuations in viscosity. For example, lotion A in Fig. 18.24 contains solids suspended in an emulsion, and the high viscosity near the top is due to non-wetted solid and creamed emulsion; the high viscosity at lower levels is due to sedimented particles. The addition of polyoxyethylene sorbitan monooleate and methylcellulose (lotion B) yields a much more uniform viscosity pattern after eight weeks of storage.

The collection of such data is certainly useful; however, it is apparent from these two examples that it is impossible to predict longterm viscosity behaviour from the data collected during the first few weeks of storage after an emulsion has been prepared. According to Sherman, the best way of using viscosity determinations for the prediction of shelf-life is to relate them to changes in particle size. As a rule, a decrease in viscosity with age reflects an increase of particle size due to coalescence and is indicative of a poor shelf-life.

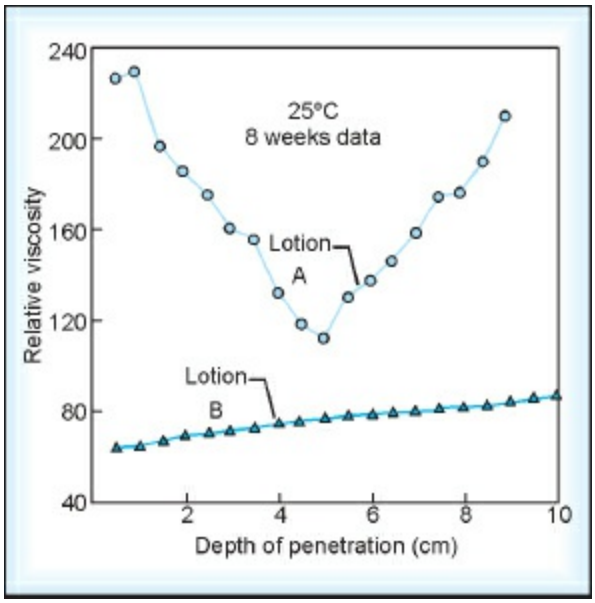



**Fig. 18.24:** Use of the helipath in the evaluation of creaming and sedimentation. Lotion B is the same as A except that 0.1% polysorbate 80 and 0.25% methyl cellulose have been added

Viscosity measurements should be carried out in undisturbed containers to avoid changes due to previous stresses. To avoid artefacts, replicate samples should be prepared in advance, or special studies must be carried out to determine the time necessary for a disturbed emulsion to recover its original viscosity. It must always be remembered that the act of measuring the viscosity disturbs the system. Also, the viscosity of emulsions at several different shear rates must be determined to obtain a clear picture of the rheologic behaviour of an emulsion. Measurements at low shear rates frequently reflect aspects of flocculation. High shear rates overcome the attractive forces between particles and can result in a precipitous loss of viscosity.

## *Electrophoretic Properties*

The zeta potential of emulsions can be measured with the aid of the moving boundary method or, more quickly and directly, by observing the movement of particles under the influence of electric current. The zeta potential is especially useful for assessing flocculation since electrical charges on particles influence the rate of flocculation. If the instability is due to coalescence, the determination of the surface charges of particles may not be relevant for the prediction of shelf-life.

The measurement of electrical conductivity has been claimed to be a powerful tool for the evaluation of emulsion stability shortly after preparation. The electrical conductivity of o/w or w/o emulsions is determined with the aid of Pt electrodes (diameter 0.4 mm; distance 4 mm) microamperometrically to produce a current of about 15 to 50 pA. Measurements are made on emulsions stored for short periods of time at room temperature or 37°C. Reportedly, the conductivity depends on the degree of dispersion. o/w preparations with fine particles exhibit low resistance; if the resistance increases, it is a sign of oil droplet aggregation and instability. A fine emulsion of water in a w/o product does not conduct current until droplet coagulation, i.e. instability occurs.

## *Particle Size Number Analysis*

Changes of the average particle size or of the size distribution of droplets are important parameters for evaluating emulsions. Particle size analysis may be carried out by a number of methods, each giving a somewhat different average for heterodisperse systems. For example, microscopic measurements of the apparent diameter give an average value dependent on the number of particles of each size. On the other hand, some electronic counting devices measure particle volume, and since the volume of a sphere is $nd^3/6$, they give greater weight to larger particles when volume is converted to diameter. Such counting devices, e.g. the Coulter counter, also require that the emulsion be diluted, sometimes with a conducting electrolyte. The changes caused by these steps as well as changes caused by sampling, even for a microscopic assessment of particle size, often make the counted sample no longer representative of the bulk emulsion.

Light scattering and related reflectance relationships have been used for particle size determinations. Thus, the change of reflectance at wavelength at which the coloured internal phase partially absorbs the incident light has been found to be inversely proportional to a power of the particle diameter.

The utility of particle size for predicting or interpreting emulsion shelf life is somewhat doubtful. Two studies utilizing fairly stable emulsions have shown that the initial increase in particle size is rather rapid, but is followed by a much slower change. Almost no change in particle size has been noted, even in the case of emulsions that show appreciable coalescence due to a low level of emulsifier. One would expect that particle size, the number of

particles, the droplet surface area, or the droplet volume should vary linearly with time. Hill and Knight claim good correlation with experimental data by plotting the total surface area of all droplets (Σ) in accordance with the following equation: $\Sigma = at + b$, where, a and b are constants and *t* equals time.

## *Practical Recommendations for Shelf-life Predictions*

The preceding discussion has pointed out that surprisingly little evidence exists to suggest that instability under stress can be related to normal shelf-life. It is most important, therefore, to set up a realistic stability program to assess the shelf-life of emulsions.

A typical test program for an "acceptable" emulsion (in the temperate zone) might establish the following: The emulsion should be stable with no visible signs of separation for at least 60 to 90 days at 45 or 50°C, 5 to 6 months at 37°C, and 12 to 18 months at room temperature. Similarly, there should be no visible signs of separation after one month's storage at 4°C and preferably after two or three freeze-thaw cycles between −20 and +25°C. An emulsion should survive at least six or eight heating/cooling cycles between refrigerator temperature and 45°C with storage at each temperature of no less than 48 h. A stable emulsion should show no serious deterioration by centrifuging at 2,000 to 3,000 rpm at room temperature. The emulsion should not be adversely affected by agitation for 24 to 48 hours on a reciprocating shaker (approx. 60 cycles per minute at room temperature and at 45°C).

During the testing period just described, the samples stored at various conditions should be observed critically for separation, and in addition, monitored at reasonable time intervals for the following characteristics:

- Change in electrical conductivity
- Change in light reflection
- Change in viscosity
- Change in particle size
- Change in chemical composition

In addition to these physical measurements, a shelf-life program for emulsions should include testing of the emulsion for microbiologic contamination at appropriate intervals. It should be remembered that the distribution of emulsifiers in a freshly prepared emulsion is different from distribution of one that has been aged for several months at 45°C. As a result,

an emulsion’s resistance to microbial contamination may be affected by redistribution or micellization of the preservative.

# 19:
# Semisolids

Pharmaceutical semisolid preparations include ointments, pastes, cream emulsions, gels and rigid foams. Their common property is the ability to cling to the surface of application for reasonable duration before they are washed or worn off. This adhesion is due to their plastic rheologic behavior, which allows the semisolids to retain their shape and cling as a film until acted upon by an outside force, in which case they deform and flow.

**Ointments**, in general, are composed of fluid hydrocarbons meshed in a matrix of higher-melting solid hydrocarbons. While most ointments are based on mineral oil and petrolatum, there are alternative types also. Polyethylene can be incorporated into mineral oil to yield a plastic matrix (e.g. Plastibase, manufactured by Squibb). Mixtures of polyethylene glycols can yield products of ointment consistency that are water-soluble. Most ointments are prepared by melting the components together. Drugs or other components are added in the fluidized state. If the solids are insoluble and to be suspended, the system is put through a milling process (a colloid mill, homogenizer, or ultrasonic mixer) so that the solids are fully dispersed.

**Pastes** are basically ointments into which a high percentage of insoluble solids has been added. They are valuable as protective barriers on the skin, such as for treating diaper rash or protecting the face and lips from the sun. Pastes are usually prepared by incorporating a solid directly into a congealed system by levigation with a portion of the base to form a paste-like mass. The remainder of the base is added with continued levigation until the solids are uniformly dispersed in the vehicle.

**Creams** are semisolid emulsion systems with opaque appearances, as contrasted with translucent ointments. Their consistency and rheologic character depend on whether the emulsion is a water-in-oil or oil-in-water type and on the nature of the solids in the internal phase. The subject of emulsions is treated in Chapter 18.

**Gels** are semisolid systems in which a liquid phase is constrained within a

threedimensional polymeric matrix (consisting of natural or synthetic gums) in which a high degree of physical (or sometimes chemical) cross-linking has been introduced. The polymers used to prepare pharmaceutical gels include the natural gums tragacanth, pectin, carrageen, agar and alginic acid, synthetic and semisynthetic materials such as methylcellulose, hydroxyethylcellulose, carboxymethylcellulose, and the Carbopols, which are synthetic vinyl polymers with ionizable carboxyl groups. Gels are prepared by a fusion process or by procedure necessitated by the gelling characteristics of the gellant.

The bulk of these semisolid preparations are applied to the skin, where they usually serve as vehicles for topically applied drugs, as emollients, or as protective or occlusive dressings. A lesser portion of topical semisolid dosage forms are applied to mucous membranes, such as rectal tissue, buccal tissue, vaginal mucosa, urethral membrane, external ear lining, nasal mucosa, and cornea. The mucous membranes permit more ready access to the systemic circulation, whereas normal skin is relatively impenetrable. The emphasis of this chapter is on the skin and on dermatologicals, but the general concepts and rationale apply to all semisolid topical therapy.

# THEORETICAL CONSIDERATION

## Skin

The skin is a large multilayered organ that in the average adult weighs about eight pounds, excluding fat. It covers a surface exceeding 20,000 $cm^2$ and has varied functions and properties. The skin serves as a barrier against physical and chemical attack. Some materials, such as nickel ions, mustard gas, and the oleoresins from Rhus toxicodendron, commonly known as poison ivy, can penetrate the barrier, but most substances cannot. The skin acts as a thermostat in maintaining body temperature, shields the body from invasion by microorganisms, protects against ultraviolet rays, and plays a role in the regulation of blood pressure.

Anatomically, the skin has many histologic layers, but in general, it is described in terms of three tissue layers: the epidermis, the dermis, and the hypodermis layer. Figure 19.1 represents an idealized section of the skin, showing the glands, hair follicles, nerves, blood vessels, and other skin accessories.

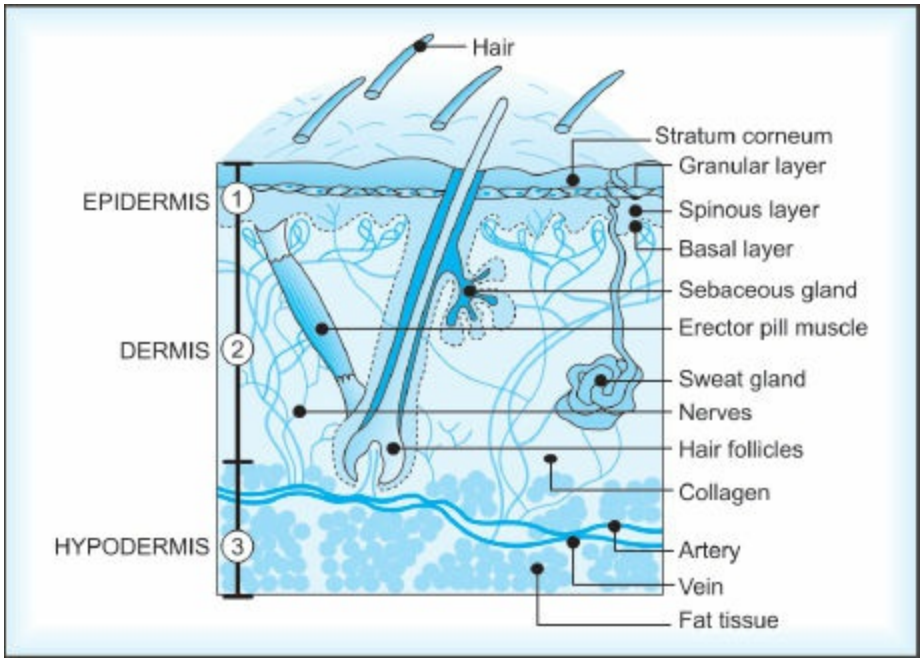



**Fig. 19.1:** Stratified organization of the skin

### *Epidermis*

******ebook converter DEMO Watermarks*******

The epidermis is approximately 50–150 μm thick and consists largely of constantly renewing, outward moving cells called keratinocytes. Apart from these cells, most of the antigen-presenting Langerhans' cells are located in the epidermis. The outermost layer of the epidermis is the stratum corneum or horny layer, which consists of compacted, dead, keratinized cells in stratified layers with a density of 1.55. Because of the dense nature of the stratum corneum, values of diffusion coefficients in this tissue are a thousand or more times smaller than in any other skin tissue, which results in higher resistance and general impenetrability.

The stratum corneum is the rate-limiting barrier that restricts the inward and outward movement of chemical substances. Structurally, the stratum corneum is a heterogeneous tissue composed of flattened keratinized cells, the outer layers of which are less densely packed than those adjacent to the underlying granular layer. The stratum corneum exhibits regional differences in thickness over the body. It is as thick as several hundred micrometers on the palms of the hand and soles of the feet in an adult, but over most of the body it is about 10 μm thick when dry, increasing to about 40 to 50 μm when fully hydrated.

There is limited knowledge of the chemical composition of the barrier. The main cellular components are the proteins, lipid, and water, combined into an ordered structure. The approximate composition in the dry state is 75 to 85% protein, 15 to 20% lipid, and 15% water. Although the surface lipids offer little resistance to the passage of compounds, studies of the removal of lipids from the cutaneous surface indicate that they participate in epidermal water function. Barrier function is restored when the extracted lipids are returned to the skin, which suggests variations in biologic membrane permeability, depending largely on the specific nature or distribution of the lipid contained in the cell membrane.

Beneath the stratum corneum are the metabolically active layers of the epidermis. The identifiable strata, top to bottom are: (a) *Stratum granulosam* —the granular layer, (b) *Stratum spinosum*—the multicellular spinous or prickle layer, (c) *Stratum germinativum*—the basal or germinal layer that lies right above the dermis. In some histological displays a fourth, upper transitional and translucent layer—*Stratum lucidum*, is also distinguishable. Epidermal cells start their mitotic journey upward to the surface; the cells flatten and shrink as they slowly the from lack of oxygen and nutrition.

## *Dermis*

The next distinctive histologic layer as shown in Fig. 19.1 is the dermis or corium, which is approximately one eighth of an inch thick and constitutes the main mass of the skin. The dermis essentially consists of about 80% of protein in a matrix of mucopolysaccharide “ground substance”.

Contained and supported within the dermis are numerous blood vessels, lymphatics, and nerves, as well as the epidermal appendages such as the hair follicles, sebaceous glands, and sweat glands. Hair follicles are distributed over the entire skin surface with the exception of the soles of the feet, the palms of the hand, the red portion of the lips, and selected portions of the sex organs. Each hair follicle is associated with one or more sebaceous glands, which are outgrowths of epithelial cells. The fractional area of the skin surface occupied by the hair follicles has been estimated to be roughly 1/1000 of the total surface. The sweat glands are divided into the eccrine and apocrine types. They are widely distributed over the surfaces of the body. The eccrine glands are particularly concentrated in the palms and soles. The principal function of the glands is for heat control, as they secrete a dilute salt solution. The apocrine glands are found in the axillae (armpits), in anogenital regions, and around nipples. They are coiled tubular glands about ten times larger than eccrine glands and extend entirely through the dermis and well into the subcutaneous layer.

## *Hypodermis*

The dermis rests on the hypodermis (subcutis) which is composed of loose fatty connective tissue. Its thickness varies considerably over the surface of the body as well as between individuals

## Percutaneous Absorption

The usual object of dermatologic drug therapy is to produce a desired therapeutic action at specific sites on the epidermal tissue. While certain topical drugs such as emollients, antimicrobials, and deodorants act primarily on the surface of the skin, the target area for most dermatologic disorders lies in the viable epidermis or upper dermis. This requires diffusive penetration of the skin or percutaneous absorption. The events governing percutaneous absorption are represented in Fig. 19.2.

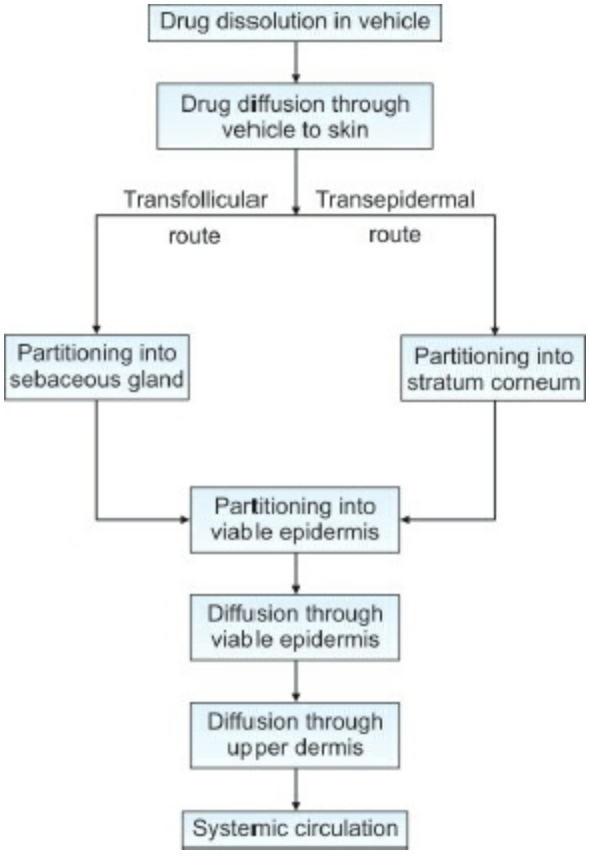



**Fig. 19.2:** Events governing percutaneous absorption

## Routes of Penetration

When a drug system is applied topically, the drug diffuses out of its vehicle onto the surface tissues of the skin. There are three potential portals of entry: through the follicular region-(transfollicular), through the sweat ducts, or through the unbroken stratum corneum between the appendages (transepidermal). There is little convincing evidence that eccrine sweat glands play any significant role in cutaneous permeability. Material may enter the ducts, and even the glands, but there appears to be no penetration from these areas to the dermis.

For substances absorbed by the transepidermal route, penetration is fairly rapid, although slower than intestinal tract absorption, and is almost always accompanied by some degree of pilosebaceous penetration as well. For substances that are absorbed through both pathways, the transepidermal route is the principal portal of entry because of the total, relatively small, absorbing surface offered by the pilosebaceous units. The epidermis presents a surface area 100 to 1000 times greater than the other routes of absorption. The appendages, sweat glands, and hair follicles are scattered throughout the skin in varying numbers, but are comparatively sparse; their total cross-sectional area is probably between 0.1 and 1.0% of the skin area.

The particular route a substance may take and the relative importance of one in contrast with the other, depends almost entirely on the physicochemical properties of the drug and the condition of the skin. Under the appropriate conditions, each of the contending routes of permeability may change and be the overwhelmingly dominant one. In particular, the transient diffusion that occurs shortly after the application of a substance to the surface of the skin is shown to be potentially far greater through the appendages than through the matrix of the stratum corneum. After steady-state diffusion has been established, the dominant diffusion mode is probably no longer intra-appendageal, but occurs through the matrix of the stratum corneum. The intact stratum corneum provides the main barrier; its ‘brick and mortar’ structure is analogous to a wall (Fig. 19.3). The corneocytes of hydrated keratin comprise the ‘bricks’, embedded in a ‘mortar’, composed of multiple lipid bilayers of ceramides, fatty acids, cholesterol and cholesterol esters. These bilayers form regions of semicrystalline domains. Most molecules penetrate through skin via this intercellular microroute and therefore many

enhancing techniques aim to disrupt or bypass its elegant molecular architecture.

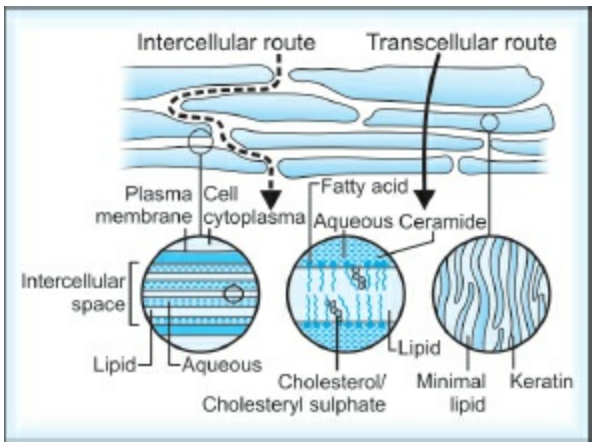



**Fig. 19.3:** Simplified diagram of stratum corneum and routes of drug penetration

Flux through shunts is difficult to measure experimentally, except possibly through hair. The recognition of transient diffusion, occurring primarily via follicles and ducts, and steady-state diffusion, occurring primarily through the intact stratum corneum, results in a considerably more self-consistent and orderly treatment of the process of percutaneous absorption.

Once a substance passes through the stratum corneum, there is apparently no significant further hindrance to penetration of the remaining epidermal layers and corium; there is then a ready entry into the circulation via the capillaries. The concentration gradient essentially ends in the dermal layer at the beginning of the circulation. The systemic circulation acts as a reservoir or “sink” for the drug. Once in the general circulation, the drug is diluted and distributed rapidly with little systemic build-up.

Diffusion through the horny layer is a passive process. There is little evidence to support specialized active transport systems for cells of the stratum corneum. The passive process is affected only by the substance being absorbed, by the medium in which the substance is dispersed, and by ambient conditions. On the other hand, percutaneous absorption is a more complicated process, of which epidermal diffusion is the first phase, and clearance from the dermis is the second. The latter depends on effective blood flow, interstitial fluid movement, lymphatics, and perhaps other factors that

combine with dermal constituents. Figure 19.4 summarizes some ways for circumventing the stratum corneum barrier.

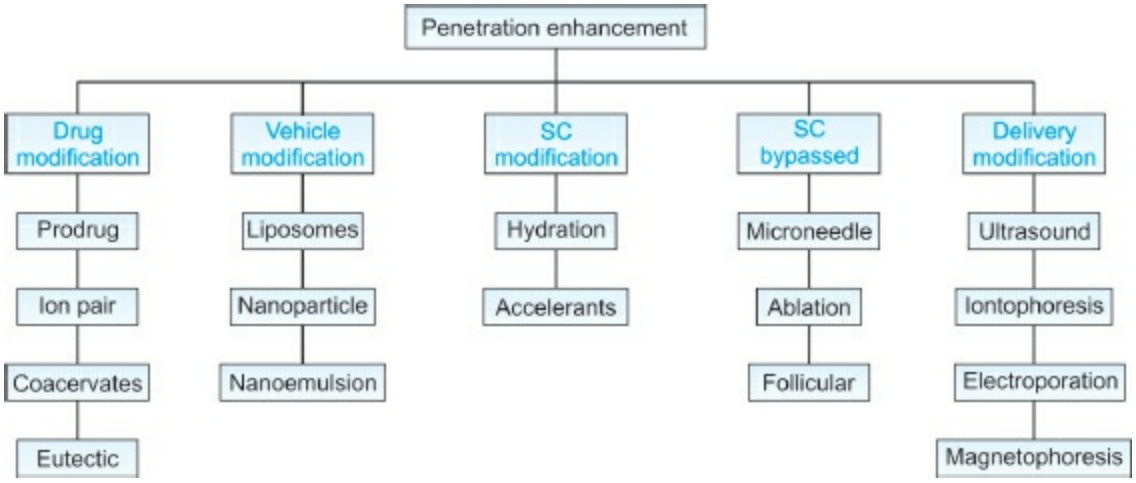



**Fig. 19.4:** Methods for increasing penetration of drug through stratum corneum

## Factors in Skin Penetration

The factors that influence skin penetration are essentially the same as those for gastrointestinal absorption, with the rate of diffusion depending primarily on the physicochemical properties of the drug and only secondarily on the vehicle, pH, and concentration. Differing physiologic variables involve the condition of the skin, i.e. whether it is intact or injured, the skin age, the area of skin treated, the thickness of the skin barrier phase, the species variation, and the skin moisture content.

The principal physicochemical factor in skin penetration is the hydration state of the stratum corneum, which affects the rate of passage of all substances that penetrate the skin. Hydration results from water diffusing from underlying epidermal layers or from perspiration that accumulates after application of an occlusive vehicle or covering on the surface. Under occlusive conditions, the stratum corneum is changed from a tissue that normally contains little water (5 to 15%) to one that contains as much as 50% water. The clinical importance of hydration can be found in the use of occlusive plastic film in steroid therapy. Here, the prevention of water loss from the stratum corneum and the subsequent increased water concentration in this skin layer apparently enhances the penetration of the steroid. The temperature of the skin and the concentration of drug play significant roles, but they are secondary to that of hydration.

The solubility of a drug determines the concentration presented to the absorption site, and the water/lipid partition coefficient influences the rate of transport. An inverse relationship appears to exist between the absorption rate and the molecular weight. Small molecules penetrate more rapidly than large molecules, but within a narrow range of molecular size, there is little correlation between the size and the penetration rate. Materials of higher molecular weight also show variable penetration. Very large molecules, such as proteins and polysaccharides, go through poorly, if at all.

# FORMULATION CONSIDERATION

## Drug Selection

The simplest approach is to choose a drug from a congeneric series or pharmacological class with the correct physicochemical properties to translocate across the barrier at an acceptable rate. A useful way to consider factors affecting drug permeation rate through stratum corneum is via the simple equation for steady state flux (Eq. 1). If we plot the cumulative mass of diffusant, m, passing per unit area through the membrane, at long times the graph approaches linearity and its slope yields the steady state flux, *dm/dt*, as described by equation:

$$\frac{dm}{dt} = \frac{DC_0K}{h} \qquad \ldots (1)$$

where, $C_0$ is the constant concentration of drug in donor solution, $K$ is the partition coefficient of solute between membrane and bathing solution, $D$ is the diffusion coefficient and $h$ is thickness of membrane. From Eq. (1), we can deduce the ideal properties of a molecule penetrating stratum corneum. These are as follows:

- Low molecular mass, preferably less than 600 *Da*, when $D$ tends to be high
- Adequate solubility in oil and water—so that the membrane concentration gradient may be high ($C_0$ is large). Saturated solutions (or suspensions having the same maximum thermodynamic activity) promote maximum flux in equilibrium systems.
- High but balanced (optimal) $K$ (too large may inhibit clearance by viable tissues).
- Low melting point, correlating with good solubility as predicted by ideal solubility theory.

  An alternative form of Eq. (1) uses thermodynamic activities when:

$$\frac{dm}{dt} = \frac{aD}{\gamma h} \qquad \ldots (2)$$

where, $a$ is the thermodynamic activity of drug in its vehicle and $\gamma$ is the effective activity coefficient in the skin barrier. For maximum penetration rate, the drug should be at its highest thermodynamic activity. Dissolved

molecules in saturated solution are in equilibrium with pure solid and such molecules posses maximum activity. Thus all vehicles containing drug as a finely ground suspension should produce the same penetration rate, provided that the system behaves ideally i.e. D, $\gamma$ and $h$ remain constant. Supersaturated solutions (i.e. nonequilibrated systems) may arise, either by cosolvent evaporating on the skin or via formulation design. The theoretical maximum flux may then increase manyfold. Ideality is difficult to maintain, as most topical vehicles interact to some extent with the horny layer.

## *Insoluble Drugs*

Insoluble drugs must be uniformly dispersed throughout the vehicle to ensure homogeneity of the product. The solid must be impalpable to the touch; otherwise, grittiness results. Particles less than 74 microns in size, equivalent to the mesh openings in a 200-mesh sieve in the US Standard Sieve series, are impalpable to most people. Milling to a finely divided state provides more surface area for contact with the dermal site and increases the rate of dissolution of poorly soluble substances.

Some powders do not disperse uniformly, but tend to aggregate in the base, whereas others present no difficulties even though the particle size is the same. The difference may be due to the electrically charged surface condition of the particles after milling. Aggregation of particles becomes a problem for those that are 5 microns or smaller in size. For particles below 0.5 microns in size, the dispersion problems increase exponentially. Different powdered substances show similar problems of aggregation in the submicron size.

Many drug substances used in topical preparations (e.g. prednisolone; fluoro-cortisone acetate) exist in several polymorphic states. Compounds that exist in different crystalline forms at room temperature possess varying amounts of free energy or thermodynamic activity. The physiologic activity and availability of a drug substance often are directly related to its thermodynamic activity and the choice of the proper crystalline form for use in the semisolid is vitally important. Following its incorporation into the semisolid, the maintenance of the selected polymorphic form in the semisolid is of equal concern. The components of the vehicle and the method of preparation of the semisolid dosage form affect the stability of the polymorphic form.

Charged molecules do not readily penetrate stratum corneum. One technique forms a lipophilic ion pair, by adding an oppositely charged species. The ion pair partitions into the stratum corneum lipids, diffuses to the aqueous viable epidermis, there to dissociate into its charged species, which partition into the epidermis and diffuse onward. Generally, enhancement is modest (two-fold). Complex coacervation is the separation of oppositely charged ions into a coacervate oil phase, rich in ionic complex. The coacervate partitions into stratum corneum, where it behaves as ion pairs, diffusing, dissociating and passing into viable tissues.

## Vehicles/Semisolid Bases

The vehicles used for a pharmaceutical semisolids are called as bases and they differ from that used for a cosmetic because for a cosmetic, penetration into the skin is not desired. Penetration or protection is desired in a pharmaceutical semisolid, and its cosmetic effect or appearance on the skin is less important. A well-formulated pharmaceutical semisolid should be both therapeutically effective and cosmetically appealing, with the major effort in the medical direction.

The therapeutic preparations included in the semisolids classification are products intended for application to the skin, scalp, and certain body orifices. These preparations include ophthalmic ointments, nasal jellies, gels, and sterile lubricants for surgical use. In this chapter, however, attention is given to those dosage forms that are used in the prevention or treatment of skin disease.

The solubility and stability of the drug in the base, as well as the nature of the skin lesion, determine the choice of the semisolid vehicle. The USP recognizes four classes of semisolid bases under the general classification of ointment: Oleaginous bases, absorption bases (anhydrous form and emulsion form), water-removable bases and water-soluble bases.

### *Oleaginous Bases*

These bases are usually anhydrous and consist of substances such as hydrocarbons, vegetable oils, silicones and certain synthetic esters, either alone or in combination. These bases are characterized by (i) low capacity to absorb water (emollient effect), (ii) exert high occlusiveness by forming water impermeable layer on skin, (iii) greasy and thus difficult to remove from skin, (iv) prolonged contact with skin and (v) poor aesthetic appeal.

**Hydrocarbons:** Hydrocarbons include soft, liquid and hard paraffins.

*Petrolatum (soft paraffin):* The most commonly used raw material in ointment vehicles is petrolatum because of its consistency, its bland and neutral characteristics, and its ability to spread easily on the skin. Petrolatum is a complex mixture of semisolid hydrocarbons, containing aliphatic, cyclic, saturated, unsaturated, branched, and unbranched substances in varying proportions. Although extensively used for more than 85 years, petrolatum

still has broad physical and chemical specifications in the USP. Wide density and melting point ranges, as well as variation in chemical composition, are permitted in the official compendia throughout the world. Petrolatum is available in the form of a short or long “fiber.” The type of fiber possessed by the petrolatum is usually determined by dipping the index finger into the petrolatum sample and then withdrawing it slowly. The long fiber type tends to form a transparent continuous film or thread joining the finger and the sample. The short fiber variety ruptures easily and does not exhibit this film. The long fiber petrolatum is preferred for an occlusive dressing because of the continuous film it forms over the surface of the skin. These bases are difficult to wash off the skin and may be used as occlusive coverings to inhibit the normal evaporation of moisture from the skin. A thin film of petrolatum produces a sensation of warmth on the skin because the insensible moisture does not evaporate. Very little water can be incorporated into these greasy bases without the addition of other substances. Addition of wax to petrolatum results in a base with firm consistency as exemplified by White ointment which is a petrolatum with 5% beeswax.

*Mineral oil (liquid paraffin):* It is obtained from petroleum, as is petrolatum, by collection of a particular viscosity-controlled fraction. It is produced in many viscosity and specific gravity ranges. Mineral oil, USP is heavy mineral oil (specific gravity, 0.848–0.905), whereas mineral oil, NF is light mineral oil (specific gravity, 0.818–0.880). The lower viscosity oils are preferred for semisolids, since they are less tacky and greasy. Blending of mineral oil with petrolatum can produce ointments of various consistencies with less drag or resistance to spreading. *Soft petrolatum* base is a mixture of 90% white petrolatum with 10% mineral oil, NF. By increasing the quantity of the mineral oil in the mixture to upto 25%, a *gelled petrolatum* base is obtained that has a gel like consistency.

*Hydrocarbon waxes (hard paraffin):* Hydrocarbon waxes frequently are employed in the manufacture of creams and ointments to increase the viscosity of mineral oil in order to prevent its separation from an ointment. Paraffin wax is obtained from petroleum and is available in a variety of melting points ranging from 35 to 75°C. Ozokerite is a mined wax with a melting point range of 65 to 75°C and consists of a mixture of saturated hydrocarbons ranging in carbon content from $C_{35}$ to $C_{55}$. Another wax that is often used is ceresin, which is a mixture of ozokerite and paraffin wax. Its

melting point varies, depending on the paraffin wax content. Ozokerite and ceresin possess the property of retaining oils within a matrix-like structure without the sweating or oozing of the oils.

Synthetic waxes have been developed from vegetable oils and naturally occurring waxes by a process of hydrogenation and catalytic splitting that involves long $C_{18}$-$C_{36}$ hydrocarbon chains. Like all true waxes, the synthetic waxes exhibit thermoplastic, crystalline properties and are not pure chemical compounds but complex mixtures of mainly long chain saturated aliphatic chemical entities. The synthetic waxes are chemically closely related to the naturally occurring waxes in that they contain long chain wax fatty acids, but are not considered to be direct replacements for them. However, they may be used in conjunction with or can replace the natural waxes in some formulations to achieve certain desired properties. *Synchrowaxes*, the brand name of a series of such waxes, have unique gelling characteristics that may be used in formulating synthetic petrolatums with occlusive properties to help moisturize the skin without the inelegant properties of natural petrolatum.

**Vegetable oils:** Vegetable oils such as peanut oil, almond oil, sesame oil, and olive oil are mono-, di-, and triglycerides of mixtures of unsaturated and saturated fatty acids. Trace metal contaminants in the oils may catalyze oxidation reactions that can be prevented by the addition of antioxidants, such as butylated hydroxyanisole, butylated hydroxytoluene or propyl-gallate, and by the addition of metal chelating agents such as salts of ethylenediamine tetraacetic acid. Antioxidants may produce problems of drug compatibility or dermal sensitivity in some patients. The exact chemical composition of a particular vegetable oil varies from lot to lot because of its natural origins. Its composition depends on the climatic conditions, the soil, the amount of rainfall during the growth of the vegetable crop, and the storage conditions of the harvested crop and the oil. The trend toward the use of synthetic material as a substitute for vegetable oil is evident in the literature and suppliers catalogs. Esters of fatty acid (stearic acid, palmitic acid) with alcohols such as glycerol, isopropyl and butyl alcohol are used as substitutes for vegetable oils in ointment bases. ***Oily silicones*** such as dimethicone and cyclomethicone have properties similar to hydrocarbon bases and are also employed as constituents of hydrophobic ointments and barrier creams because of their water-repellent properties.

## *Absorption Bases*

These bases are usually anhydrous and they do absorb aqueous solutions and can be considered water-in-oil emulsions. The bases do not absorb water on contact, but with sufficient agitation. The absorption bases are formed by the addition of substances miscible with hydrocarbons and possessing polar groupings, such as the sulfate, sulfonate, carboxyl, hydroxyl, or an ether linkage. Lanolin, lanolin isolates, cholesterol, lano-sterol and other sterols, acetylated sterols, or partial esters of polyhydric alcohols (e.g. sorbitan monosterate or monooleate) may be added to make the hydrocarbon bases hydrophilic. Such hydrophilic mixtures have been known as “absorption bases,” although the term “absorption” is a misnomer.

The absorption bases are of two types: the anhydrous form and the hydrous (emulsion) form.

**Anhydrous absorption bases:** Anhydrous lanolin (wool fat), beeswax, hydrophilic petrolatum and organosilicones are examples of anhydrous vehicles that absorb water to form water-in-oil emulsions.

*Anhydrous lanolin* is capable of absorbing about 30–50% of its weight of water to form an emulsion. *Wool fat*, which is emulsifying fraction of anhydrous lanolin along with hydrocarbon bases forms Wool alcohol ointment, BP.

*Organosilicones* such as cetyl dimethicone co-polymer, polyethylene glycol glycerides and caprylic capryl stearates have hydrophilic polyol groups and hydrophobic alkyl side chains, capable of producing w/o emulsion with high water content.

| **Formula #1** | |
|---|---|
| **Hydrophilic petrolatum (USP XX)** | **g** |
| Cholesterol | 30.0 |
| Stearyl alcohol | 30.0 |
| White wax | 80.0 |
| White petrolatum | 860.0 |
| | **1,000.0** |

The maximum amount of water that can be added to 100 g of such a base at a given temperature is known as the water number. To determine the *water*

*number*, the base is stirred continuously as the water is being added. Distilled or deionized water should be used. The end point is reached when no more water can be "absorbed" into the base, as evidenced by droplets of water remaining in the container. In a study involving the separate addition of a series of surfactants to a semisolid base, it was found that the waterabsorbing capacity of the base increased as the HLB number (hydrophilic-lipophilic number) of the surfactant decreased (Table 19.1).

**Table 19.1:** Determination of water numbers using 10 g samples

| Surfactant | HLB | Grams water absorbed | | Water number |
|---|---|---|---|---|
| | | Sample 1 | Sample 2 | |
| (Control: White petrolatum) | | 0.40 | 0.40 | 4.0 |
| Sorbitan monolaurate | 8.6 | 5.21 | 5.41 | 53.1 |
| Sorbitan monopalmitate | 6.7 | 8.20 | 8.52 | 83.6 |
| Sorbitan monostearate | 4.7 | 10.59 | 10.17 | 103.8 |
| Sorbitan monooleate | 4.3 | 24.75 | 25.25 | 250.0 |
| Sorbitan sesquioleate | 3.7 | 29.84 | 31.04 | 304.4 |
| Sorbitan trioleate | 1.8 | 41.95 | 40.31 | 411.3 |

| **Formulas #3 and #4** | | |
|---|---|---|
| | **%** | **%** |
| *Oil phase* | | |
| Lanolin, anhydrous USP | 3.1 | 15.0 |
| Petrolatum, white, USP | 25.0 | — |
| Mineral oil, heavy | 25.0 | 8.0 |
| Beeswax (white wax, USP) | 10.0 | 7.0 |
| Sorbitan sesquioleate | 1.0 | — |
| Propylparaben | 0.05 | 0.05 |
| Amerchol CAB | — | 20.0 |
| *Aqueous phase* | | |
| Sodium borate, USP | 0.7 | — |
| Polyethylene glycol 1500 | — | 5 |
| Methylparaben | 0.15 | 0.15 |
| Purified water | 35.0 | 49.8 |

*Procedure:* Heat the oil phase to 70°C, and add the aqueous solution at 72°C to the oil phase, stirring continuously.

**Hydrous absorption bases:** These are w/o emulsions with ability to absorb additional water. Hydrous lanolin was the prototype or forerunner of the absorption bases because of its ability to absorb water. Various absorption bases were developed as various lanolin isolates and derivatives became commercially available. Many of these lanolin fractions aid in the formation of water-in-oil emulsions. A typical example of a lanolin absorption base follows:

| Formula #2 | |
|---|---|
| **Lanolin absorption base** | **%** |
| Lanolin alcohols | 10 |
| Lanolin | 25 |
| Mineral oil, low viscosity | 30 |
| Purified water | 35 |

Mineral oil is added to reduce the tackiness of the base. Nonionic water-in-oil emulsifiers, such as glyceryl monostearate, cholesterol, cetyl alcohol, and the sorbitan fatty acid derivatives, may be added for improved stability and water-absorbing capacity. These vehicles have “emollient” properties and deposit an oily film upon the skin. Examples of water-in-oil emulsion vehicles that utilize the absorption base principle are given in Formulas #3 and #4.

*Cold cream base*, which reportedly dates back to Galen, was the forerunner of these water-in-oil emulsion vehicles. The cold cream type of emulsion frequently utilizes a borax-beeswax combination as the emulsifier, with mineral oil or a vegetable oil as the continuous phase. A protective oil film remains on the skin following the evaporation of the water. The slow evaporation of water gives the skin a cooling effect. Semisolid water-in-oil emulsions of the borax-beeswax type frequently exhibit poor long-term physical stability. The development and large-scale commercial manufacture of water-in-oil emulsifiers have made it possible to prepare stable semisolids that are oily to the touch. Also, relatively nongreasy water-in-oil emulsions

may be prepared by a judicious combination of raw materials.

Synthetic substances are replacing natural raw materials as the latter become restricted in availability. As an example, the supplies of natural beeswax have declined with the steady price rises that result from both supply and inflation. A number of synthetic beeswaxes have appeared with properties quite similar to the natural. Synthetic spermaceti types have replaced the natural grade since the latter was banned as a result of endangering the whale. Formula #5 illustrates the use of synthetic beeswax in a relatively nongreasy cold cream.

| **Formula #5** | |
|---|---|
| **Cold cream** | **%** |
| **A** | |
| Purified water | 34.60 |
| Borax | 1.00 |
| Methylparaben | 0.25 |
| **B** | |
| Light mineral oil | 50.00 |
| Synthetic beeswax flakes | 13.00 |
| Glyceryl monostearate, pure | 1.00 |
| Propylparaben | 0.15 |

*Procedure:* Dissolve the methylparaben and borax in water at 75 to 80°C. Dissolve the propylparaben in a well-mixed mixture of phase B heated to 75 to 80°C. Add phase A to phase B while stirring rapidly.

Hydrophilic ointment is an example of a water-in-oil absorption base type vehicle that does not have any lanolin or its derivatives in the formula.

| **Formula #6** | |
|---|---|
| **Hydrophilic ointment (USP XX)** | **%** |
| Methylparaben | 0.25 |
| Propylparaben | 0.15 |

| | |
|---|---|
| Sodium lauryl sulfate | 10.00 |
| Propylene glycol | 120.00 |
| Stearyl alcohol | 250.00 |
| White petrolatum | 250.00 |
| Purified water | 370.00 |

This ointment can be used as a vehicle for many drug substances, but is not a cosmetically elegant preparation. The high petrolatum content leaves an unctuous residue upon the skin that may be uncomfortable. Modification of the formulation by reducing the petrolatum content, and the addition of other emollients such as cetyl alcohol, hexadecyl alcohol, and fatty acid esters (isopropyl myristate or palmitate), can add cosmetic appeal to the preparation. The effect of such modifications on the activity of a drug substance incorporated in the base must be determined.

| **Formula #7** | |
|---|---|
| **Hydrophilic ointment base** | **%** |
| *Oil phase* | |
| Amerchol CAB | 50.0 |
| Cetyl alcohol | 2.0 |
| Stearyl alcohol | 2.0 |
| *Aqueous phase* | |
| Sodium lauryl sulfate | 2.0 |
| Water | 34.0 |
| Methyl gluceth-20 | 10.0 |
| Preservative | q.s. |

*Procedure:* Add the water phase at 80°C to the oil phase at 80°C. Cool while mixing to just above congealing temperature. Variations: For greater firmness, increase ratio of stearyl to cetyl alcohol.

| **Formula #8** | |
|---|---|
| **Hydrophilic ointment base** | **% (w/w)** |

| | |
|---|---|
| *Oil phase* | |
| Acetylated lanolin | 5.0 |
| Mineral oil 70 vis. | 5.0 |
| Amerchol L-500 | 10.0 |
| Amerchol CAB | 15.0 |
| Microcrystalline wax, 195°C | 5.0 |
| Cetyl alcohol | 5.0 |
| Brij 52 | 6.0 |
| Brij 58 | 4.0 |
| *Aqueous phase* | |
| Water | 40.0 |
| Methyl gluceth-20 | 5.0 |
| Preservative | q.s. |

*Procedure:* Add the aqueous phase at 80°C to the oil phase at 80°C. Cool while mixing to just above congealing temperature.

## *Water-removable Bases*

The water-removable bases are oil-in-water emulsions and are referred to as "creams." The vanishing cream bases fall into this category. The vanishing creams are so termed because upon application and rubbing into the skin, there is little or no visible evidence of their former presence. Formulas for some typical vanishing cream bases in which different types of emulsifiers are used are given in Table 19.2.

Removal of these creams from skin or clothing is facilitated by the oil-in-water emulsifiers they contain. Creams may be applied to moist skin lesions, since the oil-inwater vehicle tends to absorb any serous discharge. The water removable bases form a semipermeable film on the site of application following the evaporation of water. The semisolid water-in-oil emulsions, however, tend to form a hydrophobic layer on the skin.

Semisolid emulsions are intimate, relatively stable mixtures or dispersions of a hydrophilic phase with a lipophilic phase. The phase that is dispersed in the form of fine microscopic globules is referred to as the 

discontinuous or internal phase; the other is the continuous or external phase. The vanishing cream type vehicles are representative of the oil-in-water emulsions, whereas the absorption bases are generally water-in-oil emulsions.

**Table 19.2:** Formulas for vanishing cream bases

| | #9 Anionic stearate emulsifier % By weight | #10 Anionic emulsifier % By weight | #11 Nonionic emulsifier % By weight | #12 Cationic emulsifier % By weight |
|---|---|---|---|---|
| Stearic acid | 13.0 | 7.0 | 14.0 | |
| Stearyl alcohol | 1.0 | 5.0 | | |
| Cetyl alcohol | 1.0 | 2.0 | 1.0 | |
| Glyceryl monostearate | | | | 10.0 |
| Isopropyl palmitate | | | 1.0 | |
| Lanolin | | | | 2.0 |
| Methylparaben | 0.10 | 0.10 | 0.10 | 0.1 |
| Propylparaben | 0.05 | 0.05 | 0.05 | |
| Sorbitan monostearate | | | 2.0 | |
| Glycerin | 10.0 | 10.0 | | 15.0 |
| Sorbitol solution [70%] | | | 3.0 | |
| Potassium hydroxide | 0.90 | | | |
| Sodium lauryl sulfate | | 1.0 | | |
| Polysorbate 60 | | | 1.5 | |
| Stearyl colamino formyl methyl pyridinium chloride | | | | 1.5 |
| Purified water, q.s. ad | 100 | 100 | 100 | 100 |

## *Water-soluble Bases*

Water-soluble vehicles are prepared from mixtures of high- and low-molecular-weight *polyethylene glycols* or *macrogol*, which have the general formula: $HOCH_2$ $[CH_2OCH_2]_n$ $CH_2OH$. The low-molecular-weight glycols in this category are liquids; those with a moderately higher molecular weight are somewhat unctuous; and the higher molecular weight polyethylene glycols are solids. Suitable combinations of high- and low-molecular-weight polyethylene glycols yield products having an ointment-like consistency, which soften or melt when applied to the skin. No water is required for their preparation. They are water-soluble because of the presence of many polar groups and ether linkages. If the polyethylene glycol ointment has a high percentage of crystalline material, the softening and melting of the ointment rubbed onto the skin will not be as gradual as with petrolatum, since the

crystalline material melts sharply with an increase in temperature. The polyethylene glycol ointments are much less occlusive than in water-in-oil emulsions of the absorption base type; they mix with skin exudates and are readily washed from the skin. The polyethylene glycol vehicles are softened by the addition of water, owing to solution of the glycols. The USP states that 5% of the polyethylene glycol 4000 may be replaced with an equal amount of stearyl alcohol when 6 to 25% of aqueous solution is to be added to the vehicle.

The "water-soluble" bases are also known as greaseless ointment bases. The compatibility of these bases with drug substances and their release rate must be evaluated for each class of drugs. Table 19.3 shows the properties and classification of various ointment bases.

**Table 19.3:** Comparative representation of various ointment bases

| Base type | *Oleaginous* | *Absorption* | *Water washable* | *Water soluble* |
|---|---|---|---|---|
| **Composition** | Hydrocarbons or vegetable oil | Anhydrous or (w/o emulsion) | o/w emulsion | Soluble constituents |
| **Water content** | Anhydrous | Anhydrous/ Hydrous | Hydrous | Hydrous |
| **Water affinity** | Hydrophobic | Hydrophilic | Hydrophilic | Hydrophilic |
| **Washability** | Non-washable | Non-washable | Washable | Washable |
| **Spreadability** | Difficult | Moderate | Easy | Moderate to easy |
| **Drugs incorporated** | Solids, oils | Solids, oils, aq. sol | Solids, oils, aq. sol | Solids, oils |
| **Drug release** | Poor | Poor to fair | Fair to good | Good |
| **Occlusiveness** | Yes | Yes | No | No |
| **Stability** | Unstable, poor | Unstable, better | Unstable, better | Stable |
| **Commercial examples** | White Ointment White petrolatum | Aquabase, polysorb Eucerin (aquaphor) | Unibase, velvachol Dermabase | Polybase |

**Vehicles and skin penetration:** The efficiency of various types of vehicles in aiding penetration can be reasonably predicted by the way in which the vehicle alters the activity of water in the stratum corneum and influences the stratum corneum/vehicle partition coefficient. Greases and oils are the most occlusive vehicles and induce the greatest hydration through sweat accumulation at the skin-vehicle interface. This is accentuated if the skin is covered with occlusive bandages or plastic. Emulsions of the water-in-oil type are somewhat less occlusive than greases. Substances in the vehicle, such as humectants, which have a high affinity for water, may under certain circumstances dehydrate the stratum corneum and decrease penetration. 

Similarly, powders increase the surface area and increase the rate of evaporation of water, and so decrease the extent of hydration. Conversely, vehicles may also affect penetration by their ability to reduce loss of water vapor on the skin surface. Paraffin bases suppress transepidermal water diffusion, whereas a number of other standard vehicles cause a lesser degree of transepidermal water loss suppression.

The role of vehicles on skin penetration is often confusing and contradictory, since the emphasis has generally been placed on the compatibility, stability, and appearance of the product. Only in recent years has attention been given to the influence of components in the vehicle on the movement of the drug through the skin. The release of a substance is favored by the selection of vehicles that have a low affinity for the penetrant or in which the drug is least soluble. This is consistent with the view that the rate of release is governed by the vehicle-to-receptor phase (stratum corneum) partition coefficient. For a given concentration of drug in certain vehicles, the activity coefficient of the drug at that concentration may vary by as much as 1000-fold from one vehicle to the other. The thermodynamic activity of the drug in the vehicle is the product of the concentration of the drug and the activity coefficient of the drug in the vehicle. Solutes held firmly by the vehicle, such as those occurring when the drug forms a soluble complex with the vehicle, exhibit low activity coefficients: hence, the rate of release from such drug-vehicle combinations is slow. Solutes held “loosely” by the vehicle (with less affinity of the vehicle for the drug or solute) exhibit high activity coefficients; therefore, the rate of release from such drug-vehicle combinations is fast. Varied materials require individual formulation based on solubility characteristics, and the formulation may also need modification for different concentrations of the agent to obtain maximal release rates.

## Penetration Enhancers

Materials have been experimentally studied in attempts to increase the rate of absorption of topically applied drugs. These agents are often called “accelerants” or “sorption promoters” or “penetration enhancers”. A sample summary of enhancers includes: water, hydrocarbons, sulphoxides (especially dimethylsulphoxide) and their analogues, pyrrolidones, fatty acids, esters and alcohols, azone and its derivatives, surfactants (anionic, cationic and nonionic), amides (including urea and its derivatives), polyols, essential oils, terpenes and derivatives, oxazolidines, epidermal enzymes, polymers, lipid synthesis inhibitors, biodegradable enhancers and synergistic mixtures. Lipid-protein-partitioning concept suggests that accelerants act by one or more ways selected from three main possibilities (Fig. 19.5).

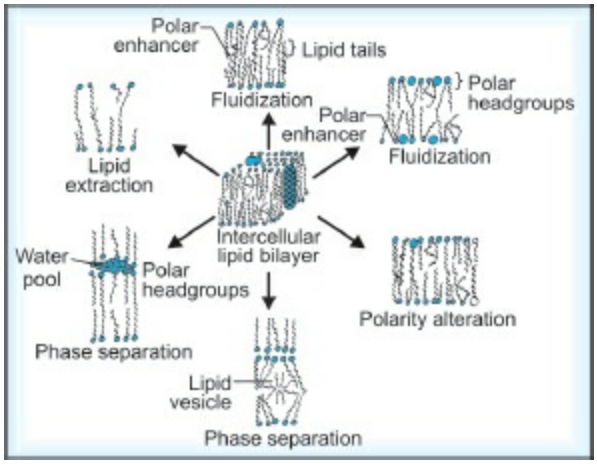



**Fig. 19.5:** Suggested mechanisms for the actions of penetration enhancers

**Lipid action:** Many enhancers such as azone, terpenes, fatty acids, DMSO and alcohols would penetrate into, and mix homogeneously with the lipids, disrupt stratum corneum lipid organisation, making it permeable. The essential action increases the drug’s diffusion coeffcient (Eq. 1). The enhancer molecules jump into the bilayer, rotating, vibrating and translocating, forming microcavities and increasing the free volume available for drug diffusion. Some solvents (e.g. DMSO, ethanol), micellar solutions and drugs (adapalene, tretenoin) may also extract lipids, making the horny layer more permeable through forming aqueous channels. However, terpenes and oleic acid, pool within lipid domains, i.e. they phase-separate and forms permeable ‘pores’ from which, polar molecules, easily permeate to viable

epidermis.

**Protein modification:** Ionic surfactants, decylmethylsulphoxide interact well with keratin in corneocytes, opening up the dense protein structure, making it more permeable, and thus increasing D (Eq. 1). Surface-active agents also appear to increase the permeability of the skin to water by altering the physical state of water in the skin in such a way as to permit greater freedom to the passage of charged hydrophilic substances. When penetration occurs, anionics penetrate best, followed by cationics and nonionic surfactants. Among anionic substances, the laurate ion is reported to have the greatest penetration and the greatest effect on the penetration of other solutes. Soaps of different fatty acids have this property in varying degrees, with penetration more significant for salts of fatty acids having a carbon length of 10 or less. The penetration of fatty acid soaps varies inversely with pH. At higher pH (approximately 11), the action of the anionic surfactant appears to be attenuated or overshadowed by the influence of the more alkaline pH itself.

**Partitioning promotion:** Many solvents enter stratum corneum, change its solution properties by altering the chemical environment, and thus increase partitioning of a second molecule into the horny layer (i.e. raise K in Eq. 1). For example, ethanol increases the penetration of nitroglycerine and estradiol.

Fick's law (Eq. 1) shows that a combination of enhancement effects on diffusivity (D) and partitioning (K) will result in a multiplicative effect. Synergistic effects have been demonstrated for many combinations, such as propylene glycol and Azone, Azone and Transcutol, propylene glycol and oleic acid, propylene glycol and terpenes. It is likely that the cosolvent, such as propylene glycol, acts to increase the concentration of both the permeant and the enhancer in the stratum corneum. In addition, the lipid fluidising effect of the enhancer will increase the free volume within the lipid bilayers thereby facilitating partitioning of both the permeant and solvent. Table 19.4 shows the list of commonly used accelerants along with their mechanisms of action.

**Table 19.4:** Penetration enhancers for topical formulations

| Accelerants | | Mechanism of action | Comment |
|---|---|---|---|
| **Category** | **Examples** | | |
| Sulphoxides | Dimethylsulfoxide<br>Dimethylacetamide<br>Dimethylformamide<br>Decylmethylsulfoxide | Skin damage, Denature proteins, Convert keratin conformation from α-helical to β-sheet, lipid fluidization | Toxic - erythema, uticaria, burning. Enhancer for hydrophilic and lipophilic drugs, non toxic, enhancer |
| Cyclic amide | Azone, N-0915, N-0539, N-0253 | Lipid fluidization N-0131, N-0721 | for hydrophilic and lipophilic drugs |
| Pyrrolidones | N-methyl-2-pyrrolidone 2 pyrrolidone | Alter solvent nature of membrane, generate reservoir within skin | Toxic erythema Hygroscopic, enhancer for hydrophilic drug |
| Fatty acids | Oleic acid<br>Lauric acid<br>Linoleic acid<br>Linolenic acid | Interact and modify SC lipid bilayers, Lipid fluidization | Enhancer for hydrophilic and lipophilic drugs, active in *cis*-configuration |
| Alcohols, glycols | Ethanol<br>Octanol<br>Propylene glycol<br>Transcutol | Increase drug solubility and thermodynamic activity, extract lipids from SC, solvent drug | Commonly used as co-solvents |
| Surfactants | Sodium lauryl sulphate<br>Cetrimide | Increase transepidermal water loss, changes membrane permeability | Irritant, enhancer for lipophilic drug marginal |
| Urea | | SC hydration, keratolytic, increase SC water content | enhancing activity |
| Terpenes | 1-8-cineole<br>L-menthol<br>Nerolidol<br>Limonene<br>Bisabolol | modify solvent nature of SC, modify drug diffusivity, disrupt or reinforce SC bilayer lipids | smaller terpenes are more active enhancers than sequiterpenes |
| Phospholipid | Phosphatidylcholine | occlude the skin surface, increase tissue hydration, interact with SC lipids | Can be employed as vesicles to carry drugs |
| Drugs | Adapalene<br>Tretinoin | Lipid fluidization | Useful for follicular transport |

## Emulsifiers

The water-soluble soaps were among the first emulsifiers used for semisolid oil-in-water emulsions. The viscosity of the cream or ointment prevents coalescence of the emulsified phases and helps to stabilize the emulsion. The addition of fatty polar substances, such as cetyl alcohol and glyceryl monostearate, tends to stabilize the semisolid oil-in-water emulsion. The interfacial film formed around the dispersed phase globules in such a system is generally solid, thereby making the emulsified preparation more rigid.

**Fatty acids and alcohols:** The commercially available fatty acids are really mixtures of related fatty acids. Stearic and palmitic acids are present in the greatest proportion in triple pressed stearic acid along with varying quantities of other fatty acids. *Stearic acid* is used in water-removable creams as an emulsifier to develop a certain consistency in the cream and to give a matt effect on the skin. When a stearate soap is used as an emulsifier, enough potassium hydroxide or triethanolamine usually is added to react with about 8 to 20% of the stearic acid. The unreacted fatty acid increases the consistency of the cream. These creams are soft and develop a sheen or luster upon aging, owing to the formation of stearic acid crystals. Creams formed with sodium stearate are much firmer in consistency. *Stearyl alcohol* and *cetyl alcohol* are used in creams as auxiliary emulsifiers and emollients. In sufficient quantity, stearyl alcohol produces a firm cream that may be softened with cetyl alcohol.

**Polyvalent ions:** Magnesium, calcium, and aluminum, tend to stabilize water-in-oil emulsions by cross-linking with the polar groups of the fatty materials. Nearly all semisolid creams and emulsified ointments require more than one emulsifier. The combination of a surface active agent with an oil-soluble auxiliary emulsifier is referred to as a *mixed emulsifier system.* Triethanolamine stearate soap combined with cetyl alcohol is an example of an oil-in-water mixed emulsifier; beeswax and divalent calcium ions or small quantities of a water-soluble surface active agent exemplify mixed emulsifiers for a water-in-oil emulsion. Maximum stability of an emulsion occurs when a complex interfacial film is formed. Such a film forms when an oil-soluble substance is added and reacts at the interface with the water-soluble surfactant. Soft water-in-oil cream bases can be made with calcium ions as an auxiliary emulsifier. The bases can be made firmer by decreasing

the mineral oil content. Formula #13 is used to make a soft water-in-oil cream base employing divalent calcium ions in the form of water-soluble saccharated lime.

| Formula #13 | |
|---|---|
| | **% (w/w)** |
| **A** | |
| Mineral oil, 65 to 75 viscosity | 30.00 |
| Lantrol | 3.00 |
| Microcrystalline wax | 2.00 |
| Acidlan 20 | 4.00 |
| Propylparaben | 0.20 |
| **B** | |
| Borax | 0.20 |
| Methylparaben | 0.20 |
| Water | 49.75 |
| **C** | |
| Saccharated lime | 0.65 |
| Purified water | 10.00 |

*Procedure:* Heat parts A, B, C separately to 78°C. Add B to A. After the emulsion has formed, add C. Cool and pass through a homogenizer

**Clays:** Magnesium aluminum silicate, has been used as a thickener, suspending agent, and oil-in-water emulsion stabilizer because of the colloidal structure of its aqueous dispersions. It also contributes to the stability of water-in-oil emulsions when used with suitable emulsifiers, probably owing to its thickening action on the internal phase whereby it inhibits coalescence. The magnesium aluminum silicate may migrate to the interfacial area, resulting in a stronger film.

| Formula #14 | |
|---|---|
| | **% (w/w)** |

| | |
|---|---|
| **A** | |
| Magnesium aluminum silicate (MAS) | 2.0 |
| Purified water | 37.0 |
| **B** | |
| Mineral oil, light | 20.0 |
| Petrolatum | 9.0 |
| Isopropyl myristate | 5.0 |
| Lantrol (lanolin oil) | 3.0 |
| 70% sorbitol solution | 20.0 |
| Arlacel 186 (glyceryl oleate and propylene glycol) | 1.0 |
| Polysorbate 80 | 1.0 |
| Preservative | q.s. |

*Procedure:* Add the MAS to the water slowly, agitating continually until smooth. Heat A to 70 to 75°C. Heat B with stirring to 70 to 75°C. Add A to B, and mix until cooled.

**Soap-type emulsion:** The soap-type emulsion may be unstable in the presence of acidic substances. Cationic or nonionic emulsifiers are preferable for drugs requiring an acid pH. Quaternary ammonium compounds like cetyl trimethyl ammonium chloride help to stabilize these emulsions in combination with such fatty alcohols as cetyl alcohol.

The nonionic emulsifiers are employed for both oil-in-water and water-in-oil emulsified pharmaceutical semisolids because they are compatible with many drug substances. The nonionic emulsifiers are versatile and may be used with strongly acidic salts or with strong electrolytes.

| **Formula #15** | |
|---|---|
| **Tripelennamine hydrochloride cream** | **% (w/w)** |
| *Oil phase* | |
| Cetyl alcohol | 5.0 |
| Glyceryl monostearate | 15.0 |
| Sorbitan monooleate | 0.3 |

| | |
|---|---|
| Polysorbate 80, USP | 0.3 |
| *Aqueous phase* | |
| Tripelennamine HCl | 2.0 |
| Methylcellulose 100 cps | 1.0 |
| Purified water, qs. ad | 100.0 |
| Preservative | q.s. |

*Procedure:* Disperse the methylcellulose in hot water in which the preservative has been dissolved, and then chill at 6°C until dissolved. Heat the oil phase to 70°C. Heat the methylcellulose solution to 72°C, and add to the oil phase, stirring continuously. Add the tripelennamine HC1 at 35°C, and stir continuously until dissolved.

**Acyl lactylates:** Recently, a series of emulsifiers have been marketed that contain chemically bonded lactic acid with fatty acids. These acyl lactylates are claimed to be mild and nonirritating to the skin and eyes, to produce an emollient feel to the skin, and to serve as oil-in-water or water-in-oil emulsifiers. The sodium salts are suggested for use in oil-in-water. The particular fatty acid lactylate that is selected should be based on the desired application of the final product as well as on the most compatible fatty acid derivative. Some of the available lactylates and the calculated *HLB* values are as follows:

| **Types of fatty acid** | **Calculated HLB** |
|---|---|
| 1. Stearic | 6.5 |
| 2. Stearic/Palmitic | 8.3 |
| 3. Lauric/Myristic | 14.4 |
| 4. Capric/Lauric | 11.3 |
| 5. Isostearic | 5.9 |

Items 3 and 4 are foamers. Item 4 shows good bacteriostatic properties, owing to the presence of the moderately short chain capric acid.

To achieve adequate stability in creams in which the oil content exceeds 10%, the supplier recommends the use of a coemulsifier to achieve adequate stability. The *HLB* system should be utilized to calculate the ratio between the two emulsifiers for the lipid(s) being used. Several ratios should be

checked to either side of the calculated *HLB* value to optimize the emulsion.

| Formula #16 | |
|---|---|
| **Antiperspirant Cream** | **% (w/w)** |
| **A** | |
| Mineral oil | 23.0 |
| Calcium stearoyl-2-lactylate | 3.2 |
| PEG 400 dioleate | 0.8 |
| **B** | |
| Glycerine | 3.0 |
| Sodium lactate (60%) | 10.0 |
| Purified water | 20.0 |
| **C** | |
| Aluminum chlorohydrate (50%) | 40.0 |

*Procedure:* Heat A, B and C to 70°C in separate vessels. Add B to C immediately before adding to A. Mix with moderate agitation while cooling.

**Promulgens:** The Promulgens are a series of nonionic emulsifiers composed of a mixture of fatty alcohols and their ethoxylates. Two types, D and G, are available and are described in the boxed area below. The two types differ in melting point and in consistency of the emulsions that they form. According to the supplier, the emulsions formed with type D are usually thicker in consistency. Since there are no ester linkages, these emulsifiers are not subject to hydrolysis. In addition, they are compatible with anionic surfactants of the sodium lauryl sulfate type or with cationics such as quaternary ammonium compounds. Type D tends to form creams, and type G tends to form liquid emulsions. It is suggested that they be used in combination to achieve a desired viscosity level.

| | **Promulgen D** | **Promulgen G** |
|---|---|---|
| CTFA adopted name — | Cetearyl alcohol and Ceteareth-20 | Stearyl alcohol and Ceteareth-20 |
| Chemical description | Cetearyl alcohol and | Stearyl alcohol and |

| | | |
|---|---|---|
| — | ethoxylated cetearyl alcohol | ethoxylated cetetryl alcohol |
| Melting point— | 47 to 55°C | 55 to 63°C |

## Humectants

**Polyols:** Glycerine, propylene glycol, sorbitol 70%, and the lower molecular weight polyethylene glycols are used as humectants in creams. The choice of a humectant is based not only on its rate of moisture exchange, but also on its effect on the texture and viscosity of the preparation. These materials prevent the cream from drying out and prevent the formation of a crust when the cream is packaged in a jar. They also improve the consistency and rub-out qualities of the cream when it is applied to the skin, permitting the cream to be spread without rolling. Increasing the humectant content tends to cause tackiness. Sorbitol 70% is more hygroscopic than glycerine and is used at a lower concentration, usually 3% as compared to 10% for glycerine. Propylene glycol and the polyethylene glycols occasionally are used in combination with glycerine, since their ability to absorb moisture is less than that of glycerine.

## Preservatives

The preservatives are added to semisolids to prevent contamination, deterioration, and spoilage by bacteria and fungi, since many of the components in these preparations serve as substrates for these microorganisms. Chemical preservatives for semisolids must be carefully evaluated for their stability with regard to the other components of the formulation as well as to the container. Plastic containers may absorb the preservative and thereby decrease the quantity available for inhibiting or destroying the microorganisms responsible for spoilage. Some preservatives may sting or irritate the mucous tissues of the eye or nasal passages. Methylparabens and propylparabens tend to be more irritating when applied in the nose than quaternary ammonium compounds (e.g. benzalkonium chloride) or the phenylmercuric salts. Boric acid may be used in the ophthalmic preparations, but is omitted from products to be used in the nose because of possible toxic effects if absorbed in large quantities.

Newer preservatives are being marketed, but all of these substances must be thoroughly evaluated for their effectiveness in the product, and their effect on the physicochemical stability of the product. As with all new dermatologicals under development, patch testing must be conducted to eliminate any possibility of skin irritation or sensitivity with the products containing these substances.

Rapid determination of preservative efficacy in semisolids can be done in 48 hours for bacteria and 7 days for molds. The method utilizes the so-called D-value, or decimal reduction time, which is calculated from a plot of the log number of surviving organisms per gram against time of inoculation of the product with specific organisms. The D-value is a numerical value of rate of destruction of a particular organism in a specific product. Since it is a quantitative expression, it can be used to compare the rate of inactivation of different organisms in one or more products. The D-value permits the calculation of the time required for the complete destruction of any size population of organisms.

The method consists of inoculating the product with known amounts of the test organisms. The products are then sampled periodically to record the population of each test organism, and the log of the surviving organisms at each sample time is plotted. The slope of the line is determined by linear

regression, and the negative reciprocal of the slope represents the D-value. The time predicted for complete destruction of the test organism in a product is calculated by linear estimate of the x-intercept.

The time required for the complete destruction of a specific organism of known population in a particular product may be predicted from the D-value. If the mean D-value for *S. aureus* in a product is 2.5 hr, the time for $10^6$ *S. aureus* per milliliter to be totally inactivated is given by the product of the log number of the organisms per milliliter multiplied by the D-value, or 6 × 2.5 h = 15 h.

## Antioxidants

Antioxidants are added to semisolids whenever oxidative deterioration is anticipated. The antioxidant system is determined by the components of the formulation, and the selection depends on several factors, such as toxicity, irritancy, potency, compatibility, odor, discoloration, solubility, and stability. Often, two antioxidants are used, since the combination is often synergistic. Listed in Table 19.5 are some physical and chemical properties of antioxidants in common use. Acids such as citric, maleic, phosphoric, or tartaric may be added to the combination to chelate trace quantities of metals.

**Table 19.5:** Commercial antioxidants

| Common name | BHA | BHT | Propyl gallate |
|---|---|---|---|
| Chemical name | (Butylated hydroxyanisole)<br>3-t-butyl-4-hydroxyanisole<br>2-t-butyl-4-methoxyphenol | (Butylated hydroxytoluene)<br>3,5-di-5-butyl-4-hydroxytoluene<br>2,6-di-t-butyl-4-methylphenol | Alkyl gallate |
| Melting point | 55°–60°C | 70°C | 150°C |
| Solubility at 25° in% [approx.] | | | |
| Propylene glycol | 70 | Insoluble | 55 |
| Peanut oil | 40 | 30 | 0.5 |

Table 19.6 shows the composition of the vehicles of several corticosteroid creams. It is designed to show how currently marketed semisolids utilize the principles described in the previous sections, namely, the different physiologically innocuous fatty materials used in the fat phase, the emulsifier systems, the humectants, preservatives, antioxidants, and chelating agents.

More raw materials are available for use on the skin than for oral use, and in turn, more are available for oral use than for parenteral use. The difference in the number of materials available for each route of administration is due to the type of absorption barrier and physicochemical environment surrounding the absorption sites. Substances such as isopropyl myristate and butyl stearate may be used topically without toxic effects, yet these esters may not be used orally, because hydrolysis of the esters by digestive enzymes yields poorly tolerated alcohols. The absence of comparable hydrolytic enzymes on the skin surface makes these compounds satisfactory for dermatologic medication.

The Federal Food and Drug Administration (FDA) approves chemical

substances and states the maximum concentration that is considered to be safe for use in a particular food or cosmetic. The information is published in the "Federal Register," and a compilation of all such substances is available. For a description of waxes of animal, insect, and vegetable origin such as lanolin, beeswax, carnauba wax, candelila wax, silicones, branched chain compounds, isopropylesters, polyols, cellulose ethers, and other raw materials suitable for creams and ointments, the reader is advised to check sources such as the CTFA, Cosmetic Ingredient Dictionary, as well as suppliers' catalogs. Another listing of cosmetic raw materials appeared in the FD&C Reports (" The Rose Sheet") which was reproduced from the Japan Cosmetic Ingredient Dictionary of 148 government-approved raw materials.

**Table 19.6:** Components of some typical corticosteroid cream bases

| | Cordran cream | Kenatog cream | Lidex cream | Locorten cream | Aristocort cream | Oxylone cream | Synalar cream | Synalar emollient cream | Valisone cream |
|---|---|---|---|---|---|---|---|---|---|
| *Emulsifiers* | | | | | | | | | |
| Sorbitan monostearate | | X | | | | | X | X | |
| Sorbitan monooleate | | | | | | | X | | |
| Sodium lauryl sulfate | | | | X | | | | | |
| Polyoxyethylene sorbitan monostearate | | | | | X | | X | X | |
| Polyoxy 40 stearate | X | | | | | | | | |
| Polyethylene glycol 100 monocetyl ether | | | | | | | | | X |
| Polysorbate 80 | | | | | X | | | | |
| *Fat phase components* | | | | | | | | | |
| Glyceryl monostearate | | X | | | X | X | | | |
| Cetyl alcohol | X | X | | X | | | | X | |
| Spermaceti | | X | | | X | X | | | |
| Stearyl alcohol | | | X | | X | | | X | |
| Stearic acid | X | | | | | | X | | |
| Petrolatum | | | | | | | | | X |
| Liquid petrolatum | X | | | | | | | X | X |
| Cetostearyl alcohol | | | | | | | | | X |
| Isopropyl palmitate | | X | | | | | | | |
| Squalane | | | | | X | | | | |
| *Polyols* | | | | | | | | | |
| Glycerin | X | | | | | | | | |
| Propylene glycol | | X | X | X | | | X | X | |
| Polyethylene glycol 400 | | | | | | X | | | |
| Polyethylene glycol 6000 | | | X | | | | | | |
| 1, 2, 6 Hexanetriol | | | X | | | | | | |
| Sorbitol solution | | | | | X | | | | |
| *Antimicrobial agents* | | | | | | | | | |
| Methylparaben | | X | | X | | X | X | | |
| Propylparaben | | X | | X | | | X | | |
| Ethylparaben | X | | | | | | | | |
| Butylparaben | | | | | | X | | | |
| 4-chloro-m-cresol | | | | | | | | | X |
| Thimerosol | | | | | | | | X | |
| Sorbic acid | | | | | X | | | | |
| Potassium sorbate | | | | | X | | | | |
| *Buffering agents* | | | | | | | | | |
| Citric acid | | | X | | | | X | X | |
| Phosphoric acid | | | | | | | | | X |
| Monobasic sodium Phosphate | | | | | | | X | | |
| Purified water | X | X | | X | X | X | X | X | X |

## PASTE, GELS AND JELLIES

Pastes are dispersions of high concentrations of insoluble powdered substances (20 to 50%) in a fatty or aqueous base. The fatty bases are less greasy as well as stiffer in consistency than ointments because of the large amount of powdered material present. These pastes adhere well to the skin and are of benefit in the treatment of chronic or lichenified lesions. Zinc gelatin paste, USP XX, for example, is used when a protective film on the skin is desired following the evaporation of water. Pastes provide a protective layer, and when covered with suitable dressings, prevent excoriation of the patient's skin by scratching.

Jellies are water-soluble bases prepared from natural gums such as tragacanth, pectin, alginates, and boroglycerin, or from synthetic derivatives of natural substances such as methylcellulose and sodium carboxymethylcellulose.

Gels are usually clear transparent semisolids containing the solubilized active substance. Carbomer 940 swells when dispersed in water in the presence of such alkaline substances as triethanolamine or diisopropanolamine to form a semisolid.

| **Formula #17** | **%** |
|---|---|
| **A** | |
| Carbomer 940 | 0.5 |
| Water | 42.5 |
| Sorbitol 70% solution | 2.0 |
| **B** | |
| Ameroxol OE 20 | 10.0 |
| Solulan 98 | 3.0 |
| Polyvinylpyrrolidone (PVP) K-30 | 1.0 |
| Triethanolamine | 1.0 |
| SD alcohol #40 | 40.0 |

*Procedure:* Phase A-Disperse Carbomer 940 thoroughly in water with good stirring. Add sorbitol solution. Phase B-Add the Ameroxol OE 20 to the alcohol, warm to 35°C, and stir until uniform. Add Solulan 98, PVP, and triethanolamine consecutively, mixing after each addition. Add phase B to phase

A with gentle mechanical mixing until gel forms.

| Formula #18 | % |
|---|---|
| Carbomer 940 | 0.75 |
| Purified water | 34.25 |
| Solulan 98 | 3.00 |
| SD alcohol #40 | 50.00 |
| Diisopropanolamine, 10% in water | 12.00 |

*Procedure:* Prepare a Carbomer slurry in water with gentle agitation, and add mixture of SDA #40 and Solulan mixture, mixing until no particles are visible. Neutralize carefully with diisopropanolamine solution to avoid incorporating air. For greater firmness, increase the concentration of the Carbomer and diso-propanolamine.

Gels are also formed with celluloses such as hydroxypropylcellulose and hydroxy-propyl-methylcellulose. A popular over-the-counter benzoyl peroxide gel contains 6% polyoxyethylene lauryl ether, 40% ethyl alcohol, colloidal magnesium aluminum silicate, hydroxypropylmethylcellulose, citric acid, and purified water.

## INDUSTRIAL PROCESSING

Pilot plant or small-scale production equipment is essential in developing a manufacturing procedure for a production-size batch. The preparation of many batches, ranging in size from 2.5 to 25.0 or more kilograms, for product evaluation and clinical testing provides opportunity to observe, correct, or improve the effects of minor but important variations in the manufacturing technique or formula. Mixing and stirring operations are critical in the preparation of emulsions, and in the laboratory these operations can be carefully controlled in 0.5 or 1.0 kg batches of finished product.

The electrically operated propeller-type mixer (*see* Chapter 1) can be manually adjusted and positioned in the laboratory mixing vessel to achieve maximum turbulence. The angle of entry of the propeller shaft and the depth of the propeller can be easily varied in the laboratory to prevent aeration. A metal spatula can be held or positioned in the beaker during mixing to serve as a baffle to increase turbulence without entrainment of air. Similar maneuverability and control of the mixing action is more limited with larger stationary equipment used for the manufacture of semisolids. High-speed agitation may introduce air into the product, and slow mixing may not form a satisfactory emulsion.

Such problems occur in large-scale manufacture, but would not be apparent in small 1 or 2 kg batches for which a beaker and a laboratory mixer are used. Small-scale equipment similar to the production models can approximate production conditions. It may not be possible to predict the exact mixing time and rotational speed of the agitator, but the overall processing characteristics can be ascertained if identical mixers are used.

Aeration of the semisolid should be avoided, since it may lead to emulsion instability and variation in density within a batch, resulting in weight variation of the ointment or cream in its container. Entrainment of air can occur during the mixing, homogenizing, or milling stage, during the transfer of the product to storage and/or filling equipment, and during the filling or packaging operation.

Aeration may be prevented at the primary emulsion step if one phase is introduced into the other in such a manner that splashing and streaming are avoided. The incoming liquid should enter the mixing kettle below the surface of the other liquid. Vortexing and splashing are overcome by careful

adjustment of the mixing conditions and liquid flow pattern.

Completely enclosed kettles are available for the manufacture of semisolids, which tend to aerate excessively. The sealed vessels can be operated under vacuum; mixing and emulsification can then be performed without entrainment of air. Loss of moisture and volatile substances may present problems, however, because of the vacuum.

A closed system prevents aeration of the product during homogenization or milling, and when the material is transferred to the storage tanks, vessels, or hoppers of the filling machines. When an auger device or a worm-drive is used in the hopper to deliver the material to the tube or jar at the filling outlet, the hopper must be kept full of product, or the rotation of the auger will drive air into the semisolid.

**Rheologic changes:** Homogenization frequently increases the consistency of a semisolid emulsion because it increases the number of emulsified particles. It can also have the opposite effect, that of decreasing the viscosity of the product owing to an electrolyte effect. Some products retain their viscosity if they are not homogenized. Consistency also is affected by the number of passes through the homogenizer, the pressures used for homogenization with the valve-type homogenizer, or the clearance between the rotor and stator if a colloid mill is used.

Some commercial creams are sensitive to agitation and stress. The continuous rotation of an auger in the hopper of the filling machine may cause a cream to liquefy. Such creams may be made more resistant to agitation by a formula change; however, the soft and easy spreading properties of the cream on the skin may then be lost. The replacement of the auger by another, gentler feeding device is of value.

## Fusion Method

Anhydrous ointments are manufactured by the fusion process. The active substance is dissolved in the melted fats and waxes, or in one of the components of the vehicle, and then mixed with the base. The melted mass must be mixed while cooling because the fatty alcohols, fatty acids, and waxes do not form true solutions with petrolatum and mineral oil, but crystallize from the melt as the temperature falls.

## Emulsification Method

### *Preparation of OH and Aqueous Phases*

The components of the oil or fat mixture are placed into a stainless steel steam-jacketed kettle, melted, and mixed. Some of the solid components (e.g. stearic acid, cetyl alcohol) are available in many different forms: cakes, flakes, or powder. The flakes are preferable because of the convenience of handling. The powder may have occasional fine metal contaminants from the pulverizing equipment. Petrolatum is inconvenient to handle unless it is melted and transferred by pumping or pouring from its drum. Transfer of large quantities of petrolatum is expedited by heating the petrolatum in the steel drum in which it is received from the supplier by means of immersion heaters, or by placing the drums in a hot room (60 to 62°C) until the petrolatum is fluid. The liquefied petrolatum can then be transferred to the mixing kettle by metering pump through metal-reinforced inert plastic hoses and insulated pipes. The oil phase is then strained through several layers of cheese cloth to remove any foreign matter. Alternatively, the petrolatum can be passed through a filter medium, particularly for an ophthalmic preparation. The oil phase is transferred by gravity or pump to the emulsion mixing kettle whose walls have been heated to the temperature of the oil phase to prevent some of its higher-melting components from congealing. The components of the aqueous phase are dissolved in the purified water and filtered. A soluble drug may be added to the aqueous phase at this time, provided the high temperature does not degrade the active substance or the emulsion is not adversely affected; otherwise, the soluble drug may be added in solution after the emulsion has formed and has cooled.

**Mixing of phases:** The phases are usually mixed at a temperature of 70 to 72°C, because at this temperature intimate mixing of the liquid phases can occur. The phase mixing temperature can be lowered a few degrees if the melting point of the fat phase is low enough to prevent the premature crystallization or congealing of its components. Decreasing the temperature at which the phases are mixed decreases the cooling time, which is a significant factor when the batch size is large. The properties of some emulsions (borax-beeswax type) depend on the temperature at which the phases are mixed. The initial mixing temperature must be raised above 70 to 72°C, because intimate mixing of the components at monolayer levels cannot occur, since the

emulsion that forms immediately has a high viscosity. The phases can be mixed in one of three ways: (1) simultaneous blending of the phases, (2) addition of the discontinuous phase to the continuous phase and (3) addition of the continuous phase to the discontinuous phase.

The simultaneous blending of the phases requires the use of a proportioning pump and a continuous mixer. This method of emulsification is satisfactory for continuous or large-batch operation. The second method may be used for emulsion systems that have a low volume of dispersed phase. The third process is preferred for many emulsion systems, since the emulsions undergo an inversion of the emulsion type during the addition of the continuous phase, which results in a finer dispersed phase globule. The disperse or aqueous phase in an oil-in-water emulsion is added slowly to the inner phase with agitation. The initial low concentration of water in relation to the concentration of oil results in the formation of a water-in-oil emulsion. The viscosity of the emulsion continues to increase as more water is added, and the volume of the oil phase also increases up to a point of its maximum expansion. Beyond this point, the viscosity decreases, and emulsion inversion is said to occur. The phases reverse themselves, and the inner phase is finely dispersed. Batch sizes are on a weight basis, which is independent of variations in temperature and density. To measure the weight in a kettle, load cells are placed onto the bases of the manufacturing kettle. The kettle exerts a pressure on the cell, which is transmitted by means of a hydraulic force exerted by a layer of oil seated on a diaphragm and can be read on a dial or recorded. Figure 19.6 shows schematic presentation of a load cell and Fig. 19.7 is a photograph of a manufacturing kettle resting on a load cell.

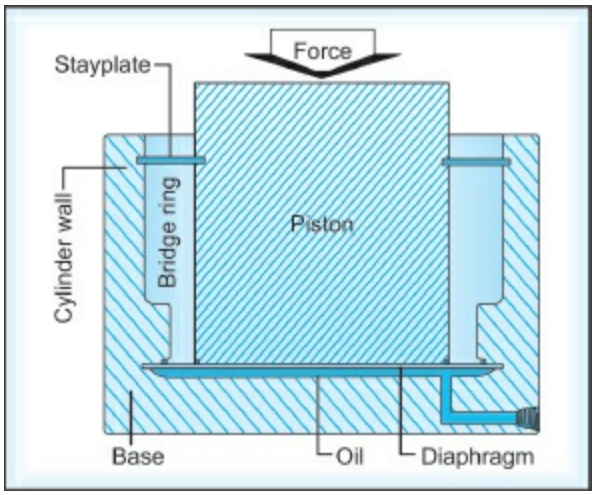


**Fig. 19.6:** Schematic diagram of a cross-section of a hydraulic load cell

**Fig. 19.7:** Photograph of a load cell set under one of the legs of a manufacturing kettle

**Cooling the semisolid emulsion:** Following the addition of the phases, the rate of cooling is generally slow to allow for adequate mixing while the emulsion is still liquid. The temperature of the cooling medium in the kettle jacket should be decreased gradually and at a rate consistent with the mixing of the emulsion and scraping of the kettle walls to prevent formation of congealed masses of the ointment or cream, especially when the semisolid contains a large percentage of high-melting substances. Figure 19.8 is a photograph of a manufacturing kettle showing agitator. Aeration may occur if the semisolid thickens considerably upon cooling, and steps should be taken to prevent this. If perfume is to be added to an oil-in-water emulsion, it is best done while the mixture is at a temperature of 43 to 45°C to avoid chilling the emulsion and to facilitate dissolution of the perfume oil in the still incompletely congealed oil phase. The perfume may be added near room temperature to a water-in-oil emulsion, since dissolution of the perfume oils is to occur in the outer phase of the system.

**Fig. 19.8:** Stainless steel jacketed mixing kettle equipped with a slow-speed, anchor-type sweep blade agitator which takes the material from the side wall and swirls it around this secondary bar type mixer

The drug is added in solution form, if not already incorporated, or as crystals, provided it is soluble in the external phase. An insoluble powder should be dispersed in the continuous phase prior to removing the semisolid from the kettle for homogenization and/or storage.

Figure 19.9 shows a homomixer after it has been withdrawn from a kettle used for dispersing solids. The cooling of the semisolid stored in a large covered vessel is not uniform, since cooling is more rapid at the surface or the walls of the container. Hence, variation in physical properties of the semisolid may occur, such as differences in the size of the fat and wax crystals and the dispersed globules.

**Fig. 19.9:** The high shear homo-mixer is attached to an electrically driven lift mechanism for raising or lowering it into the jacketed kettle used for dispersing solids in liquids

Adjustment of the final water content of a water-in-oil emulsion is not easy once the emulsion has been formed. Several batch runs help to determine the amount of water lost on heating in the particular process, and this lost water should be added to the required amount at the start of manufacture. The oil film surrounding each emulsified water droplet in a water-in-oil emulsion tends to retard evaporation, so that water loss is not excessive following this type of emulsification.

## *Low-energy Emulsification*

The high cost of energy has prompted reevaluation of manufacturing procedures in an attempt to limit the amount of thermal and mechanical energy expended in the production of emulsions. It has been shown that the use of a minimum amount of the emulsion phase in the emulsification stage can result in a considerable reduction in energy requirements and processing time without compromising the quality of the product. The major cost saving is achieved by heating both the oil phase and a portion of water or external phase to the required temperature to form a concentrated emulsion. The balance of the aqueous phase is added at room temperature during the cooling state. Thus, the energy used to heat the aqueous phase and the mechanical energy of mixing during the cooling stage is reduced.

The quality, stability, rheologic properties, and the particle size

distribution of the internal phase of the finished product prepared by this process depend on several variables. These include the temperature required for forming the concentrated emulsion, the ratio of the external phase to the internal phase forming the concentrated emulsion, the phase inversion temperature, the type and intensity of mixing, and the rate of addition of the external phase.

**Homogenization:** The creams or ointments that require further treatment are then transferred or pumped to the proper homogenizer, the selection of which is governed by the degree and rate of shear stress required. The choices include a low-shear gear pump, a roller mill, a colloid mill, a valve-type homogenizer, and a suitable sonic homogenizer (*see* Chapter 1 for detail). Uniform dispersion of an insoluble drug in a semisolid, as well as reduction of the size of the fatty aggregates can be attained by the passage of the warm (30 to 40°C) ointment or cream through a homogenizer or mill.

## EVALUATION OF SEMISOLIDS

**Extrudability and spreadability:** Ease with which semisolid come out from tube or pack and ease with which it spreads over skin is designated as extrudability and spreadability, respectively. In house tests are available for evaluation of extrudability and spreadability behavior of the formulation. Force required, making the ointment move out or extrude through uncapped tube, placed below the plunger is test of extrudability (Fig. 19.10). In spreadability test, a small quantity of ointment is placed between two slides one of which is movable. Force is applied to movable slide and the corresponding distance traveled is recorded (Fig. 19.11). Graphical plots between applied force and distance traveled indicate spreadability.

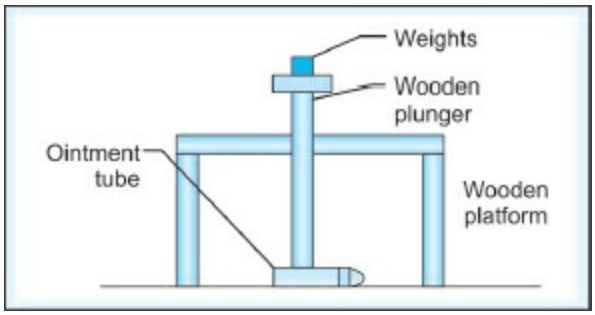



**Fig. 19.10:** Apparatus for extrudability test

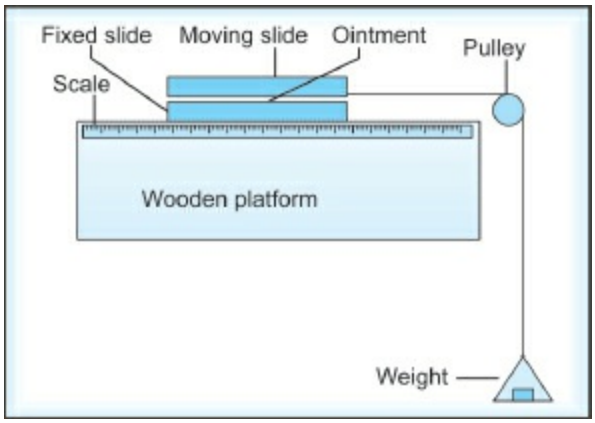



**Fig. 19.11:** Apparatus for spreadability test

**In vitro permeation:** The principal in vitro technique for studying skin penetration involves use of some variety of a diffusion cell in which animal or human skin is fastened to a holder and the passage of compounds from the epidermal surface to a fluid bath is measured. The simplicity of methods and

equipment ranges from just stretching human skin over the mouth of a funnel to using special glass chambers such as Franz diffusion cell. The penetration rates can be quantitated, particularly by radioactive measurements. The area of spread of radioactive agent on the surface is detected with autoradiographs, allowing expression in terms of quantity per unit area per unit time. Radioactive agents have not taken over completely. Many chemical agents penetrate in sufficient concentration to be determined by one or the other techniques of physical or chemical analysis.

More recently, model systems have been used that do not use membranes. Solvents such as alcohol-water have been utilized as models chosen to have negligible solubility in the phase representing the skin, but in which the drug is fairly soluble. A receptor phase or "sink" is used to receive the penetrant. Chloroform and isopropyl myristate have served as sinks. Since they are immiscible with the alcohol-water, it is not necessary to introduce an artificial membrane to separate these sinks from the vehicles. The important factors influencing the release into the receptor phase are the solubility in the vehicle and the partition coefficient of the drug between the vehicle and the receptor phase. Optimal release is obtained from vehicles containing the minimum concentration of solvent required for complete solubilization of the drug.

**Irritation test:** Allergy or irritation due to a specific ointment base component is more frequent and more important, hence a number of test procedures have been devised to test irritancy level of an ointment formulation. Irritation potential of formulation is evaluated in vitro—HET CAM test on eggs, in vivo—Draize dermal irritation test on rabbits and in humans—Draize-Shelanski Repeat Insult Patch Test. The tests should be thorough and well designed; they should include human patch tests, eye irritation studies, determination of minimum lethal dose on at least two animal species, and chronic toxicity studies.

**In vivo test:** The major in vivo methods are histologic techniques, use of tracers, analysis of body tissues and fluids, and elicitation of a biologic response. Tissue changes in skin following the application of various substances to the cutaneous surface can yield information about the specific tissue affected, so that not only absorption, *per se*, is revealed, but also the route of penetration. The method is limited to dyes and to a small number of other substances that yield perceivable colored end products with specific chemical reactions. The pentration of dyes, fluorescence and radioactive

labeling represent the most widespread tracer methods. The studies are always combined with other techniques such as histologic or chemical analysis of tissues or fluids. While radioactive methods give information on the amount of the compound that moves across the skin, they yield little or no information on the route of penetration or on the localization of the penetrant within the structure of the skin.

Urine analysis is by far the most frequently used method. Although the urinary method is extremely valuable, caution is indicated since the recovered agent will not necessarily be the original material or the amount needed to penetrate the skin. Some of the applied agent may have gone elsewhere than into the urine; some may have been metabolized and therefore may be no longer detectable. A steady state between absorption and excretion needs to be reached before measurements can be accepted. The use of topical agents that elicit a physiologic reaction when they reach the dermis makes it possible to demonstrate not only penetration, but also the time required for a reaction to occur. The method is intriguing because it is simple and has practical applications. For example, responses such as sweat secretion, vasoconstriction, vasodilation, pigmentation, and vascular permeability can be recorded with reasonable accuracy by visual observation.

Various methods have been used for studying in vitro and in vivo percutaneous absorption. There are three variables in all the methods: application of the medicament, apparatus, and measurement of the medicament. The combinations of these variables lead to numerous methods, and comparisons between investigations is difficult.

## STORAGE OF SEMISOLIDS

Unless rapid inprocess methods of analysis are developed, it is the usual practice to store the semisolid until the specified quality control tests have been completed before packaging into appropriate containers: tubes, jars, or single-dose packets. A product is considered to be “in-process” until it has been packaged. These preliminary quality control tests are time-consuming and delay the packaging process; however, it is less costly to wait for the assay and to store the material until it can be scheduled for filling than to package and then perhaps be compelled to empty the containers to recover the material, should the semisolid fail to meet the established specifications for the product. Some semisolids have a tendency to “set up” or exhibit an increase of viscosity on storage, and such products cannot be stored for any length of time. The industrial pharmacist must be aware of the delays caused by quality control requirements and packaging schedules, so that he can develop formulations that tolerate storage in bulk without undergoing marked changes in consistency which might cause filling problems. The active substance in the cream or ointment may react with the storage container unless a highly resistant #316, stainless steel, is used for bulk storage. Evaporation of water from a cream must be retarded; this can be effectively accomplished by placing non-reactive plastic sheeting in direct contact with the cream, as well as covering the storage container with a tight-fitting stainless steel lid.

## Transfer of Material for Packaging

The semisolid may be gravity fed, if it is a two-level operation, or pumped to the filling equipment. It must be able to resist the shear stress developed in the transfer of the product, as well as that due to the mechanical action of the filling equipment.

Once a formal manufacturing procedure has been established, there should be no deviation from it. If a change is necessary, however, the problem should be carefully reevaluated, first in the research and development laboratory and then at the pilot plant and manufacturing level.

Although the design and the evaluation of semisolids usually does not include the equipment cleaning operation following the manufacture and filling of the product, it is mandatory that the cleaning operation be thorough to avoid any contamination between batches. Cleaning of large-scale equipment is facilitated and labor costs and downtime of equipment can be reduced through the use of high-pressure (up to 1000 psi), low-volume pump systems now available. The cutting force of high-pressure hot water that may contain detergent can be applied like a knife edge to clean difficult-to-reach tight spots inside kettles and tanks and a variety of manufacturing and processing equipment, eliminating old-fashioned manual scrubbing. Homogenizers, pumps, and filling equipment that have areas wherein pockets of water or product may accumulate and that are ordinarily inacessible must be completely disassembled, cleaned, sanitized, and dried before reassembly. Ball valves and sanitary (Ladish) type or sanitary threaded piping should be used throughout. The packing material used as lubricant for the shafts of mixers should also be replaced during the cleanup process if there is any possibility that they may harbor microorganisms. The manufacturing and packaging equipment should be sanitized following thorough cleaning with detergents. They should be flushed with chlorinated water, formalin, or other suitable sterilant followed by a bacteria-free water rinse. Water and swab samples should be taken to verify microbial elimination.

# 20:
# Suppositories

Suppositories are ovoid or conical medicated solids intended for insertion into one of the several orifices of the body, excluding the mouth. This term derives from the Latin suppositus, meaning “to place under.” These are generally intended for use in the rectum, vagina, and to a lesser extent, the urethra for local or systemic effects.

Suppository use has been known as early as 2600 BC, and was recommended in the works of Hippocrates (ca. 400 BC). Premodern suppositories were made by hand using soap or other semisolid fatty substances as the main vehicles. They were not commonly used until the seventeenth and eighteenth centuries and did not become popular until the midnineteenth century. Oleum Theobromae was first recommended to American pharmacists by A.B. Taylor in 1852, and it soon grew in popularity as the suppository base of choice. Glycerinated gelatin mixtures did not appear as suppository vehicles until about 1875. In 1913, Bruno Solomon published a critical study of suppository bases, in which he classified them into three broad types (1) cocoa butter, (2) fat and wax combinations with cocoa butter and (3) glycerinated gelatin bases.

In the 1930s, several unwanted side effects and disadvantages inherent to oral therapy focused attention, principally in Europe, on the rectal route for administering drugs. Industrial concerns, principally in Germany and France, synthesized special lipid excipients, which were designed to replace cocoa butter. Water-soluble polyethylene glycol type bases were introduced as an improvement on glycerinated gelatin and lipid type suppository bases.

For the combined prescription and over-the-counter market, suppositories represent about 1% of all medications dispensed in the United States. The suppository is a far more popular medication in Europe and South America than in the United States.

## SUPPOSITORY TYPES

**Rectal suppositories** for adults weigh about 2 g and are usually tapered to resemble a torpedo shape. Children's suppositories weigh about 1 g and have a corresponding reduction in size. Rectal suppositories usually employ vehicles that melt or soften at body temperature and are not compressed as tablets, because the amount of liquid in the rectal cavity is inadequate for tablet disintegration. Effervescent base tablets, which contain dried sodium biphosphate, sodium bicarbonate, and to further aid disintegration, starch or finely divided cellulose, have been described as carbon dioxide releasing laxative suppositories. These effervescent base tablets require a small amount of water for rapid disintegration. This compressed rectal suppository is dipped in or sprayed with a thin coating of water-soluble polyethylene glycol to add an external film for protection of the core and for aid in insertion into the rectum.

**Vaginal suppositories** sometimes called pessaries, weigh about 3 to 5.0 g and usually are molded in the globular or oviform shape, or compressed on a tablet press into modified conical shapes. Compressed tablets weighing about 3 g are used as vaginal suppositories. The moisture level of the vagina is sufficient to allow ready dissolution of a tablet if it is formulated to require minimum water for disintegration. The compressed tablet for vaginal use is usually almond-shaped to ease insertion and to provide maximum surface area, which facilitates tablet disintegration and hastens dispersion of the drug on the vaginal wall. A typical vaginal tablet contains active ingredients, with lactose and/or anhydrous dextrose as excipients, and boric and/or phosphoric acid(s) for adjusting the acidity of the vagina to approximately pH 5. Vaginal suppositories are usually used for topical therapy, as in the treatment of vaginitis, or as a spermatocide. They also can be used for introducing drugs with systemic effects.

**Urethral suppositories**, sometimes called **bougies**, are pencil-shaped and pointed at one extremity. Urethral suppositories intended for males weigh about 4 g each and are 100 to 150 mm long; for females, they are 2 g each and usually 60 to 75 mm in length. Figure 20.1 illustrates a representative sampling of several commercially available suppositories.

**Fig. 20.1:** Types and shapes of suppositories

**Layered suppositories:** There is a patent that describes a layered suppository comprising an outer shell that has a melting point of 37 to 38°C and a core that has a melting point of 34 to 35°C and that is contained within and completely surrounded by the shell. Each layer contains different drugs. The layering also may be accomplished by multilayering the suppository in the horizontal plane. This is accomplished by partially filling the mold, allowing the mass to congeal, and pouring additional layers on those previously solidified. Multilayered suppositories serve the dual purpose of separating incompatible drugs in different layers and providing different melting characteristics for controlling the rate of drug absorption.

Coatings have been applied to suppositories to protect them from fast disintegration, to act as lubricants, and to prevent coalescing of adjacent suppositories during storage. Polyethylene glycol, cetyl alcohol, or a patented polyvinyl alcohol and Tween coating is applied for these purposes by dipping the suppository in the coating solution until the desired coating thickness is obtained.

**Soft gelatin suppositories** of varying shapes, filled with either liquid or solid mixtures of the drug, have been made for rectal and vaginal use. Gstirner described a novel procedure for making suppositories. Solutions of either gelatin, alginates, cellulose derivatives, polyvinylpyrrolidone, or silicates mixed with the desired active ingredients are poured into the appropriately shaped molds and lyophilized. The resultant suppositories are non-melting, but readily dissolve in body fluids.

**Hydrogel suppositories:** Rectal administration of a cylindrical hydrogel

for 12 hour periods after it has been soaked in an aqueous drug solution, followed by withdrawal and replacement by a second soaked cylindrical hydrogel, was described by DeLeede, DeBoer, and Breimar. Use of an osmotic delivery system was also detailed.

## THERAPEUTIC USES

Drugs may be administered in suppository form for either local or systemic effects. Such action depends on the nature of the drug, its concentration, and the rate of absorption. Emollients, astringents, antibacterial agents, hormones, steroids, and local anesthetics are dispensed in suppository form for treating local conditions of either vagina, rectum, or urethra. Many published literature illustrates the use of prostaglandin-containing vaginal suppositories for interruption of early pregnancy.

Rectal suppositories are primarily intended for the treatment of constipation and hemorrhoids. Rectal suppositories are also utilized for systemic actions in conditions where oral medication would not be retained or absorbed properly, such as during severe nausea and vomiting or in paralytic ileus. A wide variety of drugs are employed, e.g. analgesics, antispasmodics, sedatives, tranquilizers, and antibacterial agents.

# FACTORS AFFECTING DRUG ABSORPTION FROM RECTAL SUPPOSITORIES

## Physiologic Factors

A number of drugs cannot be administered orally, because either the drugs are affected by the digestive juices, or their therapeutic activity is modified by the liver after absorption. After a drug is absorbed from the small intestine, the drug is carried by the hepatic portal vein to the liver. The liver modifies many drugs chemically and thereby often reduces their systemic effectiveness. On the other hand, a major portion of the same drugs can be absorbed from the anorectal area and still retain therapeutic values. The lower hemorrhoidal veins surrounding the colon and rectum enter into the inferior vena cava and thus bypass the liver. The upper hemorrhoidal vein does connect with the portal veins leading to the liver. More than half (50 to 70%) of rectally administered drugs were reported to be absorbed directly into the general circulation. The lymphatic circulation also helps in absorbing a rectally administered drug and in diverting the absorbed drug from the liver.

The pH of the rectal mucosa plays a significant rate-controlling role in drug absorption. Schanker reported that the rat colon has a pH of approximately 6.8, a pH slightly more acidic than previously believed. Rectal fluids have virtually no buffer capacity, and as a consequence, the dissolving drugs determine the pH existing in the anorectal area. Schanker states that weaker acids and bases are more readily absorbed than the stronger, highly ionized ones. These findings suggest that the barrier separating the colonic lumen from the blood is preferentially permeable to the un-ionized forms of drugs. Thus, the absorption of a drug would be enhanced most likely by a change in the pH of the rectal mucosa that would increase the proportion of un-ionized drug. The effect of intraluminal pH on the absorption of several acidic and basic drugs is shown in Table 20.1.

**Table 20.1:** Effect of intraluminal pH on absorption from the rat colon

| Drug | pKa | pH of the perfusion solution | |
|---|---|---|---|
| | | 6.8-7.2* | 3.6–4.0† |
| | | % Absorbed | % Absorbed |
| *Acid* | | | |
| Salicylic | 3.0 | 12 | 42 ± 3 |
| Benzoic | 4.2 | 19 | 50 ± 7 |
| Phenol | 9.9 | 36 | 37 ± 1 |
| *Base* | | | |
| Aniline | 4.6 | 44 | 32 ± 5 |
| Quinine | 8.4 | 20 | 9 ± 1 |

* The solution, which entered the colon with a pH of 7.2 and left with a pH of 6.8, is a weakly buffered saline solution. +This highly buffered solution entered the colon with a pH of 3.6 and left with a pH of 4.0.

As shown in Table 20.1, absorption of acidic drugs was markedly increased when the pH of the surrounding fluids was lowered. The absorption of salicylic acid rose from 12% at a pH of approximately 7 to 42% at pH 4. In contrast, with a basic drug like quinine, which becomes more ionized at the lower pH values, absorption was decreased from 20% at pH 7 to 9% at pH 4. Phenol is a weak acid and is almost completely un-ionized at both pH 7 and pH 4. Consequently, there was little change in absorption when the pH was lowered.

Riegelman and Crowell have demonstrated that one of the rate-limiting steps in drugs absorption is the diffusion of the drug to the site on the rectal mucosa at which absorption occurs. This diffusivity is influenced not only by the nature of the medicament, such as the presence of surfactant or the water-lipoidal solubility of the drug, but also by the physiologic state of the colon, that is, the amount and chemical nature of the fluids and solids present.

The state of the anorectal membrane also plays a role in drug absorption. This membranous wall is covered with a relatively continuous mucous blanket, which can act as a mechanical barrier for the free passage of drug through the pore space where absorption occurs.

Drugs absorbed from the small and large intestines would most likely be absorbed from the anorectal area. The similarity in the patterns of drug absorption from the small and large intestines makes it unlikely that a drug that has passed through the small intestine would be significantly absorbed from the colon. Conversely, a drug that can be absorbed from the colon most likely would have been completely absorbed in the’ small intestine before

reaching the colon. It should be recognized that although average body temperature is 37°C, patient temperatures may vary from 36 to 38°C, owing to daily and monthly cycles. The suppository formulator must bear in mind the lower limit as a “worst case.”

## Physicochemical Characteristics of the Drug

The sequence of events leading to drug absorption from the anorectal area can be diagramatically represented as follows:

Drug → in vehicle drug in colon fluids → Absorption through the rectal mucosa.

In order for the drug to be available for absorption, it must be released from the suppository and distributed by the surrounding fluids to sites of absorption. By dissolving in the fluids, there is wide contact of the drug with the lumen walls, thereby increasing drug contact with a large number of absorption sites. If the drug has a lipid-water coefficient favoring fat solubility, the drug is released slowly from its suppository excipient. Allawala and Riegelman report that a drug that is very soluble in cocoa butter and present in low concentration does not escape to the surrounding aqueous solution as readily as the drug that is slightly soluble in the cocoa butter. Thus, water-soluble, oil-insoluble salts are preferred in fat-base suppositories. For water-soluble suppository type bases, from which the drug is released as the vehicle dissolves, the water-soluble type salt is the one of choice for quicker drug absorption. For example, to increase the absorption rate from suppositories, ephedrine sulfate and quinine hydrochloride, as well as sodium barbital and sodium salicylate, are preferred to their corresponding bases and acids.

The rate-limiting step in drug absorption from suppositories is the partitioning of the dissolved drug from the melted base and not the rate of solution of the drug in the body fluids. Riegelman and Crowell have shown that the rate at which the drug diffuses to the surface of the suppository, the particle size of the suspended drug, and the presence of surface-active agents are factors that affect drug release from suppositories. Solution of the drugs in solid polyethylene glycol and oleaginous bases resulted in prolonged absorption times, because the drug is slowly eluted into the surrounding fluids. As would be expected, the larger the particle size, the slower the rate of solution, and as a consequence, the drug absorption rate is decreased with an increase in drug particle size. Surfactants can both increase and decrease drug absorption rate. For instance, in the case of sodium iodide, absorption is accelerated in the presence of surfactants and appears to be proportional to the relative surface tension lowering of the vehicle. In addition, Riegelman

and Crowell state that the acceleration of sodium iodide absorption might also be attributed to the mucus-peptizing action of the vehicle. The rectal membrane is covered by a continuous mucous blanket, which may be more readily washed away by colonic fluids that have reduced surface tension. The cleansing action caused by the surfactant-containing vehicle may make additional pore spaces available for drug absorption, thus facilitating drug movement across the rectal membrane barrier. In the case of phenol-type drugs, absorption rate is decreased in the presence of surfactant, probably because of the formation of a drug-surfactant complex.

Schanker showed that the absorption of several acid and base compounds in solution, as in a retention enema, was not affected over a 10-fold range of concentration. In the case of the retention enema, the absolute amount of drug absorbed was directly proportional to the initial saturation concentration present and not to any excess beyond this amount. If the luminal concentration of drug is above a particular amount, which varies with the drug, the rate of absorption does not change with further increases in drug. Thus, colonic absorption of drugs is a matter of simple diffusion across the colonic membrane. In suppositories, however, concentration does play a role in determining the rate of release of the drug from suppository bases.

Once the drug is released from the suppository base and reaches the site of absorption on the lumen wall, the lipid-soluble undissociated drug is the most readily absorbed form. Completely ionized drugs like quaternary ammonium compounds and sulfonic acid derivatives are poorly absorbed. Unionized substances that are lipid-insoluble also are poorly absorbed.

The relation between the degree of ionization and the rate of absorption of drugs is illustrated in Table 20.1. Weak acids with a pKa below 4.3 and weak bases with a pKa below 8.5 are usually readily absorbed. Highly ionized compounds are poorly absorbed. Acids having pKa values below 3.0 and pKa values for bases above 10.0 (quaternary ammonium salts) indicate negligible absorption rates. This relation suggests that the anorectal and colonic mucosae are selectively permeable to the uncharged drug molecule, whereas the ionized drugs penetrate the mucosa poorly or negligibly. Thus, drug absorption can be increased by the use of buffer solutions or salts that convert the pH of the anorectal area to a value that increases the concentration of unionized drug.

In summary, absorption of drugs from the anorectal area is affected by

such physiologic factors as colonic contents, circulation, pH, lack of buffering capacity, physiologic state, and the mucous blanket on the lumen wall. The physicochemical characteristics of drugs affecting absorption are the lipid/water partition coefficient and the degree of ionization. When the amount of drug in the rectal fluids is above the rate-determining level, marked increases in drug concentration play no role in altering established drug absorption rates. Drug concentration is related, however, to release rates from suppository bases. The presence of surfactant may or may not aid absorption, depending on concentration and possible interaction with the drug. Drug particle size is directly related to absorption rate.

## Physicochemical Characteristics of the Base and Adjuvants

Various properties of the suppository base can affect drug absorption. Heinmann et al. reported that with use of sodium phenobarbital, the absorption rate is faster from fatty bases having a lower melting range than from those with higher melting ranges. It was also shown that absorption rate increases along with hydroxyl values. In one of the studies using polyethylene glycols (PEGs) bases, a decrease in absorption time was observed with increase in the molecular mass of the PEG. Since fatty bases may harden for several months after molding, this rise in melting range certainly would affect absorption (*see* "Examples of typical stability problems," presented later in this chapter).

Adjuvants in the formula can affect drug absorption through changes in the rheologic properties of the base at body temperature, or by affecting the dissolution of the drug in the media of the dosage form.

In emulsion type bases, it was shown that the amount of water-soluble drug released increased with the water content of the base, and that the rate of drug released could be prolonged by the addition of an aqueous polymer. Addition of hydrophobic colloidal silicon oxide to fat base suppositories dramatically changed the rheologic behaviour of the mass.

Salicylates were found to improve the rectal absorption of water-soluble antibiotics in lipophilic bases. Drug release from cylindric hydrogels of hydroxyethyl methacrylate decreased as increasing percentages of the cross-linking agent ethylene glycol dimethacrylate were used.

## Blood Levels from Different Dosage Forms

The literature is replete with conflicting information concerning the effectiveness of drugs administered in suppository form. The information is difficult to correlate because of different or inadequate methods for determining blood levels, the nature of the drug and the suppository base, as well as the inability of many patients to retain the suppository. In some studies, blood levels from suppositories were obtained, and in some of these studies, these blood levels were considered therapeutically effective.

Rudolfo *et al*. reported blood levels resulting from the oral, rectal, and intravenous administration of theophylline derivatives (Fig. 20.2). Rectal retention enemas and intravenous injections showed that these two routes are similarly effective if allowance is made for the approximately 30 min delay required for drug absorption from the rectum.

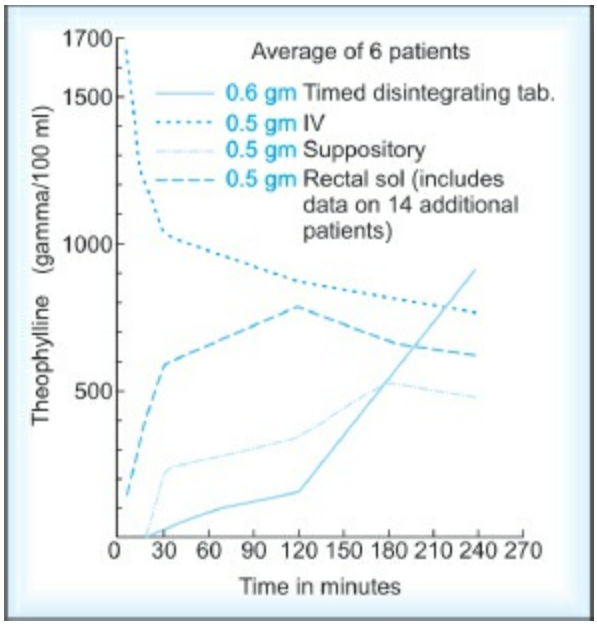



**Fig. 20.2:** Plasma concentration of theophylline following single dose of theophylline monoethanolamine (gamma per 100 ml)

The fact that the rectum or colon is a dependable site for drug absorption seems well established, but not all investigators agree that the suppository dosage form yields therapeutically adequate blood levels. Several investigators report adequate absorption of drug from suppositories. Enesco and coworkers made a comparative study on the absorption time of six drugs,

namely sodium salicylate, chloral hydrate, methylene blue, atropine, morphine, and sodium iodide. The first five drugs are absorbed rectally more quickly and at therapeutically more effective levels than with oral administration. Absorption of sodium iodide is slower by the rectal route than by the oral route, but varies considerably from one individual to another.

Shichiri et al. reported increased intestinal absorption of insulin from a suppository. Copsidas and Ward-McQuaid found pentazocine suppositories equal in relief of moderate pain to pethidine injections and saw less side effects with the suppository. Turrell, Marino, and Nerb report the same dosage requirement for sulfanilamide in glycerinated gelatin suppositories as in tablets. Higher concentrations of the sulfonamide are found in the blood following its administration by enema than by suppository. Thus, in some cases, the suppository does yield effective therapeutic blood levels, although the enema yields faster and higher concentrations of drug in the blood. To maintain the therapeutic effectiveness of a drug in a suppository requires, therefore, a wise choice of both the drug salt and the suppository base.

## SUPPOSITORY BASES

A variety of substances have been used as suppository bases throughout the history of medicine. Their uses and applications were prompted by availability rather than scientific knowledge. More exacting requirements and specifications were applied, however, to this area of pharmaceutical dosage form in the past decades. The use of chemical and physical measurements has provided the yardsticks to set standards for suppository bases, as well as for the finished suppositories. Some of these parameter of the available bases are given in Table 20.2.

**Table 20.2:** Suppository bases

| Name | Supplier | Composition | Melting range °C | Saponification value | Iodine value | Hydroxyl value | Water no. | Solidification point °C | Remarks |
|---|---|---|---|---|---|---|---|---|---|
| Adeps Solidus | | Triglycerides of saturated fattyacids with mono and diglycerides | 33.5–35.5 | 225–240 | <7 | | | 32–34 | Official in German Pharmacopoeia |
| Cebes Pharma 16 | Aarhus Oliefabrik, A/S, Aarhus, Denmark | Modified palm kernel oil | 33–35 | 240–250 | <1 | | | 30–32 | Fat-soluble |
| Cotomar | Procter and Gamble, Cincinnati, OH | Partially hydrogenated cottonseed oil | 37 | 191 | 70 | | | | |
| S-70-X095 | Best Foods/ | Rearranged hydrogenated vegetable oils | 34.4-35.6 | | | | | | |
| S-70-XXA | Refined Oils, Div. of CPC International Englewood Cliffs, NJ | | 38.2-39.3 | | | | | | |
| Hydrokote 25 | Capital City Products Company, | Higher melting fractions of coconut and palm | 33.6–36.3 | 235–245 | <4 | | | 31–33 | Narrow melt. range |
| Hydrokote 711 | Columbus, OH | | 39.5-44.5 | 230–240 | <4 | | | 35–37 | Broad melt. range |
| Hydrokote SP | | kernel oil; upon request, may contain 0.25% lecithin | 31.1-32.3 | 245–255 | <6 | | | 30–32 | Sharp melt. range |
| Idropostal (water-soluble) | Medifarma, Milan, Italy | Condensation product of polyethylene oxide | 53-60 | <1 | | | Sol. | 50 | Water-soluble; in powder form |
| Kaomel | Durkee Foods, Rockville Center, NY | Fractionated Hydrogenated Triglycerides | 35–38 | 194 | 59 | <10 | | | |
| Massa Estarinum A | Edelfett-Werke Werner Schluter, Ham-burg, Germany (Kay-Fries, Inc., Montvale, NJ) | Mixture of tri-, di-, and monoglycerides of saturated fatty acids $C_{11}H_{23}COOH$ to $C_{17}H_{35}COOH$ | 33-35 | 225-240 | <1 | 35-45 | 30-40 | 29–31 | Emulsifies aqueous solution delayed release for drugs with local effect |
| Massa Estarinum AB | | | 29-31 | 235-245 | <3 | 25-40 | | 26.5–28.5 | Low melting mass for melting point correction |
| Massa Estarinum B | | | 33.5–35.5 | 225–240 | <3 | 20–30 | | 31–33 | Melts below body temp., universal base |
| Massa Estarinum C | | | 36-38 | 225-235 | <3 | 20–30 | | 33–35 | For addition to drugs that lower MP; for tropical climates |
| Massa Estarinum D | | | 40-42 | 220-230 | <3 | 30-40 | | 38-40 | High melting base |
| Massa Estarinum E | | | 34-36 | 215-230 | <3 | 45-60 | | 30–32 | Emulsifies water, alcohol, glycerin |
| Massa MI 13 (fat-soluble) | Medifarma, Milan, Italy | Mixture of di-and triglycerides of saturated fatty acids | 36–37 | 225–235 | 2–3 | | | 33–35 | Fat-soluble; specific weight |
| Massuppol 15 | Croklaan, B V, Wormerveer, Holland | | 34–37 | 230-240 | <3 | | 50–100 | 29.5–31.5 | For compression and molding; especially suitable for mass production; fat-soluble |
| Neosuppostal-N (fat-soluble and water-dispersible | Medifarma, Milan, Italy | Hydrogenated triglyceride with fatty alcohols and emulsifiers | 37–38 | 112 | 9–11 | | unlimited | 36–37 | Suitable also for vegetable extracts; compatible with antibiotics and vitamins |
| Paramount B | Durkee Foods, Rockville Center, NY | Hydrogenated interesterified vegetable oils | 34.5-35.5 | 235–245 | <3 | <10 | | | |
| Satina III | | Fractionated hydrogenated triglycerides | 33–35 | 242 | 1 | <10 | | | |
| Suppocire AIML | Establissements | Eutectic mixtures | 33-35 | 225-245 | <2 | 20-30 | | | |
| Suppocire AM | Gattefosse, Paris, | of mono-, di-, and | 35-36.5 | 225-245 | <2 | < 6 | | | |
| Suppocire AP | France | triglycerides | 33-35 | 200-220 | <1 | 30-50 | | | |
| Suppocire BS2X | (Gattefosse Corp., | derived from natural | 36-37.5 | 220-240 | <3 | 15-25 | | | |
| Suppocire $NA_0$ | Elmsford, NY) | vegetable oils | 35.5-37.5 | 225-245 | <2 | <3 | | | |
| Novata Type AB | Henkel International Dusseldorf, Germany (Henkel, Teaneck, NJ) | Mixture of tri-, di-, and monoglycerides of saturated fatty acids | 29-31 | 235-245 | <3 | | | 26.5–28.5 | For light and voluminous additives |
| Type A | | | 33.5-35.5 | 225-240 | <3 | | | 29-31 | For extreme cooling |
| Type B | | | 33.5–35.5 | 225-240 | <3 | | | 31.5–33.5 | Standard for retail pharmacy work |
| Type BC | | | 33.5–35.5 | 225-240 | <3 | | | 30.5–32.5 | Good elasticity; for industrial production |
| Type BD | | | 33.5-35.5 | 225-240 | <3 | | | 32-34 | Low hydroxyl value; relatively nonreactive |
| Type BBC | | | 34-36 | 225-240 | <3 | | | 30.5-32.5 | For automatic processing |
| Type E | | | 34-36 | 215-230 | <3 | | | 29-31 | Emulsion-type; for water pickup |
| Type BCF | | | 35-37 | 225-240 | <3 | | | 30-32 | For heavy crystalline substances |
| Type C | | | 36-38 | 225-235 | <3 | | | 33-35 | For drugs with lower melting point; also for tropical use |
| Novata Type D | | | 40-42 | 220-230 | <3 | | | 38-40 | For increasing melting point |
| Type 229 | | | 33.5-35.5 | 240-255 | <3 | | | 31.5-33.5 | For groups that react with hydroxyl groups |
| Suppostal-N | Medifarma, Milan, Italy | Hydrogenated triglyceride with fatty alcohols and emulsifiers | 37-38 | 98 | 16-18 | | 54 | 36-37 | Compatible with chloral hydrate and other eutectic substances; fat-soluble |
| Suppostal-Es | | | 38-39 | 99 | 18-20 | | 50 | 37-38 | Takes up essential oils, balsams; fat-soluble |
| Wecobee W | PVO International, Inc., Boonton, NJ | Triglycerides | 31.7-32.8 | 242-252 | <4 | | 30-40 | 30-31 | Narrow melting range |
| Wecobee R | | Higher melting fractions of coconut oil and palm kernel oil (may contain 0.25% lecithin) | 33.9-35 | 236-246 | <4 | | 30-40 | 31-32 | Narrow melting range |
| Wecobee S | | | 38-40.5 | 236-246 | <4 | | 30-40 | 32-34 | Broad melting range above body temperature |
| Wecobee M | | | 33.3-36 | 238-248 | <3 | | 30-40 | 29-31 | Narrow melting range; low melting point |
| Wecobee FS | | | 39.4-40.5 | 236-248 | <3 | | 30-40 | 32-34 | Narrow melting range above body temperature |
| Witepsol H12 | Dynamit Nobel, Koln-Mulheim, Germany (Dynamit Nobel of America, Northvale, NJ) | Triglyceride of saturated vegetable fatty acids with monoglycerides (formerly marketed as "Imhausen bases") | 32.3-33.5 | 240-245 | <7 | | 100 | 29-31 | Universal base for industry |
| Witepsol H15 | | | 33.5-35.5 | 230-240 | <7 | | 100 | 32.5-34.5 | Universal base for industry; higher MP |
| Witepsol W35 | | | 33.5-35.5 | 225-235 | <7 | | | 29-32 | Suitable for quick cooling, small-scale production |
| Witepsol S55 | | | 33.5-35.5 | 220-230 | <7 | | | 29-32 | Good dispersibility for vaginal suppository |
| Witepsol E75 | | | 37-39 | 220-230 | <7 | | 45 | 32-34 | Special high MP for eutectic mixtures and tropical suppositories |
| Witepsol E85 | | | 42-44 | 220-230 | <7 | | 45 | | |
| Polyethylene Glycols | Dow Chemical Co., Midland, | Linear polymers of ethylene oxide | | | | | | | |
| 1000 | MI; Union Carbide Chemicals | | 38-41 | | | | Sol. | 38 | Soft, waxy |
| 1450 | | | 42-47 | | | | Sol. | 42 | |
| 4000 | Co., Dallas, TX | | 40-48 | | | | Sol. | | Solid |
| 6000 | | | 49 | | | | Sol. | | |
| Tween 61 | ICI Americas Inc., Wilmington, DE | Polyethylene glycol sorbitan mono-stearate | 35-49 | 95-115 | | 165-195 | | | Used alone or in combination with fats or emulsified bases; waxy consistency |

## Specifications

**1. Origin and chemical composition:** A brief description of the composition of the base reveals the source of origin (i.e. either entirely natural or synthetic, or modified natural products) and chemical makeup (i.e. compound, or a well-defined or poorly elucidated mixture). Physical or chemical incompatibilities of the base with the other constituents may be predicted if the exact formula composition is known, including preservatives, antioxidants, and emulsifiers.

**2. Melting range:** Since fatty suppository bases are complex mixtures of triglycerides and therefore do not have sharp melting points (MP), their melting characteristics are expressed as a range indicating the temperature at which the fat starts to melt and the temperature at which it is completely melted. Although the "melting range" is usually determined by the USP method, manufacturers of bases occasionally use different methods for determining melting characteristics, such as "Wiley melting point," "capillary melting point," "softening point," "incipient melting (or thaw) point," and others.

**3. Solid-fat index (SFI):** From this graph of the percentage of solids versus temperature, one can determine the solidification and melting ranges of fatty bases as well as the molding character, surface feel, and hardness of the bases (Fig. 20.3).

A base with a sharp drop in solids over a short temperature span proves brittle if molded too quickly. This type of base requires a reduced differential between mold temperature and mass temperature for trouble-free molding. Suppository hardness can be determined by the solids content at room temperature. Since skin temperature is about 32°C, one can predict a product that would be dry to touch from a solids content over 30% at that temperature.

Since SFI is determined by dilatometry, which necessitates melting the base to carry out measurements, this set of data is reflective only of the base immediately after molding and not of an equilibrium or hardened state. (*See* "Examples of Typical Stability Problems," in this chapter.)

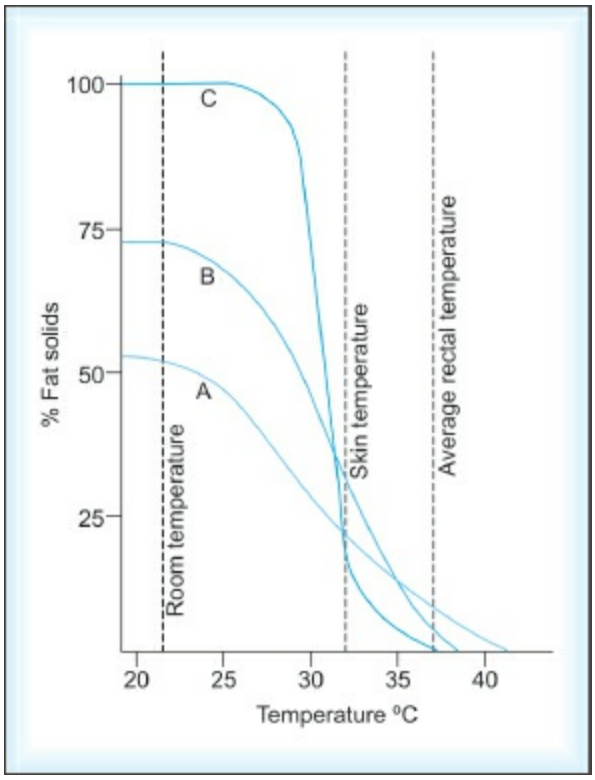



**Fig. 20.3:** Typical solid—fat indices

**4. Hydroxyl value:** This is a measure of unesterified positions on glyceride molecules and reflects the monoglyceride and diglyceride content of a fatty base. The number represents the milligrams of KOH that would neutralize the acetic acid used to acetylate 1 g of fat.

**5. Solidification point:** This value allows prediction of the time required for solidifying the base when it is chilled in the mold. If the interval between the melting range and solidification point is 10°C or more, the time required for solidification may have to be shortened by augmenting refrigeration to produce a more efficient manufacturing procedure.

**6. Saponification value:** The number of milligrams of potassium hydroxide required to neutralize the free acids and saponify the esters contained in 1 g of a fat is an indication of the type (mono-, di-, or tri-) glyceride, as well as the amount of glyceride present.

**7. Iodine value:** This value expresses the number of grams of iodine that reacts with 100 g of fat or other unsaturated material. The possibility of decomposition by moisture, acids, and oxygen (which leads to rancidity in fats) increases with high iodine values.

**8. Water number:** The amount of water, in grams, that can be incorporated in 100 g of fat is expressed by this value. The “water number” can be increased by the addition of surface active agents, monoglycerides, and other emulsifiers.

**9. Acid value:** The number of milligrams of potassium hydroxide required to neutralize the free acid in 1 g of substance is expressed by this value. Low “acid values” or complete absence of acid are important for good suppository bases. Free acids complicate formulation work, because they react with other ingredients and can also cause irritation when in contact with mucous membranes.

## *The Ideal Suppository Base*

The ideal suppository base may be described as follows: (l) Having reached equilibrium crystallinity, the majority of components melt at rectal temperature 36°C but bases with higher melting ranges may be employed for eutectic mixtures, addition of oils, balsams, and suppositories intended for use in tropical climates. (2) The base is completely nontoxic and nonirritating to sensitive and inflamed tissues. (3) It is compatible with a broad variety of drugs. (4) It has no metastable forms. (5) It shrinks sufficiently on cooling to release itself from the mold without the need for mold lubricants. (6) It is nonsensitizing. (7) It has wetting and emulsifying properties. (8) The “water number” is high, i.e. a high percentage of water can be incorporated in it. (9) It is stable on storage, i.e. does not change color, odor, or drug release pattern. (10) It can be manufactured by molding by either hand, machine, compression, or extrusion.

If the base is fatty, it has the following additional requirements (11) “acid value” is below 0.2, (12) “saponification value” ranges from 200 to 245, (13) “iodine value” is less than 7, (14) the interval between “melting point” and “solidification point” is small or the SFI curve is sharp.

A number of suppository bases are available commercially, manufactured for specific purposes. These bases are listed and described in Table 20.2. A suppository base containing all of these properties has not been found. Indeed, some of the properties are mutually exclusive and are not ideal in all situations. Often, the addition of drugs changes the desirable characteristics of the base. Judicious formulation requires the use of the physical values described, for the help in the choice of the base for the drug.

## Cocoa Butter (Theobroma Oil)

Cocoa butter is the most widely used suppository base; it is often used in compounding prescriptions when no base is specified. It satisfies many of the requirements for an ideal base, since it is innocuous, bland, and nonreactive, and melts at body temperature. Cocoa butter has several disadvantages, however. It can become rancid, melt in warm weather, liquefy when incorporated with certain drugs, and with overheating, isomerize to an undesirable lowered melting point.

Cocoa butter is primarily a triglyceride, with the predominant glyceride chains being oleopal-mitostearin and oleodistearin. It is a yellowish-white, solid, brittle fat, which smells and tastes like chocolate. Its melting point lies between 30°C and 35°C (86°F to 95°F), its iodine value is between 34 and 38, and its acid value is no higher than 4. Because cocoa butter can easily melt and become rancid, it must be stored in a cool, dry place and be protected from light.

Cocoa butter exhibits marked polymorphism (the property of existing in different crystalline forms), a phenomenon probably attributable to the high proportion of unsaturated triglycerides. Each of the different forms of cocoa butter has different melting points, as well as different drug release rates. When cocoa butter is heated above its melting temperature (about 36°C) and chilled to its solidification point (below 15°C), immediately after returning to room temperature this cocoa butter has a melting point of about 24°C, approximately 12° below its original state. A knowledge of these polymorphic states is essential for an understanding of how uniform drug release patterns can be obtained from suppository bases consisting primarily of cocoa butter.

Cocoa butter is thought to be capable of existing in four crystalline states:

1. The α form, melting at 24°C, is obtained by suddenly cooling melted cocoa butter to 0°C.
2. The *β*’ form crystallizes out of the liquefied cocoa butter with stirring at 18 to 23°C. Its melting point lies between 28 and 31°C.
3. The *β*’ form changes slowly into the stable *β* form, which melts between 34 and 35°C. This change is accompanied by a volume contraction.
4. The *γ* form, melting at 18°C, is obtained by pouring a cool (20°C) cocoa

butter, before it solidifies, into a container, which is cooled at deepfreeze temperature.

The formation of the various forms of cocoa butter depends on the degree of heating, on the cooling process, and on conditions during this process. At temperatures below 36°C, negligible amounts of the unstable forms are obtained, but prolonged heat above that critical temperature causes the formation of the unstable crystals with resulting lowered melting points. The reconversion to the stable $\beta$ form takes one to four days, depending on the storage temperature—the higher the temperature, the faster the change.

The formation of the unstable forms can be avoided by various methods (1) If the mass is not completely melted, the remaining crystals prevent the formation of the unstable form. (2) Small amounts of stable crystals added to the melted cocoa butter accelerate the change from the unstable to the stable form; this process is called “seeding.” (3) The solidified melt is tempered at temperatures between 28 and 32°C for hours or days, causing a comparatively quick change from the unstable to the stable form.

All of these properties of cocoa butter may cause considerable difficulties in the manufacturing process. As a general rule, the minimal use of heating in the process of melting the fats is recommended. Prolonged heating must be avoided as much as possible. There are several additional disadvantageous characteristics inherent to cocoa butter as a suppository base. Low contractility during solidification causes the suppositories to adhere to molds and necessitates the use of mold release agents or lubricants.

The solidification point of cocoa butter lies about 12 to 13° below its melting point. This property can be utilized in working with cocoa butter in suppository formulations, in which the mass can be kept in a fluid state at comparatively low temperatures. Constant agitation maintains cocoa butter liquid at temperatures below its solidification point.

Cocoa butter does not contain emulsifiers and therefore does not take up large quantities of water (maximum 20 to 30 g of water to 100 g of cocoa butter). The addition of emulsifiers such as Tween 61 (5 to 10%) increases the water absorption considerably. Emulsifiers also help to keep insoluble substances suspended in the fat. Suspension stability is further obtained by the addition of materials (aluminum monostearate, silica) that give melted fats thixotropic properties. There is always the possibility that the suppositories containing these additives will harden on storage. Therefore,

prolonged, careful stability observations are recommended.

Such drugs as volatile oils, creosote, phenol, and chloral hydrate lower the melting point of cocoa butter to a considerable extent. To correct this condition, wax and spermaceti were commonly used. Now special bases with high melting ranges are available for this purpose. Examples of these bases can be found in Table 20.2.

The quality of cocoa butter varies with the origin and treatment. Thus, it is quite possible to obtain different physical characteristics with two cocoa butters from different sources, although both are within all specifications of the USP. The selection of a reliable source of supply is imperative to eliminate broad variations in color and consistency between batches.

## *Cocoa Butter Substitutes*

The mechanization of suppository manufacture, as well as the disadvantages inherent to cocoa butter, have prompted a search for suitable substitutes. The satisfactory ones maintain the many desirable properties of cocoa butter, and attempts are made to eliminate the objectionable ones. Indian pharmacopoeia recommends cuccum butter as a substitute for cocoa butter, due to its lower cost.

## *Typical Treatment of Vegetable Oils to Produce Suppository Bases*

Fat-type suppository bases are produced from a variety of materials, either synthetic or natural in origin. For example, such vegetable oils as coconut or palm kernel oil are modified by esterification, hydrogenation, and fractionation at different melting ranges to obtain the desired product.

An inexpensive method involves hydrogenation of corn oil to reduce unsaturation and so increase the percentage of solid triglycerides at room temperature. The triglycerides with lower melting points are then removed by solvent extraction or by pressing. Manufacturers of fats and oils refer to this type of product as a “hard butter.”

A common method of producing fats intended for use as suppository bases involves *interesterification*. In this process, coconut oil, palm kernel oil, and/or palm oil (all chosen for their high content of lauric acid moieties) are refined to remove free fatty acids, deodorized to remove volatiles,

hydrogenated as described previously, and then interesterified. This final step, catalyzed by sodium methoxide, more equally distributes the fatty acid moieties among the glycerin molecules, creating more common triglycerides, and therefore a more narrow melting range.

A third method utilizes *re-esterification*. First, the oil is split into fatty acids and glycerin by treatment with high-pressure steam. The glycerin is removed from the mixture, and the remaining free fatty acids consist of $C_6$-$C_{18}$ chain length compounds, namely caproic, caprylic, capric, lauric, myristic, palmitic, oleic, and stearic acids. Caproic, caprylic, and capric acids are removed by fractional vacuum distillation, because they are readily rancidified and also may cause irritation of mucous membranes. The remaining fatty acids, consisting mainly of lauric acid, are hydrogenated to harden the mixture and lower its iodine value. The catalyst used in the hydrogenation process is removed, and then the fatty acid mixture is re-esterified with an excess of glycerin to form a mixture of triglycerides, diglycerides, and monoglycerides. The manufacturer controls the re-esterification to build into the base the desired characteristics, e.g. melting range, good mold release, smoothness, and viscosity. In the final steps, the base is deodorized and purified by filtration.

The solid-fat indices of bases produced by these methods are illustrated earlier in Fig. 20.3. Index A represents a typical hard butter. Indices B and C represent interesterified and re-esterified products respectively.

## Hydrophilic Suppository Bases

### *Glycerin Suppositories*

USP XX described the following formula for glycerin suppositories for use as a cathartic:

| | |
|---|---|
| Glycerin | 91 g |
| Sodium stearate | 9 g |
| Purified water | 5 g |
| To make approximately | 100 g |

The glycerin is heated in a suitable container to about 120°C. The sodium stearate is dissolved, with gentle stirring, in the heated glycerin, after which the purified water is added and mixed, and the hot mixture is immediately poured into a suitable mold.

In addition to the above official preparation, USP XX also provided an unofficial formula for glycerated gelatin suppositories:

| | |
|---|---|
| Drug and purified water | 10 g |
| Gelatin | 20 g |
| Glycerin | 70 g |

This formula is most often used in vaginal suppositories, where local application of antimicrobial agents is intended. The suppository dissolves slowly to prolong the activity of the drug. Because glycerin is hygroscopic, these suppositories are packaged in materials that protect them from environmental moisture.

Glycerinated gelatin suppositories do not melt at body temperature, but rather dissolve in the secretions of the body cavity in which they are inserted. Solution time is regulated by the proportion of gelatin/glycerin/water used, the nature of the gelatin used, and the chemical reaction of the drug with gelatin.

Glycerinated gelatin suppositories support mould or bacterial growth, and as a consequence, they are stored in a cool place and often contain agents that

inhibit microbial growth.

## *The Polyethylene Glycols*

The various polyethylene glycol polymers are marketed in the United States as Carbowax and Polyglycols and are suggested for use as suppository bases. Long-chain polymers of ethylene oxide have the general formula $HOCH_2(CH_2OCH_2)_xCH_2OH$ and exist as liquids when their average molecular weight ranges from 200 to 600, and as wax-like solids when their molecular weights are above 1000. Their water solubility, hygroscopicity, and vapor pressure decrease with increasing average molecular weights. The wide range of melting points and solubilities makes possible the formulation of suppositories with various degrees of heat stability and with different dissolution rates. They do not hydrolyze or deteriorate, are physiologically inert, and do not support mold growth.

Several combinations of polyethylene glycols have been prepared for suppository bases having different physical characteristics. Examples of these formulas can be illustrated by a few suggested in the work of Collins, Hohmann, and Zopf.

| **Base 1** | |
|---|---|
| Polyethylene glycol 1000 | 96% |
| Polyethylene glycol 4000 | 4% |

This base is low-melting and may require refrigeration during the summer months. It is useful when rapid disintegration is desired.

| **Base 2** | |
|---|---|
| Polyethylene glycol 1000 | 75% |
| Polyethylene glycol 4000 | 25% |

This base, more heat stable than Base 1, may be stored at higher temperatures than the previous one. It is useful when a slow release of active ingredients is preferred.

Bases 1 and 2, which do not contain water, are usually dipped in water before insertion, so that possible irritation to mucous membranes may be

eliminated. This irritation, or “sting,” is caused when the water is drawn from the mucosa. Most patients do not feel discomfort from the use of these suppositories. Cheymol, Buffet, and Lechat suggested the addition of 10% water to facilitate solution of the suppository after insertion.

The polyethylene glycol suppositories can be prepared by both molding and cold compression methods. A mixture of 6% hexanetriol-1,2,6 with polyethylene glycol 1540, and 12% of the polyethylene oxide polymer 4000 are especially suitable bases for the cold compression technique. The drug is incorporated by dissolving or dispersing in the molten base. Special precautions are necessary in preparing a molded suppository with the polyethylene glycol bases. The mold must be dry because of the solubility of the base in water. The melted mass must be allowed to cool almost to the congealing point before pouring, or the resultant suppository will be fissured owing to the crystallization and contraction of the polymer. Such suppositories may be easily fractured in packaging or handling. The polyethylene glycol base suppositories cannot be prepared suitably by hand rolling. Polyethylene glycol suppositories do not require a mold lubricant and are easier to prepare than cocoa butter suppositories.

Disintegration times of polyethylene glycol type suppository bases, measured in vitro by determining the rate of solution in water at body temperature, do not coincide with the human in vivo results, measured by X-ray study of suppositories containing barium sulfate. Table 20.3 shows comparative results. Thus, clinical results seem the best criterion for choosing the desired polyethylene glycol base, and in vitro test methods should be used for controlling product uniformity of different production lots.

**Table 20.3:** Comparison of in vivo solution time to in vitro disintegration time of three suppository bases

| Base | Solution Time | Disintegration Time |
|---|---|---|
| Polyethylene glycol 1000 | 13 min | 15 min |
| Cocoa butter | 3 min | 3 min |
| A base made of: Polyethylene glycol 1540–94% | | |
| Hexanetriol 1, 2, 6–6% | 18 min | 40 min |

Reports of many workers on adverse reactions of these polymers indicated little difference in sensitivity to individual bases, but a diminished reaction with the 6000 polymer. In one study, the problem of safety, sensitization, and chemical inertness was attributed to impurities and not to the base itself.

## Water-dispersible Bases

Several nonionic surface active materials, closely related chemically to the polyethylene glycols, have been developed as suppository vehicles. Many of these bases can be used for formulating both water-soluble and oil-soluble drugs. The water-dispersible bases offer the additional advantages of storage and handling at elevated temperatures, with claims of broad drug compatibility, nonsupport of microbial growth, nontoxicity, and nonsensitivity.

The surfactants most commonly used in suppository formulations are the polyoxyethylene sorbitan fatty acid esters (Tween), the polyoxyethylene stearates (Myrj), and the sorbitan fatty acid esters (Span and Arlacel). These surface active agents may be used alone, blended, or used in combination with other suppository vehicle materials to yield a wide range of melting points and consistencies.

Caution must be exercised in the use of surfactants with drugs. There are reports indicating increased rate of drug absorption, and other reports showing interaction of these surface active agents with drugs and a consequent decrease in therapeutic activity. Each formulation must be tested in vivo to evaluate its medicinal effectiveness, as well as its safety.

Gross and Becker recommend a water-dispersible, high-melting-point (50°C) suppository base consisting of polyoxyethylene 30 stearate (Myrj 51), water, white wax, and dioctyl sodium sulfosuccinate (Aerosol OT). The use of the Aerosol OT in the formula is claimed to lend synergism to the surfactant and thus aid in the rapid disintegration of the suppository. The drugs studied were phenobarbital, quinine hydrochloride, tannic acid, and chloramphenicol.

Whitworth and Larocca studied, by in vivo and in vitro methods, nineteen different formulas for suppository bases. These suppositories contained hydrogenated cottonseed oil as the main constituent and varying amounts of surfactant to increase the release of a dye, which was intended to represent a drug. Judging from the rate of release of dye, the bases containing 35 to 40% of the emulsifying agents studied (combinations of Tweens, Tweens and Spans, and Tweens and Arlacel) gave best release.

Ward reports on several polyoxyethylene sorbitan derivatives (Tweens),

which are designed to melt at body temperature into liquids that disperse readily in the body fluids. A 2.5 g suppository consisting of Tween 61 (60%) and Tween 60 (40%), and one of Tween 61 (90%) and glyceryl laurate (10%), melted at 37.5°C in about 16 min. The combination of Tween 61 (85%) and glyceryl laurate (15%) melted in about 12 min.

Another type of water-dispersible suppository vehicle reported is based on the use of water-soluble cellulose derivatives, such as methylcellulose and sodium carboxymethylcellulose.

## Hydrogels

Hydrogels are defined as macromolecule network that swell but do not dissolve in water; this swelling is a consequence of the presence of hydrophilic functional groups attached to the polymeric network. Hydrogels employed for rectal or vaginal drug delivery has been prepared from polymers like polyvinyl alcohol, hydroxyl ethyl methacrylate, polyacrylic acid, or polyoxyethylene. Although hydrogels based drug delivery systems have yet to appear in suppositories commercially, research efforts in this area are increasing, giving their potential application in controlled drug delivery, bioadhesion, site specific action and biocompatibility.

## FORMULATION CONSIDERATIONS

The first considerations of the formulator are:

1. Is the medication intended for local or systemic use?
2. Is the site of application rectal, vaginal, or urethral?
3. Is the desired effect to be quick, or slow and prolonged?

Preliminary suppository bases to be studied are first evaluated by measuring drug availability from the suppository in water at 36 to 37°C. Stability of both active ingredients and base containing drug(s) at 4°C and room temperature is the next consideration. To reduce the number of suppository bases chosen for stability studies, ease of molding and release in the manufacturing equipment are simultaneously studied. After these parameters are established, toxicity (irritancy) and drug availability are measured in animals before the medication is ready for human clinical trials.

## Suppositories for Systemic Effect

A selection of possibly desirable suppository bases should be made, e.g. by choosing from those suggested in Table 20.2. Availability and cost of the suppository bases must be considered before the formulation work is begun. Whichever base is used, the drug should be homogeneously dispersible in it, but releasable at the desired rate to the aqueous body fluids surrounding the suppository. Therefore, the solubility of the active ingredient(s) in water or other solvents should be known. If the drug favors water, a fatty base with low water number may be preferred. On the other hand, if the drug is highly fat-soluble, a water-type base, perhaps with the addition of a surfactant to enhance solubility, may be the preferred choice.

The theoretically desirable suppository formulations are molded in the laboratory and stored at room temperature (25 ± 3°C) for at least 48 hours before undergoing in vitro testing for release rate, to be described in the section “Testing of Suppositories”. To enhance the homogeneity of drug in the desired base, either a suitable solvent is used or the drug is finely comminuted before incorporation. A drug that is soluble in a minimal quantity of water, or in another liquid miscible with the base, can be dissolved and the solution added to the molten base. If the drug is to be incorporated directly into the base, it should be finely ground so that 100% can be passed through a 100-mesh USP screen. Fragility, brittleness, and ease of handling the suppository formulations on production equipment are some of the screening tests performed before the time-consuming and costly animal and human tests begin.

The fluid content in the rectum is small. Therefore, in vitro findings of release rates, in which comparatively large amounts of water are generally used, can be regarded only as a general guideline for formulation and after the formula is in production, as a quality control procedure. In many cases, there is reasonable correlation of in vitro to in vivo release rates (*see* Table 20.3), but this is not necessarily so. In vivo clinical findings in man are the ultimate criteria for choosing a desired formulation, and the suppository formulation so chosen yields the in vitro release rate pattern that is to be used as the desired standard. The clinical findings may be based on blood levels of the drug and/or desired clinical effects in man. Thus, since the suppository formula is not chosen until the desired clinical effects in man are determined,

screening of several prototype formulas by laboratory tests is the practiced procedure. Chemical and physical stability, consistency of in vitro drug release patterns within theore tically desired ranges, and animal toxicity are some characteristics studied before suppository formulas are chosen for human clinical trials.

Once several likely candidates are chosen for intense human clinical studies, at least two formulations, each containing different batches of acceptable-quality ingredients, are placed on prolonged stability tests. The parameters tested are described in the section “Specific Problems in Formulating Suppositories.” The suppositories are stored at room temperature (25 ± 3°C) and at 4°C. They are tested at regular intervals (1-, 3-, and 6 month and 1- and 2-year periods) for changes in appearance, melting and softening range, drug stability, base stability, and in vitro drug release pattern. The minimum age of the samples to be used in clinical trials should be determined by the stability of the melting range of the formula, since nearly all shift upward initially, but the time required to reach equilibrium varies.

## Suppositories for Local Effect

Drugs intended for local action are generally nonabsorbable, e.g. drugs for hemorrhoids, local anesthetics, and antiseptics. The bases used for these drugs are virtually nonabsorbable, slow in melting, and slow in drug release, as contrasted with suppository bases intended for systemic drugs. Local effects are generally delivered within a half hour and last at least 4 hours.

The base chosen is one intended for local action; several such bases are depicted in Table 20.2. The drug must be homogeneously distributed in the suppository base. This is accomplished as described previously for incorporating the drug in a systemic base. The suppository is tested 48 hours after molding by immersion in a 37°C water bath. (*See* the section "Testing of Suppositories".) The desired base should release an adequate amount of drug within a half hour, and completely melt with release of all drug between 4 and 6 hours. A suppository that does not melt within the 6-hour test period would probably not completely release its drug, cause discomfort to the patient, and be expelled by the patient before it is fully utilized.

## Water in Suppositories

Use of water as a solvent for incorporating substances in suppository bases should be avoided for the following reasons:

1. Water accelerates the oxidation of fats.
2. If the water evaporates, the dissolved substances crystallize out.
3. Unless the water is present at a level significantly higher than that required for dissolving the drug, the water has little value in facilitating drug absorption. Absorption from water-containing suppositories is enhanced only if an oil-in-water emulsion exists with more than 50% of the water in the external phase.
4. Reactions between ingredients contained in suppositories are more likely to occur in the presence of water. Sometimes, anhydrous chemicals are used to avoid this possibility.
5. The incorporation of water or other substances that might be contaminated with bacterial or fungal growth necessitates the addition of bacteriostatic agents such as the parabens.

## Hygroscopicity

Glycerinated gelatin suppositories lose moisture by evaporation in dry climates and absorb moisture under conditions of high humidity. Polyethylene glycol bases are also hygroscopic. The rate of moisture change in polyethylene glycol bases depends not only on humidity and temperature, but also on the chain length of the molecule. As the molecular weight of these ethylene oxide polymers increases, the hygroscopicity decreases, with a significant drop for the 4000 and the 6000 series.

## Incompatibilities

Polyethylene glycol bases were found to be incompatible with silver salts, tannic acid, aminopyrine, quinine, ichthammol, aspirin, benzocaine, iodochlorhydroxyquin, and sulfonamides. Many chemicals have a tendency to crystallize out of polyethylene glycol, e.g. sodium barbital, salicylic acid, and camphor. Higher concentrations of salicylic acid soften polyethylene glycol to an ointment-like consistency, and aspirin complexes with it. Penicillin G, although stable in cocoa butter and other fatty bases, was found to decompose in polyethylene glycol bases. Fatty bases with significant hydroxyl values may react with acidic ingredients.

## Viscosity

The viscosity of the melted suppository mass is important in the manufacture of the suppository and to its behaviour in the rectum after melting. Melted cocoa butter and some of its substitutes have low viscosities, whereas the glycerinated gelatin and polyethylene glycol type base have viscosities considerably higher than cocoa butter. In the manufacture of suppositories made with low-viscosity bases, extra care must be exercised to avoid the sedimentation of suspended particles. Poor technique can lead to nonuniform suppositories, particularly in the distribution of active ingredients. To prevent segregation of particles suspended in molten bases, the well-mixed mass should be handled at the lowest temperature necessary to maintain fluidity, constantly stirred without entrapping air, and quickly solidified in the mold.

The following approaches may be taken to overcome the problems caused by use of low viscosity bases:

1. Use a base with a more narrow melting range that is closer to body temperature.
2. The inclusion of approximately 2% aluminum monostearate not only increases the viscosity of the fat base considerably, but also aids in maintaining a homogeneous suspension of insoluble materials. Cetyl, stearyl, or myristyl alcohols or stearic acid are added to improve the consistency of suppositories.

## Brittleness

Suppositories made from cocoa butter are quite elastic and do not fracture readily. Synthetic fat bases with a high degree of hydrogenation and high stearate contents, and therefore a higher solid content at room temperature, are usually more brittle. Fracturing of the suppository made with such bases is often induced by rapid chilling (shock cooling) of the melted bases in an extremely cold mold. Brittle suppositories are troublesome not only in manufacturing, but also in the subsequent handling, wrapping, and use. To overcome this difficulty, the temperature differential between melted base and mold should be as small as possible. Addition of a small amount of Tween 80, Tween 85, fatty acid monoglycerides, castor oil, glycerin, or propylene glycol imparts plasticity to a fat and renders it less brittle.

## Density

To calculate the amount of drug per suppository, the density of the base must be known. The volume of the mold cavity is fixed, and therefore, the weight of the individual suppository depends on the density of the mass. Knowledge of the suppository weight can be obtained from a given mold and density of the chosen base; the active ingredients can then be added to the bulk base in such an amount that the exact quantity of drug is present in each molded suppository. If volume contraction occurs in the mold during cooling, additional compensation must be made to obtain the proper suppository weight. Thus, density alone cannot be the sole criterion for calculating suppository weight per fixed volume mold. When volume contraction occurs, the suppository weight is determined empirically by small batch runs.

## Volume Contraction

This phenomenon occurs in many melted suppository bases after cooling in the mold. The results are manifested in the following two ways.

1. *Good mold release:* This is caused by the mass pulling away from the sides of the mold. This contraction facilitates the removal of the suppositories from the mold, eliminating the need for mold release agents.
2. *Contraction hole formation at the open end of the mold:* This undesirable feature results in lowered suppository weight and imperfect appearance of the suppository. The contraction can be eliminated by pouring a mass slightly above its congealing temperature into a mold warmed to about the same temperature. In volume production using standard molds, where adequate control of temperature may not be feasible, the mold is overfilled so that the excess mass containing the contraction hole can be scraped off.

## Lubricants or Mold Release Agents

Cocoa butter adheres to suppository molds because of its low volume contraction. These suppositories are difficult to remove from the molds, and various mold lubricants or release agents must be used to overcome this difficulty. Mineral oil, an aqueous solution of sodium lauryl sulfate, various silicones, alcohol, and tincture of green soap are examples of agents employed for this purpose. They are applied by wiping, brushing, or spraying. The release of suppositories from damaged molds was improved by coating the cavities with polytetrafluoroethylene (Teflon).

## Dosage Replacement Factor

The amount of base that is replaced by active ingredients in the suppository formulation can be calculated. The replacement factor, *f*, is derived from the following equation:

$$f = \frac{100\,(E - G)}{(G)\,(X)} + 1 \qquad \ldots (1)$$

where, *E* is weight of pure base suppositories, *G* is weight of suppositories with *X*% active ingredient.

Most commonly used drugs are tabulated by replacement factor, using cocoa butter arbitrarily assigned the value 1 as the standard base:

| | |
|---|---|
| Boric acid | 0.67 |
| Phenobarbital | 0.81 |
| Mild silver protein | 0.61 |
| Balsam Peru | 0.83 |
| Bismuth subgallate | 0.37 |
| Bismuth subnitrate | 0.33 |
| Camphor | 1.49 |
| White or yellow wax | 1.0 |
| Spermaceti | 1.0 |
| Chloral hydrate | 0.67 |
| Quinine hydrochloride | 0.83 |
| Digitalis leaves, powdered | 0.61 |
| Ichthammol | 0.91 |
| Castor oil | 1.0 |
| Phenol | 0.9 |
| Procaine hydrochloride | 0.8 |
| Resorcin | 0.71 |
| Salol | 0.71 |
| Sulfanilamide | 0.6 |

| | |
|---|---|
| Sulfathiazole | 0.62 |
| Theophylline sodium acetate | 0.6 |
| Zinc oxide | 0.15–0.25 |

## Weight and Volume Control

The amount of active ingredient in each suppository depends on (1) its concentration in the mass, (2) the volume of the mold cavity, (3) the specific gravity of the base, (4) the volume variation between molds-good machining of the molds should keep the volume of each cavity within 2% of a desired value, (5) weight variations between suppositories due to the inconsistencies in the manufacturing process, e.g. incomplete closing of molds, uneven scrapings. Regardless of the reason for the variation in weight, it should be within ± 5%.

The German and Russian Pharmacopoeias state individual weight variations of rectal suppositories at ± 5% of the average weight. The Pharmacopeia Nordica allows ± 10% of the average weight for 90% of the suppositories, but these deviations must not exceed ± 20%.

## Rancidity and Antioxidants

Many investigators confuse the acidity of fats with rancidity. The presence of free fatty acids in either small or large quantities is no indication of rancidity, or that such a product may necessarily become rancid.

Rancidity results from the autoxidation and subsequent decomposition of unsaturated fats into low- and medium-molecular-weight ($C_3$-$C_{11}$) saturated and unsaturated aldehydes, ketones, and acids, which have strong, unpleasant odors. The lower the content of unsaturated fatty acid constituents in a suppository base, the greater is its resistance to developing rancidity. Since this reaction begins with the formation of hydroperoxides, a measure of autoxidation in progress is the peroxide value. This peroxide or active oxygen value is a measure of the iodine liberated from an acidified solution of potassium iodide by the so-called “peroxide oxygen” of the fats.

Examples of effective antioxidants are phenols, such as m- or p-diphenols; *α*-naphthol; quinones, such as hydroquinone or β-naphthoquinone; tocopherols, particularly the *β* and *α* forms; gossypol present in cottonseed oil; sesamol present in sesame oil; propyl gallate and gallic acid; tannins and tannic acid; ascorbic acid and its esters; butylhydroxyanisole (BHA); and butylhydroxytoluene (BHT).

## MANUFACTURE OF SUPPOSITORIES

Four methods are used in preparing suppositories, namely molding by hand, compression, pour molding, and compression on a regular tablet press.

## Hand Molding

The simplest and oldest method of preparing a suppository is by hand, i.e. by rolling the well-blended suppository base containing the active ingredients into the desired shape. The base is first grated and then kneaded with the active ingredients by use of a mortar and pestle, until the resultant mass is plastic and thoroughly blended. The active ingredients are usually finely powdered, or dissolved in water, or sometimes mixed with a small amount of wool fat to help incorporation with the suppository base. The mass is then rolled into a cylindric rod of desired length and diameter, or into vaginal balls of the intended weight. Starch or talcum powder on the rolling surface and hands prevent the mass from adhering. The rod is cut into portions, and then one end is pointed. This method is practical and economical for the manufacture of small numbers of suppositories.

## Compression Molding

A more uniform and pharmaceutically elegant suppository can be made by compressing the cold-grated mass into a desired shape. A hand-turned wheel pushes a piston against the suppository mass contained in a cylinder, so that the mass is extruded into molds (usually three). Hand-operated instruments and procedures for making suppositories in this manner are described in basic pharmacy texts.

The cold compression method is simple and results in a more elegant appearance than does hand molding. It avoids the possibilities of sedimentation of the insoluble solids in the suppository base, but is too slow for large-scale production. One of the major disadvantages in the use of the cold compression technique for molding fat type base suppositories is air entrapment. This unavoidable inclusion of air makes close weight control impossible and also favors the possible oxidation of both the base and active ingredients.

## Pour Molding

The most commonly used method for producing suppositories on both a small and a large scale is the molding process. First, the base material is melted, preferably on a water or steam bath to avoid local overheating, and then the active ingredients are either emulsified or suspended in it. Finally, the mass is poured into cooled metal molds, which are usually chrome-or nickel-plated.

## Automatic Molding Machine

The molding operations (pouring, cooling, and removal) can be performed by machine. All filling, ejecting, and mold-cleaning operations are fully automated. The output of a typical rotary machine ranges from 3500 to 6000 suppositories an hour.

In the rotary molding machine, as illustrated in Fig. 20.4, chrome-plated brass molds are installed radially in the cooling turntable. First, the prepared mass is fed into a filling hopper where it is continuously mixed and maintained at constant temperature. The suppository mold is lubricated by brushing or spraying and then filled to a slight excess. After the mass solidifies, the excess material is scraped off and collected for re-use. All pumping and scraping units are heated electrically at controlled temperatures. The cooling cycle is adjusted, as required by the individual suppository mass, by adjusting the speed of the rotary cooling turntable. The solidified suppositories are moved to the ejecting station, where the mold is opened and the suppositories are pushed out by steel rods. The mold is closed, and then moved onto the spraying station for lubrication and a repeat of the cycle.

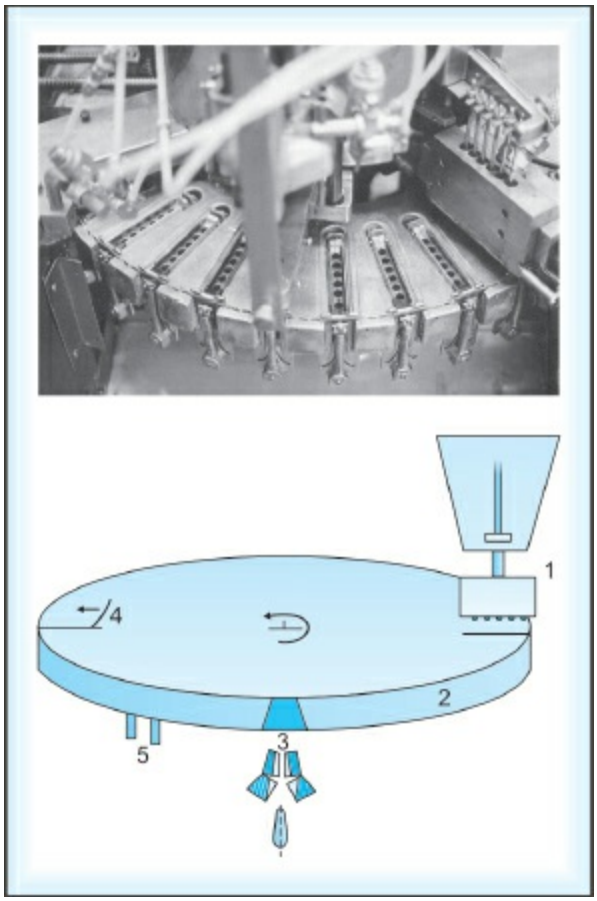



**Fig. 20.4:** Rotary suppository molding machine, (1) feeding device and filling hopper, (2) rotating cooling turntable, (3) suppository ejection station, (4) scraping device, (5) refrigerant inlet and outlet

Temperatures and output speeds are regulated to create optimal conditions for continuous, automatic production. Molds must be kept clean to prevent any deposition of mass from interfering with their proper closure. Incomplete closure of the molds results in overweight suppositories with mold marks. Air jets blow loose particles out of the molds and thus help to minimize machine downtime for cleaning.

On the straight-line machine, molds are arranged for increased productivity, that is, up to about 10,000 suppositories per hour. The straight-line machine carries out the same steps as the rotary model. The individual molds are carried on a track through a cooling tunnel, where scrape-off and ejection occur.

## PACKAGING OF MOLDED SUPPOSITORIES

Suppositories must be packaged so that each suppository is overwrapped, or they must be placed in a container in such a manner that they do not touch each other. Staining, breakage, or deformation by melting caused by jostling or adhesion can result from poorly wrapped and packaged suppositories.

Suppositories in direct contact with one another are marred by fusion resulting from changes in ambient temperature. Partially melted suppositories stain the outer package unless they are over wrapped or are packaged with some other barrier that prevents contact with the outer container. Suppositories usually are foiled in tin or aluminum; paper and plastic strips are also used.

Overwrapping of suppositories is done by hand or machine. Hand packaging is slow and yields a nonuniform and generally inelegant preparation. Modern packaging machines overcome these difficulties. They are capable of wrapping uniformly about 8,000 suppositories per hour. In one type of machine, the chill-hardened suppositories are placed in a notched turntable and then fed to the packaging station, where the foil is unwound from a roll, cut to size, and finally rolled around each suppository. In other machines, the suppositories are enclosed in cellophane or heat-sealing aluminum foils. Plastic may be thermoformed into two package halves, with the suppository placed mechanically in one half and the second half sealed on afterward. The heat sealer makes contact with the plastic strips only momentarily and at a sufficient distance, so that the suppository is not affected by the heat.

Many suppositories are not individually wrapped. In such cases, they are placed into cardboard boxes or plastic containers that have been molded to provide compartments for 6 to 12 suppositories.

The individually wrapped suppositories are usually packaged in slide, folding, or setup boxes. Occasionally, hygroscopicity or volatility of ingredients necessitates packaging the suppositories in glass or plastic containers. In the case of glycerinated gelatin suppositories, a well-sealed package is required. Changes in weight of suppositories depend on the types of packaging materials used.

## In-Package Molding

A significant advance in suppository manufacturing was the development of automated methods for molding suppositories directly in their wrapping material. This is currently accomplished with either plastic or aluminum foil. Machines utilizing plastic either thermoform the mold and fill the mold in sequence, or simply fill the mass into previously thermoformed molds. Machines using aluminum foil/polypropylene/lacquer laminates emboss two parallel strips of foil so that when they are sealed together, molds are formed.

In both plastic and aluminum approaches, the tops of the molds are left open for the entrance of filling nozzles. After the mass has been injected, usually by means of small, variable-throw piston pumps, the tops are sealed. The strips are then passed in an upright position through a cooling station. Using these techniques, one machine can make 12,000 to 20,000 suppositories per hour. The advantages of in-package molding include high production rates, no generation of scrapings, no bulk handling or storage of unwrapped suppositories, and maintenance of strict temperature controls. The disadvantages are dependence on the shape of the formed mold and seal completeness for the shape of the suppository and depression formation in the rear of the suppository since no scraping takes place.

Disposable molds have the additional advantage of being suited for suppositories intended for tropical climates. If the mass should melt at the high storage temperatures, the mold still retains it in its proper shape, so that upon cooling it can be dispensed without distortion.

## STORAGE

Suppositories should be protected from heat, preferably by storing in the refrigerator. Polyethylene glycol suppositories and suppositories enclosed in a solid shell are less prone to distortion at temperatures slightly above body temperature. Glycerinated gelatin suppositories should be protected from heat, moisture, and dry air by packaging in well-sealed containers and storing in a cool place.

## TYPICAL STABILITY PROBLEMS

The suppository, including active ingredients and the base, must be chemically and physically stable at refrigerator temperatures as well as at room temperature storage conditions for at least two years. Storage stability studies are normally conducted at 4°C and at room temperature (25 ± 3°C).

Cocoa butter suppositories in storage sometimes "*bloom*", i.e. form a white powdery deposit on the surface. This is unsightly and usually can be avoided if the suppositories are wrapped in foil, and stored at uniform cool or refrigerator temperatures.

Fat base suppositories have been shown to *harden* for a period of time after manufacture. This upward shift in melting range is due to slow crystallization to the more stable polymorphic forms of the base. Depending on the initial melting range and the formula of the suppository, this phenomenon may affect the melting of the suppository and subsequent drug absorption rates. The softening time test and differential scanning calorimetry can be used as stability-indicating test methods to predict problems of this sort. Storage immediately after manufacture at an elevated temperature below the melting range speeds up the aging process. Since the hardening phenomenon is a finite process, this tempering approach can minimize further changes in melting range, which may be worth the addition to manufacturing cycle time.

The suppository overwrap foil also can cause problems in time. For example, if the suppository contains an acid, the foil wrapping may be attacked and develop pinholes.

Stability studies of suppositories intended for tropical climates must be conducted in the final package at temperatures at which the suppositories will eventually be kept. High-melting bases, water-soluble bases, and special polyethylene shell packages must be considered. Labeling should emphasize storage in a cool place. Efforts should be made in formulating suppositories for the tropics to maintain the physical and chemical stability of these suppositories in their final package, even when they are stored at temperatures as high as 50°C (122°F).

Storage studies also should include anticipated problems resulting from shipment. To test the effects of handling the product in the field,

suppositories often are shipped by the desired transport facilities to several areas in the country, and then tested physically, and occasionally chemically, for stability. Cool conditions for shipment often are required.

## EVALUATION OF SUPPOSITORIES

The literature is well documented with test methods to assure that each manufactured lot of suppositories consistently meets the standards established during the manufacture of early experimental lots. Finished suppositories are routinely inspected for appearance, and after being sliced lengthwise, for uniformity of the mix. They are assayed for active ingredients to ensure that they individually conform to labeled content. Melting range tests are performed to check the physical and absorption characteristics of each manufactured batch. Fragility tests are carried out to ascertain that the suppositories can be packaged and shipped with minimal breakage.

## Melting Range Test

This test is also called the *macromelting* range test and is a measure of the time it takes for the entire suppository to melt when immersed in a constant-temperature (37°C) water bath. In contrast, the *micromelting* range test is the melting range measured in capillary tubes for the fat base only. The suppository melting point apparatus by ERWEKA® consists of a graduated tube like glass test chamber. The sample to be tested is placed in a spiral shaped glass test basket inside the test chamber which itself is surrounded by a water jacket heated by circulation thermostat (Fig. 20.5). The time for the entire suppository to melt or disperse in the surrounding water is measured.

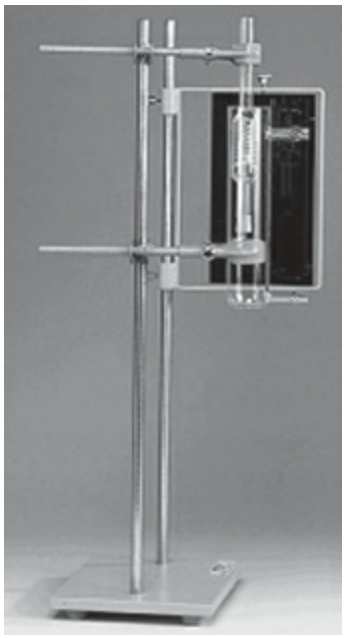


**Fig. 20.5:** Suppository melting point apparatus (*Courtesy* by permission from ERWEKA®)

## Softening Time Tests of Suppositories

Softening time test apparatus consists of a U-tube partially submersed in a constant-temperature water bath. A constriction on one side holds the suppository in place in the tube. A glass rod is placed on top of the suppository, and the time for the rod to pass through to the constriction is recorded as the "softening time". This can be carried out at various temperatures from 35.5 to 37°C, as a quality control check and can also be studied as a measure of physical stability over time. A water bath with both cooling and heating elements should be used to assure control within 0.1° C. The PM 30 Penetration Tester by ERWEKA® has been designed to carry out reproducible measurements of the softening time of suppositories at predetermined temperatures (Fig. 20.6). The construction of the tester makes visual observation of the melting characteristics extremely simple. The unit comprises 3 Test Stations.


**Fig. 20.6:** Assembly for measuring the softening time of suppositories (*Courtesy* by permission from ERWEKA®)

## Breaking Test

Brittleness of suppositories is a problem for which various solutions have already been described. The breaking test is designed as a method for measuring the fragility or brittleness of suppositories. The apparatus used for the test consists of a double-wall chamber in which the test suppository is placed (Fig. 20.7). Water at 37°C is pumped through the double walls of the chamber, and the suppository, contained in the dry inner chamber, supports a disc to which a rod is attached. The other end of the rod consists of another disc to which weights are applied. The test is conducted by placing 600 g on the platform. At 1 min intervals, 200 g weights are added, and the weight at which the suppository collapses is the breaking point, or the force that determines the fragility or brittleness characteristics of the suppository. Differently shaped suppositories have different breaking points. The desired breaking point of each of these variously shaped suppositories is established as the level that withstands the break forces caused by various types of handling, i.e. production, packaging, shipping, and patient in-use handling.

**Fig. 20.7:** Breaking test apparatus (*Courtesy* by permission from ERWEKA®)

## Disintegration/Dissolution Testing

Testing for the rate of in vitro release of drug substances from suppositories has always posed a difficult problem, owing to melting, deformation, and dispersion in the dissolution medium. Early testing was carried out by simple placement in a beaker containing a medium.

In an effort to control the variation in mass/medium interface, various means have been employed, including a wire mesh basket, or a membrane, to separate the sample chamber from the reservoir. Samples sealed in dialysis tubing or natural membranes have also been studied. Flow cell apparati have been used, holding the sample in place with cotton, wire screening, and most recently with glass beads. The basic suppository disintegration tester by ERWEKA® that complies with the latest EP specifications is shown in Fig. 20.8.

**Fig. 20.8:** Disintegration test apparatus (*Courtesy* by permission from ERWEKA®)

# 21:
# Pharmaceutical Aerosols

The packaging of therapeutically active ingredients in a pressurized system is not new to the pharmaceutical industry. According to present day usage, an aerosol or pressurized package is defined as “a system that depends on the power of a compressed or liquefied gas to expel the contents from the container.” It is in light of this definition that the terms aerosol, pressure package, pressurized package, and other similar terms are used in this chapter.

Although pressurized packages existed during the early 1900s, it was not until 1942, when the first aerosol insecticide was developed by Goodhue and Sullivan of the United States Department of Agriculture, that the aerosol industry was begun. The principles of aerosol technology were applied to the development of pharmaceutical aerosols in the early 1950s. These aerosol products were intended for topical administration for the treatment of burns, minor cuts and bruises, infections, and various dermatologic conditions. Aerosol products intended for local activity in the respiratory tract appeared in 1955, when epinephrine was made available in a pressurized package. Based on their acceptability to both patient and physician, and their widespread use, pharmaceutical aerosols represent a significant dosage form and should be considered along with other dosage forms, such as tablets, capsules, solutions, etc.

An examination of the aerosol dosage form reveals the following specific advantages over other dosage forms:

1. A dose can be removed without contamination of remaining material. Stability is enhanced for those substances adversely affected by oxygen and/or moisture. When sterility is an important factor, it can be maintained while a dose is being dispensed.
2. The medication can be delivered directly to the affected area in a desired form, such as spray, stream, quick-breaking foam, or stable foam.
3. Irritation produced by the mechanical application of topical medication is

reduced or eliminated.

4. Rapid onset of action, circumvention of the first pass effect and avoidance of degradation in the GI tract is achieved.
5. Dose lowering in case of steroid therapy and dose titration to individual needs can be achieved by using metered dose and dry powder inhalers.
6. Alternate route of administration is provided in case of drugs which shows erratic pharmacokinetics upon oral or parenteral administration and which may interact chemically or physically with other medicinals needed concurrently.

Other advantages are tamperproof systems, ease and convenience of application and application of medication in a thin layer.

## COMPONENTS OF AEROSOL PACKAGE

An aerosol product consists of the following component parts (1) propellant, (2) container, (3) valve and actuator and (4) product concentrate.

## Propellants

The propellant is responsible for developing the proper pressure within the container, and it expels the product when the valve is opened and aids in the atomization or foam production of the product. Various types of propellants are utilized which can be broadly classified as liquefied gases, hydrocarbons, hydrocarbon ether and compressed gases. Listed in Table 21.1 are the commonly used propellants together with several of their physicochemical properties. Those properties of particular interest to the industrial pharmacist have been included.

**Table 21.1:** Physicochemical properties of commonly used propellants

| Propellant class and chemical name | Molecular formula | Numerical designation | Solubility in water (weight %) 77°F | Vapour pressure 70°F | Vapour pressure 130°F | Boiling point °F | Boiling point °C | Liquid density |
|---|---|---|---|---|---|---|---|---|
| **Chlorofluorocarbons (CFCs)** | | | | | | | | |
| Trichloromonofluoromethane | $CCl_3F$ | 11 | 0.11 | 13.4 | 39.0 | 74.7 | 23.7 | 1.49 |
| Dichlorodifluoromethane | $CCl_2F2$ | 12 | 0.028 | 84.9 | 196.0 | –21.6 | –29.8 | 1.33 |
| Dichlorotetrafluoroethane | $CClF_2$ $CClF_2$ | 114 | 0.013 | 27.6 | 63.5 | 38.4 | 3.6 | 1.47 |
| **Hydrochlorofluorocarbons (HCFCs)** | | | | | | | | |
| Difluoroethane | $CH_3CHF_2$ | 152a | < 1.0 | 76.4 | 191.0 | –11.2 | –24.0 | 0.91 |
| Hydrofluorocarbons (HFCs) | | | | | | | | |
| Tetrafluoroethane | $CF_3CH_2F$ | 134a | 0.150 | 71.1 | 198.7 | –15.0 | –26.2 | – |
| Heptafluoropropane | $CF_3CHFCF_3$ | 227 | 0.058 | – | – | –3.2 | –16.5 | – |
| **Hydrocarbons (HCs)** | | | | | | | | |
| Butane | $C_4H_{10}$ | A-17 | – | 31.6 | 82.0 | 31.1 | -0.6 | 0.58 |
| Isobutane | $C_4H_{10}$ | A-31 | – | 45.8 | 111.0 | 10.9 | –11.8 | 0.56 |
| Propane | $C_3H_8$ | A-108 | – | 122.8 | 270.7 | –43.7 | –44.6 | 0.50 |
| Hydrocarbon Ethers (HEs) | | | | | | | | |
| Dimethyl ether | $CH_3OCH_3$ | – | – | 63.0 | – | –13.0 | – | 0.66 |
| **Compressed gases (CGs)** | | | | | | | | |
| Carbon dioxide | $CO_2$ | – | 0.7[a] | 852.0 | – | –109[c] | – | – |
| Nitrous oxide | $N_2O$ | – | 0.5[a] | 735.0 | – | –127 | – | – |
| Nitrogen | $N_2$ | – | 0.014[a] | 492.0[b] | – | –320 | – | – |

[a]Volume of gas soluble in one volume of water at atmospheric pressure

[b]At the critical point (–233°F)

[c]Sublimes

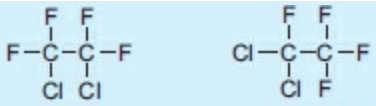
Liquefied gases are gases at room temperature and atmospheric pressure and can be liquefied easily by lowering the temperature or by increasing the pressure. When liquefied gases are placed into a sealed container, they immediately separates into a liquid and a Vapour phase. They have

widespread use as propellants, since they are very effective in dispersing the active ingredients in to a fine mist or foam, are relatively inert, nontoxic and non flammable. They also have the advantage of maintaining a constant pressure within the container. Chlorofluorocarbons (CFCs), hydro-chlorofluorocarbons (hydrochlorofluoro-carbons (HCFCs) and hydrofluorocarbons (hydrofluorocarbons (HFCs) are commonly used liquefied gases. All the fluorinated propellants are identified by a numerical designation. The rules governing this nomenclature system are as follows:

1. All the propellants are designated by three digits (000).
2. The first digit represents the one less than the number of carbon atoms ('C'-1) in the compound.
3. The second digit represents the one more than the number of hydrogen atoms ('H'+1) in the compound.
4. The last third digit represents the number of fluorine atoms in the compound.
5. The number of chlorine atoms (for CFCs) in the compound can be found by subtracting the sum of the fluorine and the hydrogen atoms from the total number of atoms that can be added to saturate the carbon chain.
6. In case of isomers, the most symmetrical one is indicated by the number alone. As the isomer becomes more and more asymmetrical, the letter a, b, c, etc. follows the number.
7. For cyclic compounds, a lowercase c is used before the number.

The use of this system can be illustrated by following example:

Dichlorotetrafluoroethane ($C_2Cl_2F_4$) has two carbon atoms, so first digit will be 1 ('C'-1), has no hydrogen atom, so second digit will be 1 ('H'+1) and has four fluorine atoms, so third digit will be 4. But it has two chlorine atoms too, which can be arranged in two different ways, leading to two isomeric forms.




Isomer (i) Isomer (ii)

Isomer (i) is more symmetric than isomer (ii), so it will be numerically designated as propellant 114 and the other one as propellant 114a.

While the fluorinated hydrocarbons find widespread use in most aerosols

for oral and inhalation use, topical pharmaceutical aerosols utilize hydrocarbons, hydrocarbon ether and compressed gases. As compared to fluorinated hydrocarbons, hydrocarbons are flammable, less toxic, less dense, more economic, more soluble, and chemically more stable. However, compressed gases have little expansion power and will produce a fairly wet spray and less stable foam as compared to liquefied gases. Compressed gases are used in products like hair preparations, ointments, dental creams, aqueous germicidal and antiseptic aerosols, contact lenses cleaner and food products.

Blends of various fluorocarbon propellants are generally used for pharmaceutical aerosols and are indicated in Table 21.2. By varying the proportion of each component, desired vapour pressure can be achieved within the limits of the vapour pressure of the individual propellants.

**Table 21.2:** Blends of fluorocarbon propellants for pharmaceutical aerosols

| Propellant blend* | Composition | Vapour pressure (psig) 70°F | Density (g/ml) 70°F |
|---|---|---|---|
| 12/11 | 50:50 | 37.4 | 1.412 |
| 12/11 | 60:40 | 44.1 | 1.396 |
| 12/114 | 70:30 | 56.1 | 1.368 |
| 12/114 | 40:60 | 39.8 | 1.412 |
| 12/114 | 45:55 | 42.8 | 1.405 |
| 12/114 | 55:45 | 48.4 | 1.390 |

*It is generally understood that the designation “propellant 12/114 (70:30)” indicates a composition of 70% by weight of propellant 12 and 30% by weight of propellant 114

As with the fluorocarbons, a range of pressures can be obtained by mixing the various hydrocarbons in varying proportions. Since the hydrocarbons are naturally occurring products, however, their purity varies, and the blending is done on the basis of the desired final pressure and not on the percentage of each component present. The pressure of each individual component varies somewhat, depending on the degree of purity. Table 21.3 illustrates some of the commonly used blends that are commercially available.

**Table 21.3:** Vapour pressure of hydrocarbons

| Propellant blend* | Composition | Vapour pressure (psig) 70°F | Density (g/ml) 70°F |
|---|---|---|---|
| 12/11 | 50:50 | 37.4 | 1.412 |
| 12/11 | 60:40 | 44.1 | 1.396 |
| 12/114 | 70:30 | 56.1 | 1.368 |
| 12/114 | 40:60 | 39.8 | 1.412 |
| 12/114 | 45:55 | 42.8 | 1.405 |
| 12/114 | 55:45 | 48.4 | 1.390 |

The vapour pressure of a mixture of propellants can be calculated according to Dalton's law, which states that the total pressure in any system is equal to the sum of the individual or partial pressures of the various components. Raoult's law, which regards lowering of the vapour pressure of a liquid by the addition of another substance, states that the depression of the vapour pressure of a solvent upon the addition of a solute (something added to the solvent) is proportional to the mole fraction of solute molecules in the solution. Given ideal behavior, the vapour pressure of a mixture consisting of two individual propellants is equal to the sum of the mole fraction of each component present multiplied by the vapour pressure of each pure propellant at the desired temperature. This relationship can be shown mathematically:

$$P_a = \frac{n_a}{n_a + n_b} P_A o = N_A P_A o \qquad \text{... (1)}$$

where:

$P_a$ = partial vapour pressure of propellant A

$P_A o$ = vapour pressure of pure propellant A

$n_a$ = moles of propellant A

$n_b$ = moles of propellant B

$N_A$ = mole fraction of component A

To calculate the partial vapour pressure of propellant B:

$$P_b = \frac{n_b}{n_a + n_b} P_B o = N_B P_B o \qquad \text{... (2)}$$

The total vapour pressure of the system is then obtained from:

$$P = P_a + P_b \text{ ... (3)}$$

where, $P$ is the total vapour pressure of the system.

When one component is present in relatively low concentration, ideal behavior is approached. For practical purposes, however, the calculated pressure is sufficiently accurate for most determinations. The application of Raoult's law for calculation of vapour pressure can best be illustrated by the following example.

Calculate the vapour pressure at 70°F of a propellant blend consisting of propellant 12/11 (30:70) (see Table 21.2).

*Moles of each substance:*

$$moles_{11} = \frac{Weight_{11}}{MW_{11}} = \frac{70}{137.38} = 0.5095$$

0.5095 moles of propellant 11.

$$moles_{12} = \frac{Weight_{12}}{MW_{12}} = \frac{30}{120.93} = 0.2481$$

0.2481 moles of propellant 12.

*From Roault's law:*

$$P_{11} = \frac{n_{11}}{n_{11} + n_{12}} P_{11}o$$

$$P_{11} = \frac{0.5095}{0.5095 + 0.2481} 13.4 = 9.01\ psia$$

9.01 psia* partial pressure of propellant 11.

$$P_{12} = \frac{n_{12}}{n_{12} + n_{11}} P_{12}o = \frac{0.2481}{0.2481 + 0.5095} \times 84.9$$

= 27.80 psia

Vapour pressure of propellant 12/11 (30:70) then equals 27.80 + 9.01, or 36.81 psia. Gauge pressure is obtained from:

psia – 14.7 = psig (Or) 36.81 – 14.7 = 22.11 psig

A difference is noted in comparing the calculated value for the vapour pressure with the experimental value (*see* Table 21.2). The difference is due to deviation from ideal behavior.

## Containers

Various materials as indicated in the following outline have been used for the manufacture of aerosol containers, which must withstand pressures as high as 140 to 180 psig at 130°F.

A. Metal
   1. Tinplated steel
      a. Side-seam (three-piece)
      b. Two-piece or drawn
      c. Tin-free steel
   2. Aluminum
      a. Two-piece
      b. One-piece (extruded or drawn)
   3. Stainless steel

B. Glass
   1. Uncoated glass
   2. Plastic-coated glass

## *Tinplate Containers*

The tinplated steel container consists of a sheet of steel plate that has been electroplated on both sides with tin. The thickness of the tin coating is described in terms of its weight, for example, #25, #50 and #100. The size of the container is indicated by a standard system, which is a measure of the diameter and height of the container. A container said to be 202 × 214 if, it is 2 3/16 inches in diameter and 2 14/16 inches in height.

Brief discussion of the procedure used in the manufacture of tinplated containers might be advantageous for a better appreciation of the quality control aspects. Tinplated steel is obtained in thin sheets and when required, it is coated with an organic material. These sheets are lithographed at this point. After the sheet is cut into sizes to make a body, a top, and a bottom, each piece is fabricated into the desired shape. The body is shaped into a cylinder and seamed via a flanging and soldering operation. The top and bottom are attached to the body and a side seam stripe is added to the inside

seam area when required. The organic coating also can be added to the finished container rather than to the flat sheets. This procedure is slower and somewhat more expensive, but a more continuous and durable coating is produced. The use of sealing compounds, types of solder, and organic coatings are discussed in this chapter under the heading “Formulation of Pharmaceutical Aerosols”.

A recent development in metal tinplate containers is the welded side-seam. Welding eliminates the soldering operation, saves considerable manufacturing time and decreases the possibility of product/container interaction. In general, two processes are used: the Soudronic system (American Can Company, Crown Cork and Seal, and the Southern Can Company) and the Conoweld system (Continental Can Company). The Soudronic system is based on an electronically controlled resistance welding method that uses a copper wire as an electrode. The rounded bodies are welded and then sent to the conventional line, where the top and bottom ends are flanged as indicated previously. The Conoweld system passes the folded body through two rotating electrode rings. The rest of the container is manufactured in the usual manner.

## *Aluminum Containers*

Aluminum is used to manufacture extruded (seamless) aerosol containers. Many existing pharmaceuticals are packaged in aluminum containers, probably because of the lessened danger of incompatibility due to its seamless nature and greater resistance to corrosion. Aluminum can be rather unpredictable, however, in that it is corroded by pure water and pure ethanol. The combination of ethanol and propellant 11 in an aluminum container has been shown to produce hydrogen, acetyl chloride, aluminum chloride, propellant 21 and other corrosive products.

## *Stainless Steel Containers*

These containers are limited to the smaller sizes, owing to production problems as well as cost. They are extremely strong and resistant to most materials. Stainless steel containers have been used for inhalation aerosols. In most cases, no internal organic coating is required.

## *Glass Containers*

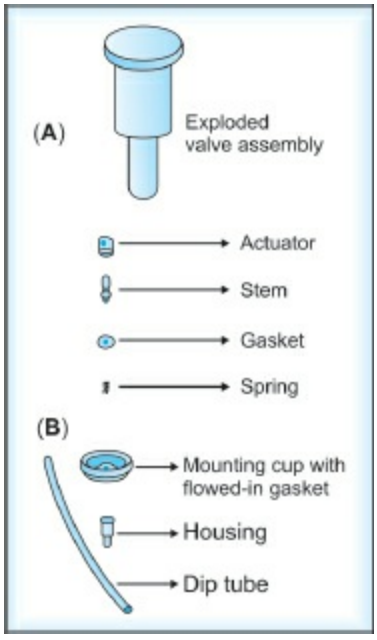
Glass aerosol containers have been used for a large number of aerosol pharmaceuticals. Glass containers are available with or without plastic coatings. The plastic coating may be totally adhered (except for the neck ring) or nonadhered and vented. By adjusting the formulation and limiting the type and quantity of propellant, satisfactory aerosol products can be formulated and packaged in glass containers. Glass aerosol containers are preferable from a compatibility viewpoint, since corrosion problems are eliminated. The use of glass also allows for a greater degree of freedom in design of the container.

## Valves

The presentday aerosol valve is multifunctional in that it is capable of being easily opened and closed and in addition, is capable of delivering the content in the desired form. Furthermore, especially in the case of pharmaceuticals, the valve is expected to deliver a given amount of medication. Valves for pharmaceuticals usually do not differ from the valves used for nonpharmaceutical aerosol products, but the requirements for pharmaceuticals are usually more stringent than for most other products. The materials used in the construction of the valve must be approved by the Food and Drug Administration. Pharmaceutical aerosols may be dispensed as a spray, foam, or solid stream, and they may or may not require dosage control. The need for several different types of valves becomes apparent.

### *Continuous Spray Valves*

An aerosol valve consists of many different parts and is assembled using high-speed production techniques. The valve manu facturers adhere to relatively close tolerances during manufacture and assembly of the valve. Various materials are used to manufacture the many components of the valve. Figures 21.1 and 21.2 illustrate typical aerosol valve assemblies for use with cans or bottles, respectively. These valve assemblies consist of the following parts:



**Figs 21.1A and B:** (A) Vapour tap body; (B) Aerosol can valve assembly (*Courtesy*of Precision Valve Corporation, Yonkers, NY)

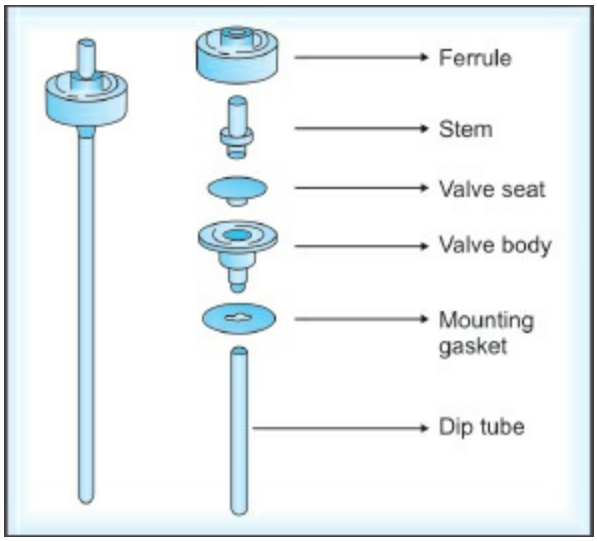



**Fig. 21.2:** Aerosol bottle valve assembly (*Courtesy* of Risdon Manufacturing Co., Naugatuck, CT)

## *Ferrule or Mounting Cup*

The ferrule or mounting cup is used to attach the valve proper to the container. For use with containers having a one-inch opening, the cup is made from tinplate steel, although aluminum also can be used. Since the underside of the valve cup is exposed to the contents of the container and to the effects of oxygen trapped in the head space, a single or double epoxy or vinyl coating can be added to increase resistance to corrosion. Ferrules are used with glass bottles or small aluminum tubes and are usually made from a softer metal such as aluminum or brass. The ferrule is attached to the container either by rolling the end under the lip of the bottle or by clinching the metal under the lip.

## *Valve Body or Housing*

The housing is generally manufactured from Nylon or Delrin and contains an opening at the point of the attachment of the dip tube, which ranges from about 0.013 inch to 0.080 inch.

The housing may or may not contain another opening referred to as the "vapour tap". The vapour tap allows for the escape of vapourized propellant

along with the liquid product (*see* Fig. 21.1A). The vapour tap further produces a fine particle size, prevents valve clogging with products containing insoluble materials, allows for the product to be satisfactorily dispensed with the container in the inverted position, reduces the chilling effect of the propellant on the skin, and in the case of hydrocarbon propellants, allows for a decrease in flame extension. These vapour tap openings are available in sizes ranging from about 0.013 inch to 0.080 inch.

**Stem:** The stem is made from Nylon or Delrin, but metals such as brass and stainless steel can also be utilized. One or more orifices are set into the stem; they range from one orifice of about 0.013 inch to 0.030 inch, to three orifices of 0.040 inch each.

**Gasket:** Buna-N and Neoprene rubber are commonly used for the gasket material and are compatible with most pharmaceutical formulations.

**Spring:** The spring serves to hold the gasket in place, and when the actuator is depressed and released, it returns the valve to its closed position. Stainless steel can be used with most aerosols.

**Dip tube:** Dip tubes are made from polyethylene or polypropylene. Both materials are acceptable for use although the polypropylene tube is usually more rigid. The inside diameter of the commonly used dip tube is about 0.120 inch to 0.125 inch, although capillary dip tubes are about 0.050 inch, and dip tubes for highly viscous products may be as large as 0.195 inch. Viscosity and the desired delivery rate play an important role in the selection of the inner diameter of the dip tube.

**Metering valves:** Metering valves are applicable to the dispensing of potent medication. These operate on the principle of a chamber whose size determines the amount of medication dispensed. This is shown in Fig. 21.3. Although these have been used to a great extent for aerosol products, they are limited in both size and accuracy of dosage. Approximately 50 to 150 mg ±10% of liquid material can be dispensed at one time with the use of such valves.

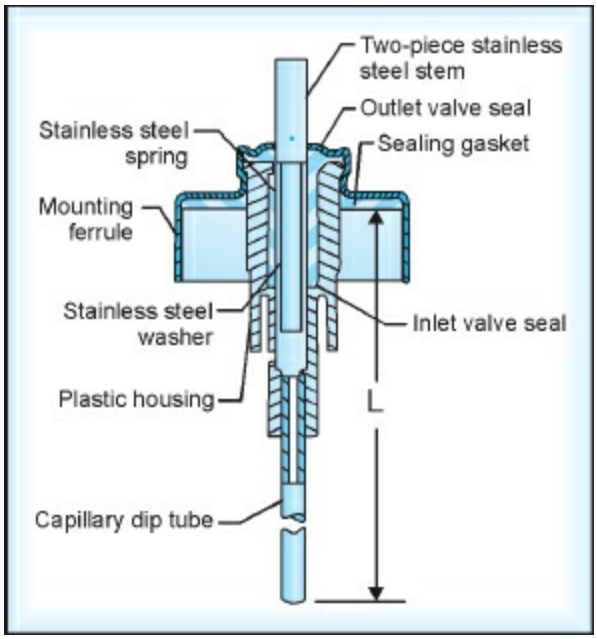



**Fig. 21.3:** Metering valve for pharmaceutical aerosols (*Courtesy* of Emson Research, Bridgeport, CT)

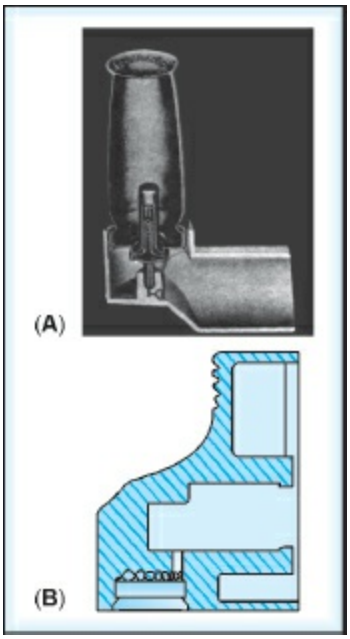
## Actuators

To ensure that the aerosol product is delivered in the proper and desired form, a specially designed button or actuator must be fitted to the valve stem. The actuator allows for easy opening and closing of the valve and is an integral part of almost every aerosol package. It also serves to aid in producing the required type of product discharge.

There are many different types of actuators. Among them are those that produce (1) spray, (2) foam, (3) solid stream and (4) special applications.

### *Spray Actuators*

Figure 21.4 illustrates actuators that are capable of dispersing the stream of product concentrate and propellant into relatively small particles by allowing the stream to pass through various openings (of which there may be one to three on the order of 0.016 inch to 0.040 inch in diameter-Fig. 21.4A). Where there is a large percentage of propellant mixture containing a sufficient quantity of a low boiling propellant such as propellant 12 or propane, actuators having relatively large orifices can be used. The combination of propellant vapourization and actuator orifice and internal channels can deliver the spray in the desired particle size range. A spray type actuator can be used with pharmaceuticals intended for topical use, such as spray-on bandages, antiseptics, local anesthetics, and foot preparations. When these actuators are used with aerosol products containing relatively low amounts of propellants (50% or less), the product is dispensed as a stream rather than as a spray, since the propellant present in the product is not sufficient to disperse the product fully. For these products, a mechanical breakup actuator is usually required (Fig. 21.4B). These actuators are capable of “mechanically” breaking a stream into fine particles by causing the stream to “swirl” through various channels built into the actuator.




**Figs 21.4A and B:** Actuators for pharmaceutical aerosols: (A) Inhalation spray type (*Courtesy* of Riker Laboratories); (B) Mechanical breakup type (*Courtesy* of Risdon Manufacturing Co., Naugatuck, CT)

## *Foam Actuators*

These actuators consist of relatively large orifices ranging from approximately 0.070 inch to 0.125 inch and greater (Fig. 21.5). The orifices allow for passage of the product into a relatively large chamber, where it can expand and be dispensed through the large orifice.


**Fig. 21.5:** Various actuators and applicators for pharmaceuticals (*Courtesy* of Pechiney Ugine Kuhlman Development, Inc., Greenwich, CT)

## *Solid-Stream Actuators*

The dispensing of such semisolid products as ointments generally requires these actuators. Relatively large openings allow for the passage of product through the valve stem and into the actuator. These are essentially similar to foam type actuators.

## *Special Actuators*

Many of the pharmaceutical and medicinal aerosols require a specially designed actuator to accomplish a specific purpose. They are designed to deliver the medication to the appropriate site of action-throat, nose, eye, or vaginal tract. Several are shown in Fig. 21.5.

## Formulation of Pharmaceutical Aerosols

An aerosol formulation consists of two essential components: product concentrate and propellant. The product concentrate consists of active ingredients, or a mixture of active ingredients, and other necessary agents such as solvents, antioxidants, and surfactants. The propellant may be a single propellant or a blend of various propellants; it can be compared with other vehicles used in a pharmaceutical formulation. Just as a blend of solvents is used to achieve desired solubility characteristics, or various surfactants are mixed to give the proper HLB value for an emulsion system, the propellant is selected to give the desired vapour pressure, solubility and particle size.

Since one must be familiar with the physicochemical properties of surfactants, solvents, and suspending agents, it follows that the formulator of aerosol preparations must be thoroughly familiar with propellants and the effect the propellant will have upon the finished product. Propellants can be combined with active ingredients in many different ways, producing products with varying characteristics. Depending on the type of aerosol system utilized, the pharmaceutical aerosol may be dispensed as a fine mist, wet spray, quick-breaking foam, stable foam, semisolid, or solid. The type of system selected depends on many factors, including the following (1) physical, chemical, and pharmacologic properties of active ingredients and (2) site of application.

## TYPES OF SYSTEMS

### Solution System/Two-phase System

A large number of aerosol products can be formulated in this manner. This system is also referred to as a two-phase system and consists of a vapour and liquid phase. When the active ingredients are soluble in the propellant, no other solvent is required. Depending on the type of spray required, the propellant may consist of propellant 12 or A-70 (which produce very fine particles), or a mixture of propellant 12 and other propellants as indicated in Tables 21.1 to 21.3. As other propellants with vapour pressures lower than that of propellant 12 are added to propellant 12, the pressure of the system decreases, resulting in the production of larger particles. A lowering of the vapour pressure also is produced through the addition of less volatile solvents such as ethyl alcohol, propylene glycol, ethyl acetate, glycerin and acetone. The amount of propellant used may vary from 5% (for foams) to 95% (for inhalation products) of the entire formulation. When a spray is produced with larger particles, a decrease is noted in the number of fine particles, decreasing the danger of inhaling these materials through formation and subsequent inhalation of airborne particles. These sprays are also useful for topical preparations, since they tend to coat the affected area with a film of active ingredients. Depending on the boiling point of the solvent used, the rate of vapourization of the propellant is decreased, thereby increasing any chilling effect that may be present. This system can best be exemplified by the following general formulations:

| | **Weight %** |
|---|---|
| Active ingredients | 10–15 |
| Propellant 12/11 (50:50) up to | 100 |

Propellant 12/11 (30:70), propellant 12/114 (45:55), or propellant 12/114 (55:45) also can be utilized for oral inhalation aerosols or other FDA-exempted products such as contraceptive foams. As the amount of propellant 12 is increased, the pressure increases. With the exception of propellant 12/11 (30:70), the pressure of these systems necessitates packaging the contents in a metal container. If the product is to be packaged in a glass container, a mixture of propellant 12/114 (20:80) or (10:90) can be used. Table 21.4

indicates the pressure limitations in using various aerosol containers.

Aerosols intended for inhalation or for local activity in the respiratory system in the treatment of asthma may be formulated as follows:

| | **Weight %** |
|---|---|
| Isoproterenol HCl | 0.25 |
| Ascorbic acid | 0.10 |
| Ethanol | 35.75 |
| Propellant 12 | 63.90 |

This type of formulation is usually packaged in a 15 to 30 ml stainless steel, aluminum, or glass container. Since propellant 12 has a relatively high vapour pressure, the addition of propellant 114 is recommended in order to reduce the pressure, as illustrated by the following example:

**Table 21.4:** Pressure limitations of nonrefillable aerosol containers

| **Container** | **Maximum pressure (psig)** | **Temperature °F** |
|---|---|---|
| Tinplated steel | | |
| Low pressure | up to 140 | 130 |
| DOT “2P” | from 140 to 160 | 130 |
| DOT “2Q” | from 160 to 180 | 130 |
| Uncoated glass | less than 18 | 70 |
| Coated glass | less than 25 | 70 |
| Aluminum | up to 180 | 130 |
| Stainless steel | up to 180 | 130 |
| Plastic | less than 25 | 70 |

Department of Transportation (DOT) Regulations

| | **Weight %** |
|---|---|
| Octyl nitrite | 0.1 |

| | |
|---|---|
| Ethanol | 20.0 |
| Propellant 114 | 49.2 |
| Propellant 12 | 30.7 |

Hydrocarbons in topical aerosol pharmaceutical preparations are used as follows:

| | **Weight %** |
|---|---|
| Active ingredients | up to 10–15 |
| Solvents such as ethanol or propylene glycol | up to 10–15 |
| Distilled water | 10–15 |
| Hydrocarbon propellant A-46 | 55–70 |

Depending on the amount of water present, the final product may be a solution or a three-phase system. Solution aerosols produce a fine to coarse spray, depending on the concentration of the other ingredients. Hydrocarbon propellant A-70 produces a drier particle, while propellants A-17 and A-31 tend to produce a wetter spray. Hydrocarbon propellants can also be used for products packaged in plastic-coated glass bottles, provided that the amount of flammable hydrocarbon propellant present does not exceed 15% of the total product weight and that the container has a volumetric capacity not exceeding 5 fluid ounces. In addition, one of every 1000 bottles must be tested to 250 psig without failure. The manufacturer of the aerosol product must test one bottle out of each lot of 20,000 bottles to the bursting point, and the bursting pressure must not be less than 300 psig. One fully charged bottle from this lot must be dropped to an unyielding surface from a height of 4 feet without producing flying glass or a shattering effect. Should either of these two bottles fail, then 10 additional bottles must be tested for each failed test. Failure of any of these 10 samples would cause the entire lot to be rejected. Finally, one fully charged bottle out of each 1000-bottle lot must be heated so that the pressure in the container is equivalent to the equilibrium pressure of the contents at 130°F, without evidence of leakage or other defect.

## Water-based System

### *Three-phase System*

Relatively large amounts of water can be used to replace all or part of the nonaqueous solvents used in aerosols. These products are generally referred to as “water-based” aerosols, and depending on the formulation, are emitted as a spray or foam. To produce a spray, the formulation must consist of a dispersion of active ingredients and other solvents in an “emulsion” system in which the propellant is in the external phase. In this way, when the product is dispensed, the propellant vapourizes and disperses the active ingredients into minute particles. Since propellant and water are not miscible, a three-phase aerosol forms (propellant phase, water phase, and vapour phase). Ethanol has been used as a cosolvent to solubilize some of the propellant in the water. By virtue of its surface-tension-lowering properties, ethanol also aids in the production of small particles.

Surfactants have been used to a large extent to produce a satisfactory homogeneous dispersion. Surfactants that possess low water solubility and high solubility in nonpolar solvents have been found to be the most useful. Long-chain fatty acid esters of polyhydroxylic compounds including glycol, glycerol, and sorbitol esters of oleic, stearic, palmitic, and lauric acids exemplify this series. In general, about 0.5 to 2.0% of surfactant is used. The propellant content varies from about 25 to 60%, but can be as low as 5%, depending on the nature of the product.

To achieve the desired fine particle size with products containing large amounts of water and a low proportion of propellant, a mechanical breakup actuator must be used along with a “vapour tap valve”.

### *Aquasol System*

A recent development that is useful for pharmaceutical aerosols is the Aquasol valve. The new Aquasol system allows for the dispensing of a fine mist or spray of active ingredient dissolved in water, which is not possible with the usual three-phase system. Since only active ingredient and water are dispensed (propellant is in vapour state and present only in extremely small quantity), there is no chilling effect as occurs with the hydrocarbon propellant.

The Aquasol system is illustrated in Figs 21.6 and 21.7. It is designed to dispense pressurized products efficiently and economically using relatively small amounts of hydrocarbon propellant; however, it can also function effectively using fluorocarbon propellants. This system, which is essentially a "three-phase" aerosol, permits the use of fairly large quantities of water in the formulation.

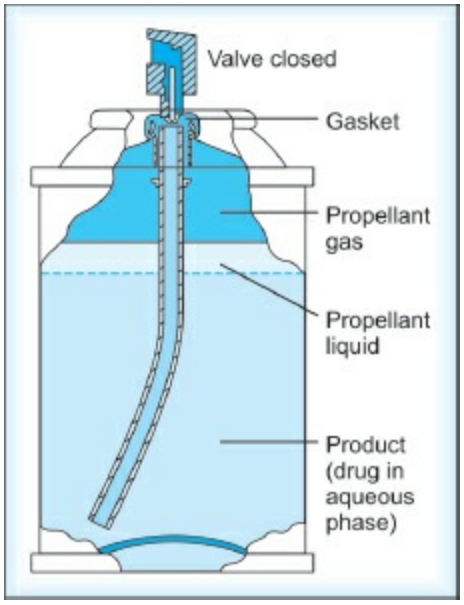



**Fig. 21.6:** The Aquasol dispenser system-valve closed (*Courtesy* of Precision Valve Corporation, Yonkers, NY)

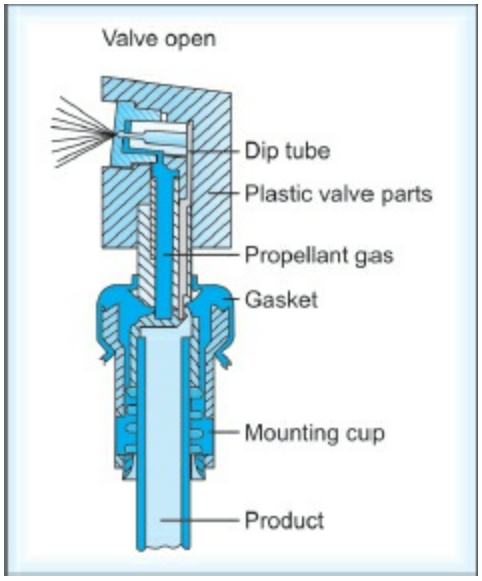



**Fig. 21.7:** The Aquasol dispenser system-valve open (*Courtesy* of Precision

Valve Corporation, Yonkers, NY)

The chief difference between the Aquasol system and the “three-phase” system is that the former dispenses a fairly dry spray with very-small particles. This relative dryness and small particle size are due chiefly to the design of the valve, which dispenses vapourized propellant rather than liquefied propellant. In addition, the vapourized propellant contributes to the nonflammability of the stream of product as it is dispensed. For example, a fine, almost dry spray is obtained using six parts of water with one part of hydrocarbon propellant. Not only is the resulting spray nonflammable, but it actually extinguishes an open flame.

As can be noted from Fig. 21.6, the active ingredient is dissolved or suspended in water or in a mixture of alcohol and water. The hydrocarbon propellant floats on top of the aqueous layer and exists as both a liquid and a vapour. Depending on the amount of alcohol present in the aqueous layer, the propellant and water/alcohol layer may or may not be immiscible. As the amount of alcohol increases, the miscibility of these two layers increases. As a pure alcohol system is approached, complete miscibility occurs, at which time a two-phase system that can function satisfactorily is produced. Flammability is increased, however, owing to the large amount of alcohol present, as well as to the fact that liquid propellant is now also being dispensed.

In the Aquasol system, the vapour phase of the propellant and the product enter the mixing chamber of the actuator through separate ducts or channels. Moving at tremendous velocity, the vapourized propellant enters into the actuator while the product is forced into the actuator by the pressure of the propellant. At this point, product and vapour are mixed with violent force, resulting in a uniform, finely dispersed spray.

Depending on the configuration of the valve and actuator, either a fine dry spray or a coarse, wet spray can be obtained. Previous studies have shown that a fine dry spray is obtained when a ratio of about 6 parts of product to 1 part of propellant is used. Up to 30 parts of product to one part of propellant has also produced a satisfactory spray, but in this case, a more coarse spray results. In the Aquasol system, it is almost impossible to dispense only the pure propellant until the package is depleted of the aqueous product.

## Suspension or Dispersion Systems

Various methods have been used to overcome the difficulties encountered that are due to the use of a cosolvent. One such system involves a dispersion of active ingredients in the propellant or a mixture of propellants. To decrease the rate of settling of the dispersed particles, various surfactants or suspending agents have been added to the systems. These systems have been developed primarily for use with oral inhalation aerosols. Several examples follow:

| | **Weight %** |
|---|---|
| Epinephrine bitartrate (within 1 to 5 microns) | 0.50 |
| Sorbitan trioleate | 0.50 |
| Propellant 114 | 49.50 |
| Propellant 12 | 49.50 |

The epinephrine bitartrate hasa minimumsolubility in the propellant system, but issufficiently soluble in the fluidsin the lungsto exert a therapeutic activity.

| | **Weight %** |
|---|---|
| Isoproterenol sulfate | 33.3 mg |
| Oleyl alcohol | 33.3 mg |
| Myristyl alcohol | 33.4 mg |
| Propellant 12 | 7.0 g |
| Propellant 114 | 7.0 g |

The isoproterenol sulfate remains dispersed in the propellant vehicle for a sufficient length of time to allow for the dispensing of a suitable dose. The physical stability of an aerosol dispersion can be increased by (1) control of moisture content, (2) use of derivatives of active ingredients having minimum solubility in propellant system (the pharmacologic activity must also be considered), (3) reduction of initial particle size to less than 5 microns, (4) adjustment of density of propellant and/or suspensoid so that they are equalized and (5) use of dispersing agents.

A formulation for oral inhalation that contains a steroid and is used to alleviate the symptoms of asthma would include:

| | Weight % |
|---|---|
| Steroid compound | 8.4 mg |
| Oleic acid | 8.4 mg |
| Propellant 11 | 4.7 g |
| Propellant 12 | 12.2 g |

The oleic acid is present as a dispersing agent for the steroid and is an aid in the prevention or reduction of particle growth or agglomeration. In addition, it serves as a valve lubricant and prevents the metered valve from "sticking" in the open position.

Particles of certain materials tend to agglomerate immediately following suspension or shortly thereafter, owing to solubility, moisture, or particle size growth. Caking results when the aggregates become massive. The degree of agglomeration is accelerated at elevated temperatures. This phenomenon may range in degree from flocculation, in which the particles are loosely bound, to aggregation, in which the particles are held together tightly and often fused. In severe cases, the particles may adhere to the walls of the container. Agglomeration may result in valve clogging, inaccuracy of dosage, and depending on the nature of the active ingredients, damage to the liner and possibly to the metal container.

The moisture content of both the suspensoid and the propellant affects the stability of the aerosol system and must be below 300 ppm. Higher moisture levels generally result in particle agglomeration. So that this level is not exceeded, rigid control over the conditions of manufacture must be exerted, and drying both the suspensoid and the propellant prior to manufacture becomes necessary. This can be accomplished by passing the propellant through various desiccants.

Materials suspended in a vehicle in which they are partially soluble show signs of particle size growth. To decrease this phenomenon, a chemical derivative of the drug that shows minimum solubility in the propellant vehicle must be selected. From a physiological viewpoint, the drug must show some solubility in the fluids surrounding the lung tissue. For example, a

derivative such as epinephrine bitartrate is preferred when preparing a suspension of the drug in the propellant system. The hydrochloride or sulfate, however, is preferred when a solution aerosol is being formulated, since they are soluble in a hydroalcohol solution, which is miscible with the propellant.

The physical stability of a dispersed system depends primarily on the rate of agglomeration of the suspensoid. This rate is affected by the initial particle size of the drug, which must be in the range of 1 to 5 microns; the exact size depends on the nature of the active ingredient and intended use of the product. This particle size range is necessary to ensure that the particles reach the intended site of action. Suspensions containing materials intended for topical administration do not require particles of the same degree of dispersion as the inhalation aerosols, but for reasons of physical stability, the particles are seldom over 50 microns in size. The active ingredients may be reduced to required particle sizes through use of suitable grinding equipment.

Consideration must also be given to the density of the suspensoid and the propellant vehicle. By decreasing the difference between these two densities, the rate of settling of the suspensoid can be decreased. The density of both the propellant and/or suspensoid may be changed by the addition of a compound of higher or lower density, so that the density of the suspensoid may be made equal to the propellant density.

Various surfactants and lubricants have been investigated in an attempt to control the rate of agglomeration. Such agents as isopropyl myristate and mineral oil have been added to these aerosols to reduce agglomeration and to act as a lubricant for the particles in passing through the valve orifices. They have met with moderate success. The addition of surfactants to aerosol suspensions has been most successful. These surfactants exert their activity by coating each of the particles in suspension and become oriented at the solid-liquid interface. Agglomeration is reduced, thereby increasing stability by providing a physical barrier. According to investigations carried out by Young, Thiel, and Laursen, nonionic surfactants were found to be most effective. Those surfactants having an HLB less than 10, such as sorbitan trioleate, could be utilized for aerosol dispersions. Other agents that were found to be useful are sorbitan monolaurate, sorbitan monooleate, and sorbitan sesquioleate. These surfactants are effective in a concentration of 0.01 to 1%, depending on the concentration of the suspensoid and the intended use of the product.

Vapour tap valves also have been used with dispersion aerosols to decrease the danger of valve clogging. The added propellant, escaping as a vapour, aids in clearing the valve of solid particles.

## Foam Systems

Emulsion and foam aerosols consist of active ingredients, aqueous or nonaqueous vehicle, surfactant, and propellant, and are dispensed as a stable or quick-breaking foam, depending on the nature of the ingredients and the formulation. The liquefied propellant is emulsified and is generally found in the internal phase. Nonaerosol emulsions are usually in lotion or viscous liquid form, but aerosol emulsions are dispensed as foams, and this can be advantageous for various applications involving irritating ingredients, or when the material is applied to a limited area.

### *Aqueous Stable Foams*

These can be formulated as follows:

| | % w/w |
|---|---|
| Active ingredients<br>Oil-waxes<br>Oil in water surfactant<br>Water | 95.0–96.5 |
| Hydrocarbon propellant | 3.5–5.0 |

While the total propellant content may be as high as 5% in certain cases, it usually is about 8 to 10% v/v or 3 or 5% w/w. As the amount of propellant A-70, A-46, etc. increases, a stiffer and dryer foam is produced. Lower propellant concentrations yield wetter foams. One, type of system producing a stable foam may be illustrated by the following example:

| | % w/w |
|---|---|
| Myristic acid | 1.33 |
| Stearic acid | 5.33 |
| Cetyl alcohol | 0.50 |
| Lanolin | 0.20 |
| Isopropyl myristate | 1.33 |
| Triethanolamine | 3.34 |
| Glycerin | 4.70 |
| Polyvinylpyrrolidone | 0.34 |

| | |
|---|---|
| Water, purified | 82.93 |

This can be used in a pharmaceutical aerosol according to the following formula:

| | % w/w |
|---|---|
| Active ingredients | 2.0 |
| Emulsion base | 94.0–95.0 |
| Hydrocarbon propellant A-46 | 3.0–4.0 |

Several different steroids, antibiotics and other agents may be dispensed in this manner. Both hydrocarbons and compressed gas propellants may be used. Unless an exemption has been granted by the FDA, fluorocarbons are no longer used for these products. (Contraceptive foams have been exempted.)

The techniques used in preparing an aerosol emulsion are the same as those used for non-aerosol emulsions. Surfactants that showed some solubility in the propellants were most effective.

## *Nonaqueous Stable Foams*

Nonaqueous stable foams may be formulated through the use of various glycols such as polyethylene glycol, which may be formulated according to the following:

| | % w/w |
|---|---|
| Glycol | 91.0–92.5 |
| Emulsifying agent | 4.0 |
| Hydrocarbon propellant | 3.5–5.0 |

The emulsifying agents found most effective were from the class of glycol esters, for example, propylene glycol monostearate. Various medicinal agents can be incorporated into this base.

## *Quick-breaking Foams*

In this system, the propellant is in the external phase. When dispensed, the ******ebook converter DEMO Watermarks*******

product is emitted as foam, which then collapses into a liquid. This type of system is especially applicable to topical medication, which can be applied to limited or to large areas without the use of a mechanical force to dispense the active ingredients. Quick-breaking aerosol foams may be formulated starting with the following:

| | % w/w |
|---|---|
| Ethyl alcohol | 46.0–66.0 |
| Surfactant | 0.5–5.0 |
| Water | 28.0–42.0 |
| Hydrocarbon propellant | 3.0–15.0 |

The surfactant can be of the nonionic, anionic, or cationic type. It should be soluble in both alcohol and water. If the proportion of ingredients is varied, foams may be obtained having a wide range in stability.

A specific formulation for quick-breaking foam would consist of:

| | % w/w |
|---|---|
| Polawax | 2.0 |
| SDA 40 Anhyd. | 61.0 |
| Perfume | 1.5 |
| Menthol | 0.1 |
| Water, purified | 35.4 |

This is pressurized by mixing 90% concentrate and 10% propellant. The pressure should be below 25 psig and the product can be packaged in glass aerosol containers using a valve and foam actuator.

## *Thermal Foams*

These foams were developed several years ago and were used to produce a warm foam for shaving. They were not readily accepted by the consumer, however, and were soon discontinued, owing to inconvenience of use, expense, and lack of effectiveness. The same technology was used to dispense hair colors and dyes, but unfortunately, they were subject to some of

the same problems, as well as some corrosion problems, and were therefore unsuccessful. It has been reported that these systems would be advantageous in dispensing medicated foams in which the application of heat would be desirable. The technology is available and is fully discussed in the previous edition of this book.

## Selection of Components

### *Propellant*

Prior to 1978, fluorinated hydrocarbons were used almost exclusively as the propellants for all types of pharmaceutical aerosols. Their chemical inertness, lack of toxicity, lack of flammability and explosiveness, and their safe record of use made them ideal candidates for use. The publication of the "ozone depletion theory" in the mid-1970s, however, and the alleged implication of the fluorocarbons in depleting the ozone levels in the atmosphere, led to the phasing out and ban of the use of fluorocarbon propellants in aerosols (with few exceptions) in 1978. This ban, promulgated by the Environmental Protection Agency (EPA), Food and Drug Administration (FDA), and the Consumer Products Safety Commission, became fully effective in April 1979, when manufacturers could no longer ship aerosol products containing fluorocarbons unless the product carried a specific federal exemption.

While some propellant manufacturers indicated that there were other suitable replacements for propellants 11, 12 and 114, the only ones that have survived the necessary toxicity tests (long and short-range) are fluorocarbons 152a, 142B and 22, which may be of limited value. The other alternatives include hydrocarbons, compressed gases, and mechanical devices and pumps. Of these alternatives, however, hydrocarbons were restricted to use with foams and water-based aerosols, and compressed gases were of limited value in aqueous products where the propellant and water were not miscible. While compressed gases overcame the immiscibility of the components, other problems such as loss of propellant, and to a lesser degree, dispersion of the spray became apparent. Since compressed gas systems do not have a chilling effect, they are applicable to topical preparations.

With the development of newer valve technology (the vapour tap and the Aquasol valve), it was found that hydrocarbon propellants, such as butane, propane, isobutane, and their mixtures, could be safely used not only with aqueous products, but with solvent-based aerosols as well. At present, hydrocarbons can be used for all types of topical aerosols, and their flammability properties (as measured by the flame extension test) can be reduced to within safe limits as allowed by the Department of Transportation (DOT) and can meet hazard labeling requirements. As a result of this new technology, all non-exempted topical pharmaceutical aerosols have been 

satisfactorily reformulated using a hydrocarbon or blend of hydrocarbon as the propellant.

Inhalation aerosols for oral or nasal use have been exempted from the FDA ban, and the fluorinated hydrocarbons—namely, propellants 12, 12/114, and in some cases 12/11 are still used. Contraceptive foams are exempted and still utilize fluorocarbons as the propellant. Recently the FDA granted an exemption for a topical antibiotic aerosol. It should also be indicated that these new products will require a new drug application (NDA).

The compressed gases-nitrogen, nitrous oxide, and carbon dioxide—can be used but are of limited value. The pump system is used for liquid antiseptics, germicides, and nasal sprays.

## *Containers*

Both glass and metal containers have been used for pharmaceutical aerosols. Glass is preferred, but its use is limited, owing to its brittleness and the danger of breakage should the container accidentally be dropped. When the total pressure of the system is below 25 psig and there is not more than 15% propellant, glass can be safely used. Pressures up to 33 psig can be utilized in conjunction with a glass container having a double plastic outer coating.

Most nonaqueous products can be placed into an unlined tinplated metal container. Depending on the degree of protection desired, ${}^{X}A$, ${}^{X}A$, or 1 pound of tin per base box can be used. This type of container has been found to be satisfactory for many alcohol-based pharmaceuticals, e.g. spray on bandages.

Products having a low pH and containing water utilize organic linings of epoxy and/or vinyl resins. Although the vinyl resin forms a tough film, it is poorly resistant to steam and cannot be used for products that must be heat sterilized or filled hot (about 200°F). For this purpose, an epoxy resin can be used since it has a greater degree of heat stability. Both materials are essentially odorless and flavorless, although less odor and flavor are found in vinyl as compared with the epoxy material. A commonly used organic coating consists of an undercoat of vinyl and a top coat of epoxy resin. This has been used to best advantage for those aqueous products of low pH.

Those products containing soaps utilize similar liners, but special attention must be given to the solder that is used for the “side-seam” of some containers. A mixture of 2% tin and 98% lead is used when the pressure is

below 40 psig. When soaps are involved, lead reacts with the fatty acids present to form insoluble lead salts, which cause valve clogging. For this purpose, pure tin or combinations of tin, silver, and other metals are used, as well as containers with a welded side-seam. It is important that specifications for containers include these considerations, and that product stability versus composition of solder be evaluated. For this reason, the welded side-seam container is now preferred for use with pharmaceuticals.

## *Chemical Reactions*

Special attention should be given to those products containing alcohol and packaged in aluminum containers. Anhydrous ethanol is extremely corrosive to aluminum and reacts according to the following equation:

$$6Al + 6C_2H_5OH\ (Anhyd.) \rightarrow 2(C_2H_5O)_3Al + 3H_2$$

The hydrogen, which is slowly liberated, increases the pressure in the container. This increase in pressure, along with a general dissolving of some of the aluminum, may result in rupture of the container. This reaction can be reduced or prevented by anodizing the aluminum and adding water to formula. The presence of 2 to 3% water tends to inhibit the reaction. Nonpolar solvents appear to be relatively safe in aluminum containers. Polar solvents tend to be corrosive to bare aluminum, but the reaction can be controlled by the addition of water and/or other inhibitors.

## *Other Containers*

Creams and ointments may be dispensed from a pressurized container by the use of an aluminum container that has been fitted with a piston and nitrogen or hydrocarbon as the propellant. When the valve is opened, the pressure against the piston pushes the piston, causing the product to be dispensed. The viscosity of the finished product plays an important role in the satisfactory dispensing of the product.

Many attempts have been made to separate product from the propellant. One such system utilizes an “accordion-pleated” plastic bag as shown in Fig. 21.8. The product is placed inside the bag and propellant is injected into the outer container through the rubber-plugged hole located in the bottom of the container.


**Fig. 21.8:** Sepro container (*Courtesy*of Continental Can Company, Chicago, IL)

Another development in barrier packs uses a laminated film made into a flexible bag. A perforated dip tube is attached to the valve and functions to prevent the bag from collapsing as the contents are used. This system, termed "Powr Flo" (American Can Company), is useful for dispensing pharam-ceutical ointments and creams.

The Preval system not only separates the propellant from the product, but allows the product to be dispensed as a spray or powder. This system consists of an aluminum cartridge containing propellant 12 and an aerosol valve. A dip tube extends from the propellant chamber to the bottom of the container and allows for the flow of product when the valve is opened. Around the top of the valve housing, a vapour tap is placed, which extends into the propellant chamber. The product is added to the container (which need not be a pressure container) and then the valve with propellant cartridge is inserted. When the specially designed actuator is depressed, propellant vapour escapes from the propellant chamber. This creates a "venturi" effect, drawing up some of the product. At this point, mixing of propellant vapour and product takes place. The vapourized propellant then aids in carrying the product through the actuator, where it is dispersed into the desired spray. Varying ratios of propellant to product can be achieved, with results ranging from a fine to a coarse spray (Fig. 21.9).

A nonaerosol package that is self-evacuating has been developed by Plant Industries and is known as Selvac. This system utilizes a resilient bladder, which is filled with the product. As the product is filled into the bladder, the bladder stretches, thereby causing mechanical energy to squeeze out the contents as the valve is released. Since this system does not contain a gas, there is little internal pressure. Two bladders are used, assembled one inside the other. The outer bladder is usually made from natural rubber latex. The valve is inserted into the bladder, and this unit is then fitted into an outer nonpressurized container. The product is filled through the valve by means of a piston-type filler, which forces the product into the bladder.

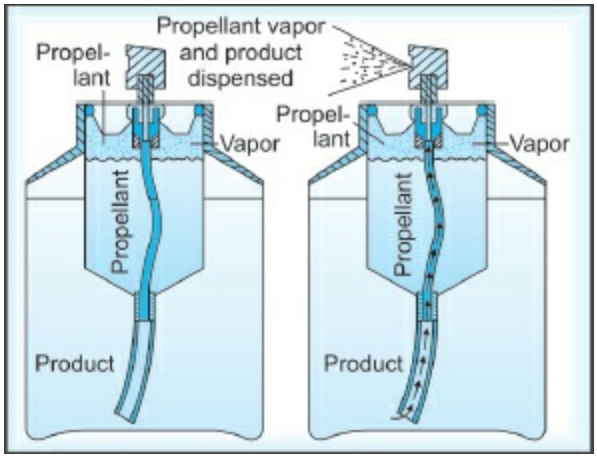



**Fig. 21.9:** Schematic view of Preval (*Courtesy* of Precision Valve Corporation, Yonkers, NY)

## *Valves, Actuators and Applicators*

Valves are selected on the basis of materials of construction and size of various orifices. Although it is extremely difficult to indicate the proper valve for each product, suggested valve designs for specific applications are available.

Various applicators have been specially designed for use with aerosol pharmaceuticals. Inhalation actuators must have all the characteristics of spray actuators and must allow escape of propellant vapours so that the vapours are not inhaled in appreciable amounts by the user. Throat applicators must be capable of depositing the medication directly into the throat area. Elongated tubes having small internal orifices, which permit a breakup of the spray, are generally used. Nasal actuators are designed to fit

into the nose and deliver the product as a fine mist. Other applicators have been designed for specific uses, including vaginal application, ophthalmic application, and others (*see* Figs 21.4 and 21.5).

## Manufacture of Pharmaceutical Aerosols

To prepare and package pharmaceutical aerosols successfully, special knowledge, skills and equipment are required. As with other pharmaceutical products, these operations must be carried out under strict supervision and adherence to rigid quality controls. Since part of the manufacturing operation (addition of propellant to concentrate) is carried out during the packaging operation, the quality control system must be modified to account for this difference. In addition to the equipment used for the compounding of liquids, suspensions, emulsions, creams, and ointments, specialized equipment capable of handling and packaging materials at relatively low temperatures (about −40°F) or under high pressure must be available. This equipment is usually limited to aerosol or pressurized packaging, and in most instances, cannot be used for other pharmaceutical operations.

### *Pressure Filling Apparatus*

Pressure filling apparatus consists of a pressure burette capable of metering small volumes of liquefied gas under pressure into an aerosol container. The propellant is added through the inlet valve located at the bottom or top of the burette. Trapped air is allowed to escape through the upper valve. The desired amount of propellant is allowed to flow through the aerosol valve into the container under its own vapour pressure. When the pressure is equalized between the burette and the container (this happens with low-pressure propellants), the propellant stops flowing. To aid in adding additional propellant, a hose leading to a cylinder of nitrogen or compressed air is attached to the upper valve and the added nitrogen pressure causes the propellant to flow. Another pressure filling device makes use of a piston arrangement so that a positive pressure is always maintained. Fig. 21.10 illustrates typical laboratory pressure filling equipment. This equipment cannot be used to fill inhalation aerosols fitted with a metered valve. Pressure filling equipment that fills through “pressure-fillable” metered valves is available.

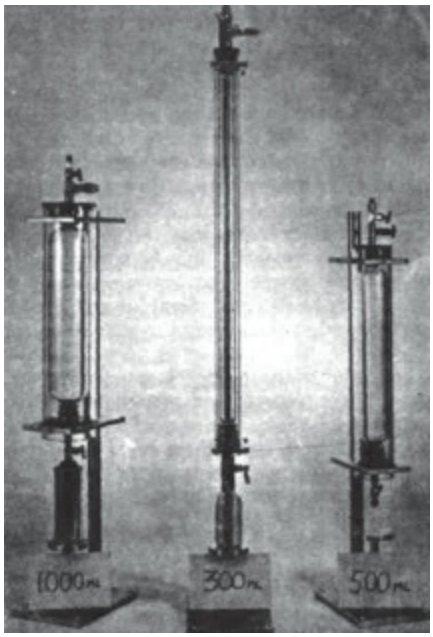



**Fig. 21.10:** Pressure burets for laboratory filling of aerosols (*Courtesy* of Aerosol Laboratory Equipment Corporation, Walton, NY)

## *Cold Filling Apparatus*

Cold filling apparatus is somewhat simpler than the pressure filling apparatus. All that is needed is an insulated box fitted with copper tubing that has been coiled to increase the area exposed to cooling. Fig. 21.11 illustrates such a unit, which must be filled with dry ice/acetone prior to use. This system can be used with metered valves as well as with non-metered valves; however, it should not be used to fill hydrocarbon aerosols since an excessive amount of propellant escaping and vapourizing may form an explosive mixture at the floor level (or lowest level). Fluorocarbon vapours, although also heavier than air, do not form explosive or flammable mixtures.

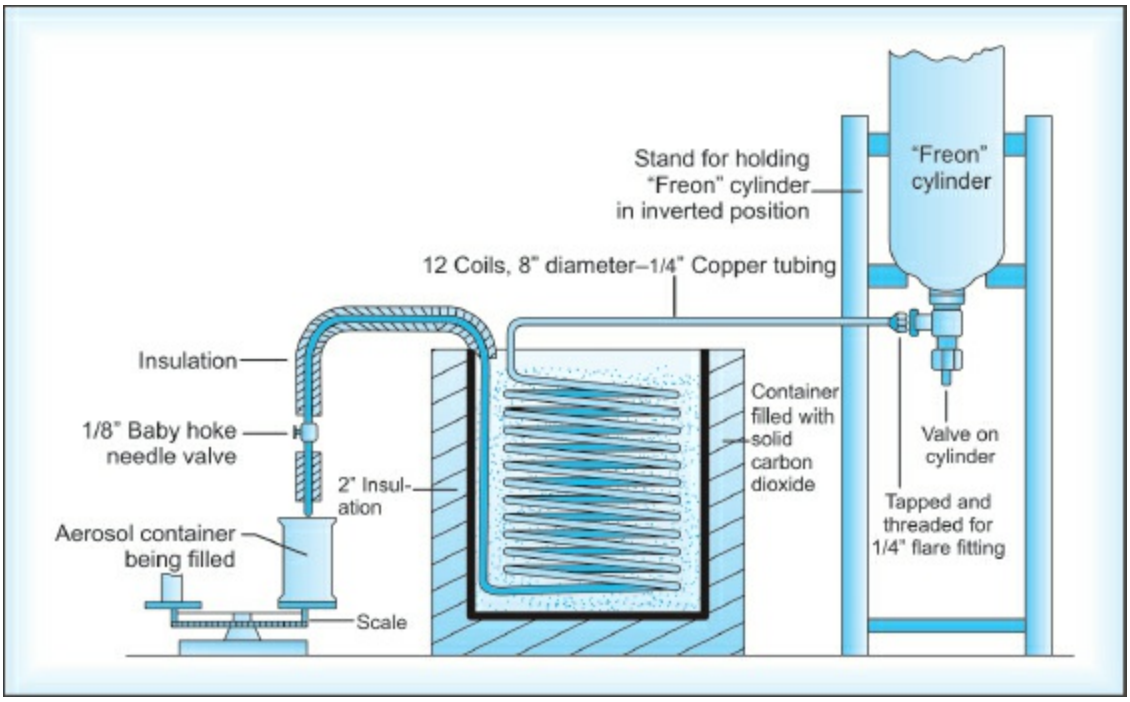



**Fig. 21.11:** Apparatus for cold filling process (*Courtesy* of EI. duPont de Nemours and Co., Inc., Wilmington, DE)

## *Compressed Gas Filling Apparatus*

Compressed gases can be handled easily in the laboratory without the use of elaborate equipment. Since the compressed gases are under high pressure, a pressure-reducing valve is required. Attached to the delivery gauge is a flexible hose capable of withstanding about 150 pounds per square inch gauge pressure and fitted with a filling head. More elaborate units utilize a flow indicator between the gauge and the flexible hose. To use this equipment for filling aerosols with compressed gases, the concentrate is placed in the container, the valve is crimped in place, and the air is evacuated by means of a vacuum pump. The filling head is inserted into the valve opening, the valve is depressed, and the gas is allowed to flow into the container. When the pressure within the container is equal to the delivery pressure, the gas stops flowing. For those products requiring an increased amount of gas, or for those in which solubility of the gas in the product is necessary, carbon dioxide and nitrous oxide can be used. To obtain maximum solubility of the gas in the product, the container is shaken manually during and after the filling operation. Mechanical shakers are also available for this purpose.

## LARGE-SCALE EQUIPMENT

Good pharmaceutical manufacturing practice requires that the filling of pharmaceutical aerosols be conducted under conditions that ensure freedom from contamination. Only equipment used specifically for aerosols is further discussed in this chapter.

**Concentrate filler:** This can range from a single-stage single hopper to a large straightline multiple-head filler or a rotary type multiple-head filler. Production schedules dictate the type of filler required. Most of these fillers deliver a constant volume of product and they can be set to give a complete fill in one or more operations. Usually, only part of the product is added at each stage, assuring a more accurate fill.

**Valve placer:** The valve can be placed over the container either manually or automatically. High-speed equipment utilizes the automatic valve placer. This orients the valve and places it in position prior to the crimping operation.

**Purger and vacuum crimper:** Aerosols are packaged in both metallic and glass containers, each requiring their own style of crimper. Combination can and bottle cappers can be used for most laboratory procedures and operate manually or on air pressure (about 80 pounds per square inch). These are capable of producing more than 10 to 12 cans per minute.

Most crimpers serve a dual function, that is, to evacuate the air within the container to about 24 inches of mercury and then seal the valve in place. Single-head crimpers or multiple-head rotary units capable of vacuum crimping up to 120 cans per minute are available. These usually require both air pressure (90 to 120 pounds per square inch) and vacuum.

**Pressure filler:** These units are capable of adding the propellant either through the valve stem, body, and dip tube, around the outside of the stem, or under the valve cup before crimping. They are either single or multiple-stage units arranged in a straight line or as a rotary unit. To speed production, a positive pressure is used to force the liquid propellant into the container.

Evacuation of air from the container, crimping the valve, and addition of the propellant can be achieved in basically one operation through the use of an “under the cap” filler. This unit operates as follows. A seal is made by lowering the crimping bell onto the container, air is removed by vacuum, and

propellant is then metered into the container at room temperature and high pressure. The crimping collet expands and crimps the valve into the opening. This unit can be fitted with three to nine filling heads.

**Leak test tank:** This consists of a large tank filled with water and containing heating units and a magnetized chain so that the cans or pucks for glass, aluminum and plastic containers are carried through and submerged into the water. The length of the tank is such that the temperature of the product before it emerges from the tank is 130°F.

According to DOT regulations, “ each completed container filled for shipment must have been heated until contents reached a minimum of 130°F, or attained the pressure it would exert at this temperature, without evidence of leaking, distortion, or other defects”.

## MANUFACTURING PROCEDURE

In general, the manufacture of aerosol products takes place in two stages: Manufacture of concentrate and addition of propellant. For this reason, part of the manufacturing operation takes place during the filling operation, which is quite different from nonaerosol pharmaceutical products. This necessitates special quality control measures during the filling operation to ensure that both concentrate and propellant are brought together in the proper proportion.

The aerosol concentrate is prepared according to generally accepted procedures, and a sample is tested. Testing at this point can save both time and money should the concentrate prove to be unacceptable. Once the propellant is added and the product is sealed into a container with a valve, complete rejects must be discarded, obviously a more costly solution. Early detection prevents the loss of the other components. This would also have made it possible to correct the rejected batch instead of discarding it. In many instances, this can be accomplished by making adjustments to the concentrate prior to aerosol filling.

Two methods have been developed for the filling of aerosol products. The *cold filling method* requires the chilling of all components, including concentrate and propellant, to temperatures of −30 or −40°F, whereas the *pressure filling method* is carried out at room temperature utilizing pressure equipment. The type of product and size of container usually influence the method to be used.

The cold filling method is restricted to nonaqueous products and to those products not adversely affected by low temperatures in the range of −40°F. In this method, product concentrate is chilled to −40°F and added to the chilled container. The chilled propellant is then added in one or two stages, depending on the amount. An alternating method of cold filling is to chill both concentrate and propellant in a pressure vessel to −40°F and then add their mixture to the aerosol container. A valve is then crimped in place. The container passes through a heated water bath in which the contents of the container are heated to 130° F to test for leaks and strength of container. The container is air-dried, spray-tested if necessary, capped, and labeled. (Containers may be lithographed, and as a consequence, the latter step is omitted.) This filling method is no longer used to any great extent and has been replaced by the pressure filling process. Metered-dose aerosols can be

filled by either process.

The pressure filling method, when first developed, was generally slower than the cold filling method. With the development of newer techniques, the speed of this method has been greatly increased to make it comparable in rate of production to the cold filling method. The concentrate is added to the container at room temperature, and the valve is crimped in place. The propellant is added through the valve or “under the cap”. Since the valve contains extremely small openings (0.018 inch to 0.030 inch), this step is slow and limits production. With the development of rotary filling machines and newer filling heads, which allow propellant to be added around and through the valvestem, the speed has been increased. For those products adversely affected by the air that may be trapped within the container, the air in the headspace is evacuated prior to adding the propellant. Following the addition of the propellant, the method becomes similar to the cold filling method.

For the most part, the pressure method is also preferred because some solutions, emulsions, suspensions and other preparations cannot be chilled. Various factors determine the method to be used. The pressure method is usually preferred to the cold method, because there is less danger of contamination of the product with moisture; high production speeds can be achieved; less propellant is lost; and the method is not limited, except for certain types of metering valves that can only be handled by the cold filling process or through use of an “under the cap” filler and valve crimper. Some metered valves that are pressure-fillable are now available.

Following the development of the aerosol product, an initial production of about 500 to 1000 units is scheduled. The initial fill is made according to the specifications of the pharmaceutical concern. These units are used for additional stability studies, and for the determination of incompatibilities with various components (containers, valves, gaskets, dip tubes). This run is also used to determine some of the problems that may become apparent in developing the product from the laboratory to full production.

A larger run of 10,000 to 25,000 units is scheduled next. At this time, all materials are identical to those utilized for the production run, and the equipment used must be the same as the production equipment. These samples can be used for clinical studies and further testing if necessary. This test run should give the following information (1) suitability of scale-up

operation, (2) number of rejects to be expected (valve, container, and other components), (3) limitations of filling process (tolerances for filling, crimping, and other operations), (4) determination of equipment to be used and (5) check on effectiveness and acceptability of final product. Should satisfactory results be obtained at this point, arrangements for full-scale production can be made.

A typical cold filling aerosol line contains the following units arranged in the order given: Unscrambler, air cleaner, concentrate filler (capable of being chilled), propellant filler (also capable of being chilled), valve placer, valve crimper, water bath, labeler, coder, and packing table. The comparable pressure filling line would be identical in arrangement except that (1) no refrigeration for chilling is required, (2) valve placer is located after “concentrate filler” and a purger and vacuum crimper are added and (3) this equipment is followed by a pressure filler.

Where “under-the-cap” filling is used, the purger, vacuum crimper and pressure filler are replaced with a single unit.

## Quality Control for Pharmaceutical Aerosols

Basically, there is no difference between methods used to produce pharmaceutical aerosols and those used to produce nonpharmaceutical aerosols, but there are differences in the standards and specifications for their production.

### *Propellants*

Propellants used in medicinal and pharmaceutical aerosols require special handling, and in many instances, special test procedures. All propellants are shipped to the user with accompanying specification sheets; however, before the propellant is used (in fact, before it is even piped into a storage tank), it is subjected to the same rigid tests necessary for all other raw materials. A sample is removed and sent to the laboratory, where its vapour pressure is determined and compared to specifications. When necessary, the density is also determined, and this is used as a further check. Gas chromatography is used to determine the identity of the propellant, and when a blend of propellants is used, to determine the composition. The purity and acceptability of the propellants is tested by moisture, halogen, and nonvolatile residue determinations. Depending on the end use of the propellants, several of these tests may be more important than others. All suppliers of propellants utilize the aforementioned tests in their own laboratories, and the tests that are run by the user are generally a check on these results, and more important, they ensure that the propellants have not become contaminated during shipment. Monographs for propellants 11, 12, and 114 are included in USP XX/NF XV (1980); monographs for the hydrocarbons are currently being written.

### *Valves, Actuators and Dip Tubes*

These parts are subjected to both physical and chemical inspection. The problem is more complex than with nonaerosol components since a valve is a multicomponent assembly consisting of various parts made to close tolerances. The examination at this point must determine whether the valves are fit to be used. They are sampled according to standard procedures as found in Military Standard Mil-STD-105D. One manufacturer of aerosols for this purpose actually assembles valves, using component parts having similar tolerances to ensure that parts having the minimum tolerance do not engage

with parts approaching maximum tolerance.

To provide the means for determining the acceptance of metered-dose aerosol valves for pharmaceutical use, a suitable test method was developed by the Aerosol Specifications Committee, Industrial Pharmaceutical Technology Section, Academy of Pharmaceutical Sciences. The object of this test is to determine the magnitude of the valve delivery and the degree of uniformity between the individual valves as related to the acceptance of any given lot of metered aerosol valves. The test is not designed to determine the suitability or the lack of suitability of the valves for a specific formulation and/or application. Detailed specifications for metered aerosol valves is a matter to be resolved between the pharmaceutical manufacturers and the aerosol valve suppliers.

The following three test solutions were proposed to rule out variations in valve delivery brought about by different formulations. These solutions were selected since they represent the range of propellants and propellant concentrations most often used in pharmaceutical aerosols. Since a metered valve delivers a specific volume of liquid with each actuation, it was proposed that metered valve delivery be designated in terms of valve delivery —volume expressed in microliters. In such a case, the test solutions recommended would apply to the control of valve delivery and uniformity for a great variety of formulations of different densities.

The test solutions may be prepared in bulk and stored in hermetically sealed containers with suitable fitments for transferring the test solution into the test units. The transfer of the test solution should be made in such a manner that no change occurs in the proportions of the ingredients of the test solution.

| **Test solution A** | |
|---|---|
| | **% w/w** |
| Isopropyl myristate | 0.10 |
| Dichlorodifluoromethane | 49.95 |
| Dichlorotetrafluoroethane | 49.95 |

Specific gravity at 25°C = 1.384

| Test solution B | |
|---|---|
| | % w/w |
| Isopropyl myristate | 0.10 |
| Alcohol USP | 49.9 |
| Dichlorodifluoromethane | 25.0 |
| Dichlorotetrafluoroethane | 25.0 |

Specific gravity at 25°C = 1.092

| Test solution C | |
|---|---|
| | % w/w |
| Isopropyl myristate | 0.10 |
| Trichloromonofluoromethane | 24.9 |
| Dichlorodifluoromethane | 50.25 |
| Dichlorotetrafluoroethane | 24.75 |

Specific gravity at 25°C = 1.388

**Testing procedure:** A representative sampling of the valves from each shipment is made according to existing methods of sampling. Twenty-five valves are selected and placed onto suitable containers, into which has been placed the specified test solution. Where possible, the containers may be filled by the pressure process. A button-type actuator with a 0.021-inch or larger unrestricted orifice is attached. This button remains in place throughout the test procedure. The containers are placed in a suitable atmosphere at a temperature of 25 ± 1°C. When the product has attained this temperature, the valve should be actuated to the fullest extent for at least 2 sec following complete dispensing of a single delivery. This procedure is repeated for a total of ten times.

The test unit is weighed to the nearest milligram. The valve is actuated to the fullest extent for at least 2 sec following complete dispensing of a single delivery. The test unit is reweighed, and the difference between it and the previous weight represents the delivery in milligrams. The test procedure is repeated for a total of two individual deliveries from each of the twenty-five

test units. The individual delivery weights in milligrams are divided by the specific gravity of the test solution to obtain the valve delivery per actuation in microliters.

**Valve acceptance:** The test procedure applies to two categories of metered aerosol valves having the following limits. For valves delivering:

54 µl or less, the limits are ± 15%.

55 to 200 µl, the limits are ± 10%.

1. Of the 50 individual deliveries, if four or more are outside the limits for the specified valve delivery, the valves are rejected.
2. If three individual deliveries are outside the limits, another twenty-five valves are sampled and the test is repeated. The lot is rejected if more than one delivery is outside the specifications.
3. If two deliveries from one valve are beyond the limits, another twenty-five valves should be taken. The lot is accepted if not more than one delivery is outside the specifications.

## *Containers*

Containers are sampled according to standard sampling procedures and in a manner similar to valves. Both uncoated and coated metal containers must be examined for defects in the lining. Several quality control aspects include specifications for the degree of conductivity of an electric current as a measure of the exposed metal. Glass containers must be examined for flaws. The dimensions of the neck and other parts must be checked to determine conformity to specifications. The weight of the bottle also should be determined.

## *Weight Checking*

This is usually accomplished by periodically adding to the filling line tared empty aerosol containers, which after being filled with concentrate, are removed and then accurately weighed. The same procedure is used to check the weight of the propellant that is being added. When a propellant blend is being utilized, checks must be made to ensure a proper blend of propellants. As a further check, the finished container is weighed to check the accuracy of the filling operation.

## *Leak Testing*

A means of checking the crimping of the valve must be available to prevent defective containers due to leakage. For metal containers, this is accomplished by measuring the “crimp” dimensions and ensuring that they meet specifications.

Final testing of the efficiency of the valve closure is accomplished by passing the filled containers through the water bath. Periodic checks are made of the temperature of the water bath, and these results are recorded.

## *Spray Testing*

Many pharmaceutical aerosols are 100% spray tested. This serves to clear the dip tube of pure propellant (for products filled by pressure through the stem, body, and dip tube), to clear the dip tube of pure concentrate (for products filled by pressure under the cap or around the stem) and to check for defects in the valve and the spray pattern. For metered valves, it serves to prime the valve so that it is ready for use by the consumer. Several of the basic aspects of a quality control system have been included in this section.

## Testing of Pharmaceutical Aerosols

Aerosols are “pressurized packages” and many tests are necessary to ensure proper performance of the package and safety during use and storage. All aerosol products that are shipped in interstate commerce are subject to the regulations of the DOT. These regulations impose limitations on the pressure within the container, flash points, flame extension, and flammability. The provisions of the Hazardous Substances Labeling Act and the Food, Drug and Cosmetic Act must be applied. In addition to these federal regulations, many local officials impose further restrictions upon aerosols.

Pharmaceutical aerosols can be evaluated by a series of physical, chemical, and biologic tests, including:

A. Flammability and combustibility
    1. Flash point
    2. Flame extension, including flashback
B. Physicochemical characteristics
    1. Vapour pressure
    2. Density
    3. Moisture content
    4. Identification of propellant(s)
    5. Concentrate-propellant ratio
C. Performance
    1. Aerosol valve discharge rate
    2. Spray pattern
    3. Dosage with metered valves
    4. Net contents
    5. Foam stability
    6. Particle size determination
    7. Leakage
D. Biologic characteristics

The flammability and combustibility of aerosol pharmaceuticals may be determined by the following procedures.

## *Flash Point*

This is determined by use of the standard Tag Open Cup Apparatus. The aerosol product is chilled to a temperature of about 25°F and transferred to the test apparatus. The test liquid is allowed to increase slowly in temperature and the temperature at which the vapours ignite is taken as the flash point. Although the test is still used, the results are of limited value because the flash point obtained is usually the flash point of the most flammable component, which in the case of topical pharmaceuticals is the hydrocarbon propellant.

## *Flame Projection*

This test indicates the effect of an aerosol formulation on the extension of an open flame. The product is sprayed for about 4 sec into a flame. Depending on the nature of the formulation, the flame is extended, the exact length being measured with a ruler.

## *Vapour Pressure*

The pressure can be measured simply with a pressure gauge or elaborately through use of a water bath, test gauges and special equipment. It is important that the pressure variation from container to container be determined, since excessive variation indicates the presence of air in the headspace. A can puncturing device is available for accurately measuring vapour pressure. Methods are available for aerosols packaged in both metal and glass containers.

## *Density*

The density of an aerosol system may be accurately determined through the use of a hydrometer or a pycnometer. These methods, which have been used for the density of nonaerosols, have been modified to accommodate liquefied gas preparations. A pressure tube is fitted with metal flanges and a Hoke valve, which allow for the introduction of liquids under pressure. The hydrometer is placed into the glass pressure tube. Sufficient sample is introduced through the valve to cause the hydrometer to rise halfway up the length of the tube. The density can be read directly. Specific gravity can be determined through the use of a high-pressure cylinder of about 500 ml capacity.

## *Moisture Content*

Many methods have proven useful for this purpose. The Karl Fischer method has been accepted to a great extent. Gas chromatography has also been used.

## *Identification of Propellants*

Gas chromatography and infrared spectrophotometry have been used to identify the propellants and also to indicate the proportion of each component in a blend.

## *Aerosol Valve Discharge Rate*

This is determined by taking an aerosol product of known weight and discharging the contents for a given period of time using standard apparatus. By reweighing the container after the time limit has expired, the change in weight per time dispensed is the discharge rate, which can then be expressed as grams per second.

## *Spray Patterns*

A method for comparing spray patterns obtained from different batches of material or through the use of different valves is available and is shown in Fig. 21.12. The method is based on the impingement of the spray on a piece of paper that has been treated with a dye-talc mixture. Depending on the nature of the aerosol, an oil-soluble or water-soluble dye is used. The particles that strike the paper cause the dye to go into solution and to be absorbed onto the paper. This gives a record of the spray, which can then be used for comparison purposes. To control the amount of material coming into contact with the paper, the paper is attached to a rotating disk that has an adjustable slit.

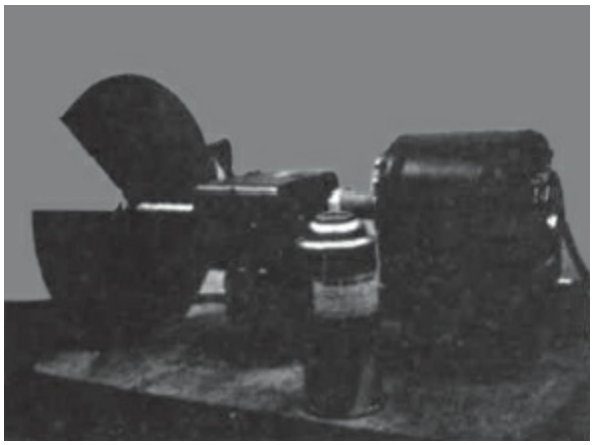


**Fig. 21.12:** Apparatus for the determination of spray patterns

## *Dosage with Metered Valves*

Several points must be considered (1) reproducibility of dosage each time the valve is depressed and (2) amount of medication actually received by the patient. Reproducibility of dosage may be determined by assay techniques whereby one or two doses are dispensed into a solvent or onto a material that absorbs the active ingredients. These solutions can then be assayed, and the amount of active ingredients determined. Another method that can be used involves accurate weighing of filled container followed by dispensing of several doses. The container can then be reweighed and the difference in weight divided by the number of doses dispensed gives the average dose. This must then be repeated and the results compared. Determination of the dose received by a patient is a rather difficult procedure, since all of the material dispensed is not carried to the respiratory tract. An artificial respiratory system has been developed and is satisfactory for this purpose.

## *Net Contents*

Several methods can be used to determine whether sufficient product has been placed into each container. The tared cans that have been placed onto the filling line are reweighed, and the difference in weight is equal to the net contents. The other method is a destructive method and consists of weighing a full container and then dispensing the contents. The contents are then weighed, with provision being made for the amount retained in the container. Other modifications consist of opening the container and removing as much of the product as possible. These tests are not indicated in determining the

actual net weight of each container as related to the amount that can actually be dispensed. The National Bureau of Standards has issued a method that can be used for foam type, low-viscosity, high-viscosity, and food aerosols. These methods standardize the manner in which the containers are to be dispensed.

## *Foam Stability*

Various methods have been suggested for the determination of foam stability. The life of a foam can range from a few seconds (for some quickbreakmg foams) to one hour or more depending on the formulation. Richman and Shangraw have indicated some of the factors involved in controlling the stability of a foam. Several methods have been used, which include a visual evaluation, time for a given mass to penetrate the foam, time for a given rod that is inserted into the foam to fall and the use of rotational viscometers.

## *Particle Size Determination*

Many methods have been advanced for the measurement of particle size of aerosols. Among, those that have been used to a great extent are the Cascade impactor and “light scatter decay” methods. The Cascade impactor operates on the principle that in a stream of particles projected through a series of nozzles and glass slides at high velocity, the larger particles become impacted first on the lower velocity stages, and the smaller particles pass on and are collected at higher velocity stages. Fig. 21.13 illustrates a unit suitable for the analysis of particles whose diameters range from 0.1 to 30 microns. This unit is specific for sampling aerosols comprised of particles that might be retained in the respiratory tract. Fig. 21.14 shows the efficiency of this unit for determining particle size. Various modifications have been made to improve its efficiency. Porush, Thiel, and Young used the light scatter decay method for determination of particle size in epinephrine aerosols. As the aerosol settles under turbulent conditions, the change in light intensity of a Tyndall beam is measured. They noted that the mass median diameter of an epinephrine aerosol ranged from 2.7 to 3.5 microns; that 70 to 78% of the particles were less than 5 microns, that 88 to 93% were less than 7 microns and that 98 to 100% were less than 10 microns. Sciarra and Cutie developed a method for the evaluation of different actuators based on the particle size distribution obtained.

**Fig. 21.13:** The cascade impactor for determining the particle size distribution of aerosols (*Courtesy* of Battelle Laboratories, Inc., Columbus, OH)

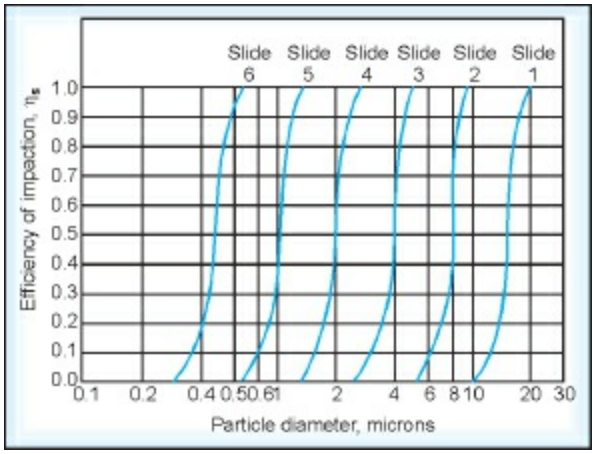



**Fig. 21.14:** Impact efficiency of cascade impactor (model Cl-S-6)

Other test methods have been covered by Johnsen, Dorland, and Dorland. In addition, a test procedure for leak testing, delivery rate, and pressure testing has been included in the USP XX/NF XV. In several cases, specific test procedures are indicated in the monographs for the aerosol preparation.

## *Biologic Testing*

The final phase involved in a comprehensive research and development

program for pharmaceutical aerosols must involve biologic testing. A limited number of these tests have been used to evaluate the efficiency of many products, including various antibacterial agents. These tests are similar to tests performed on nonaerosol pharmaceuticals. Biologic testing of aerosol products should include a consideration of therapeutic efficacy and toxicity.

## *Therapeutic Activity*

Various testing procedures are available to determine the therapeutic activity of aerosols. With the exception of consideration given to the aerosol feature of the package, these procedures are similar to existing tests used for nonaerosols. The dosage of the product has to be determined for inhalation aerosols, and this must be related to particle size distribution. Topical preparations are applied to the test areas in the usual manner, and adsorption of therapeutic ingredients can be determined.

## *Toxicity*

Toxicity testing should include both topical and inhalation effects. Aerosols applied topically may be irritating to the affected area and/or may cause a chilling effect. The degree of chilling effect depends on the type and amount of propellant present. There is really no good test available at the present time, although thermistor probes attached to recording thermometers have been used to indicate the change in skin temperature when skin is sprayed with an aerosol for a given period of time.

Inhalation toxicity must also be considered even though the product may be intended for topical administration. This can be accomplished by exposing test animals to vapours sprayed from an aerosol container.

## Stability Testing

One of the most important considerations in the formulation of a pharmaceutical aerosol is its stability. Those aspects directly concerned with the components used to prepare the pharmaceutical aerosol must be fully studied. The effect that the container has upon the product and conversely, the effect of product upon container, must be considered.

The same considerations apply to the valve components. Even slight changes in the various components of the valve may result in an inoperative package. The valve component of the aerosol package has several working parts made of different materials, such as natural or synthetic rubber, plastic, and stainless steel. All these materials may produce an adverse effect on the product and must be fully studied.

Since a variety of different materials are used in the make-up of the container, valve, and dip tube, it is difficult to determine whether a reaction takes place between the materials and the drug. Many of the components come into intimate contact with the medicinal agent. To determine whether these reactions do occur, all materials must be studied separately and collectively. Several container coatings as well as valves with different subcomponents may be studied so that any reaction between the component and the product may be detected. Samples are prepared and packaged in glass aerosol containers as controls.

The testing of these aerosols must cover three areas (1) concentrate and propellant, (2) container and (3) valve. Evidence of decomposition or deterioration in any of these areas could result in an ineffective product. A more detailed discussion of the testing of aerosols can be found elsewhere in the literature.

### *Concentrate and Propellant*

Immediately after preparation, several of the important physicochemical constants of the product are determined. These vary, depending on the nature of the product, but may well include vapour pressure, spray rate of valve, pH, density or specific gravity, refractive index, viscosity, total weight, assay of active ingredients, infrared and/or gas chromatography curves, color, and odor. These are then used for comparison during each evaluation of the product.

These samples are usually stored on their sides so that the product comes into contact with both the valve mounting cup and the container. When three-piece metal cans are used, care should be taken to ensure that some samples have liquid as well as gaseous contact with the side-seams (soldered or welded).

## *Container*

The contents of the container are removed by chilling the contents to a temperature of 0°F or less and opening the container. The container is then examined for signs of corrosion. These changes can be detected without much difficulty, since attack upon tinplate, tin-free steel, or aluminum is generally visible to the naked eye and under a microscope. Small pinholes can easily be detected. For those containers that have internal lacquering, the examination must ensure that the lacquer is not softened, dissolved, peeled, or blistered by the concentrate. Special attention should be paid to the side-seam and headspace, as there is a greater danger of attack upon these areas.

When glass is used for the container, an examination of the container can be omitted. Plastic containers may require special testing to determine whether leaching or sorption has taken place.

## *Valve*

The valve should be examined to ensure that it is functional and will satisfactorily dispense the product and be easily closed. This can readily be determined during the dispensing of the product. The valve cup should be examined for evidence of corrosion. The various valve subcomponents should also be examined for evidence of softening, cracking, elongation, or distortion. Several of these effects can result in defective valves that will not operate properly. Elongation and cracking of the dip tube should be noted, and if present, corrected.

## AEROSOLS FOR PULMONARY DRUG DELIVERY

The two possible mechanisms for delivery of drugs to the lung include an aerosol or direct instillation. The most commonly used is the aerosol, which consists of finely divided liquid droplets or solid particles in a gaseous suspension. A device which generates an aerosol is termed as an atomizer. An atomizer may be electrically, pneumatically or mechanically powered. The main types of device used at present to produce aerosols are nebulizers, metered dose inhalers (MDI) and dry powder inhalers (DPI), although development of the technology is causing the distinctions between these devices to become blurred.

## Nebulizers

Pharmaceutical nebulizers can be divided into two main groups, pneumatic and electric. An electric generator (ultrasonic) operates from an electric source, while a pneumatic generator (hydrodynamic and jet) derives its power from pressurized gas source.

The ultrasonic nebulizer consists of a piezoelectric crystal which produces high frequency sound waves in the liquid in the nebulizing unit. The surface waves produce small droplets which are conducted away by an airstream for inhalation. The hydrodynamic nebulizer uses a system that prepares a film of water for aerosol formation by flowing it over a hollow sphere. A small orifice in the sphere expels gas at supersonic velocity. This high-velocity gas ruptures the thin film of water and produces a continuous dispersion of fine, liquid particles. A gas cylinder or compressor supplies the gas pressure. In case of jet nebulizer, high-velocity gas flow is directed over a tube that is immersed in a water reservoir (Fig. 21.15). The expansion of the driver gas decreases the pressure over the tube, which draws the formulation into the gas stream. The high shear rate in the jet stream then nebulizes it.

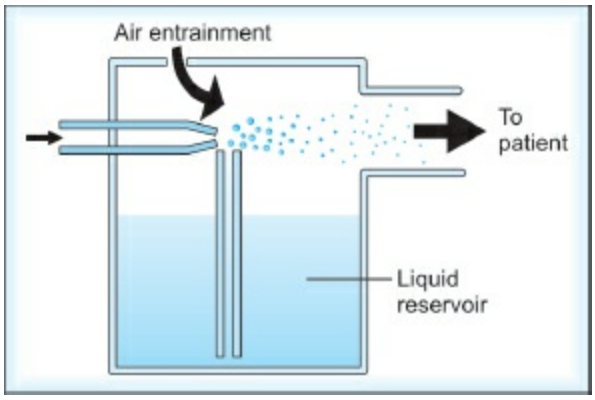



**Fig. 21.15:** An air-driven nebulizer

The properties of nebulizers vary widely; while all produce droplets with sizes in the range 1–10 μm, they vary significantly in droplet size distribution and pulmonary deposition. Consequently most use some sort of baffle system in the airstream; coarse droplets impact on this and are returned to the reservoir for re-nebulization, while the smaller particles avoid the baffle and are passed to the patient.

Despite this they have a number of advantages that is causing a renewal

of interest in their use. Continuous nebulization can deliver very large quantities of drugs if necessary, without major formulation problems. Moreover, MDIs and DPIs have a relatively high gas flow rate resulting in high oropharyngeal impaction. This problem is reduced in nebulizers since the airflow can be adjusted to suit the patient's inhalation rate. Unfortunately most nebulizers are bulky and require a fixed power source, which limits their use severely.

Boehringer Ingelheim have developed the Respimat®, spring-driven spray with a similar outward appearance to a conventional MDI. Unlike an MDI, the Respimat® delivers its spray in a slow low-velocity cloud leading to increased central pulmonary deposition.

## Metered Dose Inhalers

Pressurized metered-dose inhalers for delivery of medications have been available since the mid 1950s. In these systems, the drug is usually a polar solid which has been dissolved or suspended in a non-polar liquefied propellant. If the preparation is a suspension, as is most commonly the case, the powder is normally micronised by fluid energy milling and the suspension is stabilized by the addition of a surfactant. Lecithin, oleic acid and the Span and Tween series surfactants have been widely used for this type of formulation. Metered dose inhalers (MDIs) are the most commonly used drug delivery system for inhalation (Fig. 21.16). The propellants have a high vapour pressure of around 400 kPa at room temperature, but since the device is sealed, only a small fraction of the propellant exists as a gas. The canister consists of a metering valve crimped on to an aluminum can. Individual doses are measured volumetrically by a metering chamber within the valve.

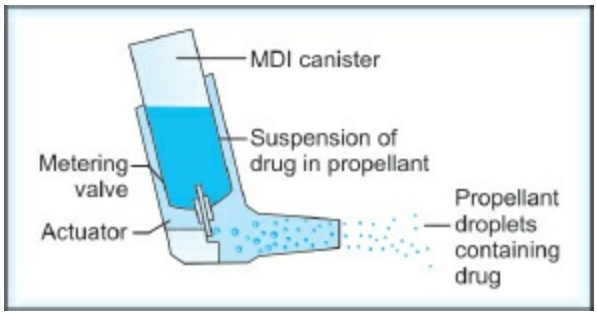



**Fig. 21.16:** The metered-dose inhaler

Each MDI canister can hold between 100 and 200 doses of between 20 pg and 5 mg of drug, which is released within the first 0.1 s after actuation. The valve stem is fitted into an actuator incorporating a mouthpiece. The aerosol, consisting of propellant droplets containing drug, is delivered from the actuator mouthpiece at very high velocity, probably about 30 $ms^{-1}$. There is partial (15–20%) evapouration of propellant prior to exit from the atomizing nozzle (“flashing”) and further break up of droplets beyond this point caused by the violent evapouration of the propellant. This results in a wide droplet size distribution from 1 to 5 μm. Only 10% of the particulates delivered in a single dose released by a metered-dose inhaler actually reach the lungs, since the bulk of it impacts in the oropharynx and the mouthpiece. Reduction of the

plume velocity, for example in the Gentlehaler® device causes a significant reduction in oropharyngeal deposition.

In order to be effective, metered dose aerosols must be triggered as the patient is inhaling. Some patients have difficulty with this feat of coordination, and breath actuated inhalers such as the Autohaler® have been designed to overcome this by triggering the valve as the patient breathes in. The Mist-Assist® inspiratory flow control device (IFCD, Ballard Medical, Draper, UT) is a compact device (similar in size to a spacer) through which both an MDI or medication from a nebulizer can be administered. By use of a floating ball within the inspiratory chamber, it provides visual and auditory (clicking sound) feedback to optimize timing of medication delivery and rate of inspiratory flow. Most important, the inspiratory flow rate (and therefore inspiratory resistance) can be adjusted on the device. This inspiratory flow control enhances laminar flow of particles and gas and increases the lung deposition.

There has been an increased interest in modifying metered-dose inhalers (MDIs) to minimize the number of administration errors and to improve the drug delivery of aerosolized particles into the nasal passageways and respiratory airways. Some of these modifications have included the introduction of tube spacers, breath actuators, and portable plastic reservoirs with inhalation aerosols. In the case of intranasal preparations, new propellant-free metered pumps have been introduced to replace the traditional propellant delivery systems.

During the late 70s and early 80s, there were a number of in vivo and in vitro studies evaluating the differences between the conventional adaptors and the expanded-chamber adaptors (Fig. 21.17), referred to as “spacers,” or “tube spacers”. At present, many conventional short-stem MDIs deliver at best only 10 to 15% of the dose actuated into the respiratory airways. The balance of the dose is either lost to the inner surface of the adaptor (approximately 10%), or is deposited through inertial impaction in the oropharynx area (80%). The latter leads to swallowing and possible systemic absorption of the therapeutic agent(s). To reduce this fraction that has been lost to the oropharynx and swallowed, a number of tube spacers of various geometric shapes and dimensions were considered, since they should, at least in theory, minimize some of the effects produced by inertial compaction, which contributes significantly to this problem.

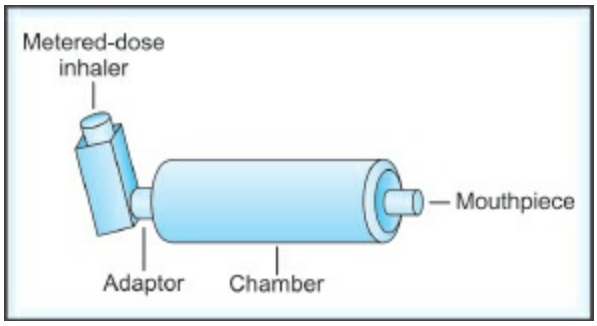



**Fig. 21.17:** A typical spacer device

Tobin et al., in an attempt to further improve and simplify the delivery of aerosolized drug from an MDI, developed a new reservoir aerosol-type delivery system (RADS), consisting of a 700-ml collapsible plastic bag into which the aerosol can be injected. This unit is designed to allow patients more time to inhale the medication after actuating than did the conventional MDI, and it eliminates some of the loss of medication associated with too rapid propulsion or inhalation of aerosolized drugs. This unit (InspirEase by Key Pharmaceuticals, Inc.) consists of a collapsible reservoir bag and a special mouthpiece that is fitted with a reed that produces a warning sound when patients are inhaling too quickly.

A new metered propellant-free intranasal pump has recently been introduced to deliver flunisolide, an effective steroidal agent in the relief of symptoms associated with seasonal or perennial rhinitus. The pump permits the administration of a metered dose of steroid without utilizing propellants, which by their cooling effects often cause smarting and irritation to the nasal mucosa. The metered aerosol pump also ensures accurate dosage and eliminates many of the administration problems associated with nose drops. The concept of propellant-free metered delivery offers a new dimension to intranasal delivery of potent therapeutic agents.

## Dry Powder Inhalers

The environmental concerns surrounding the use of chlorofluorocarbons have led to a resurgence of interest in dry powder inhaler devices. The dry powder inhalers rely on inspiration to withdraw drug from the inhaler to the lung and hence the effect of inhalation flow rate through various devices has been extensively studied. The major problem to be overcome with these devices is to ensure that the finely micronized drug is thoroughly dispersed in the airstream. It has been recommended that patients inhale as rapidly as possible from these devices in order to provide the maximum force to disperse the powder. The quantity of drug and deposition pattern varies enormously depending on the device. Rotahaler® used individual capsules of micronized drug which were difficult to handle. Modern devices use blister packs (e.g. Diskus®) or reservoirs (e.g. Turbuhaler®) (Fig. 21.18).

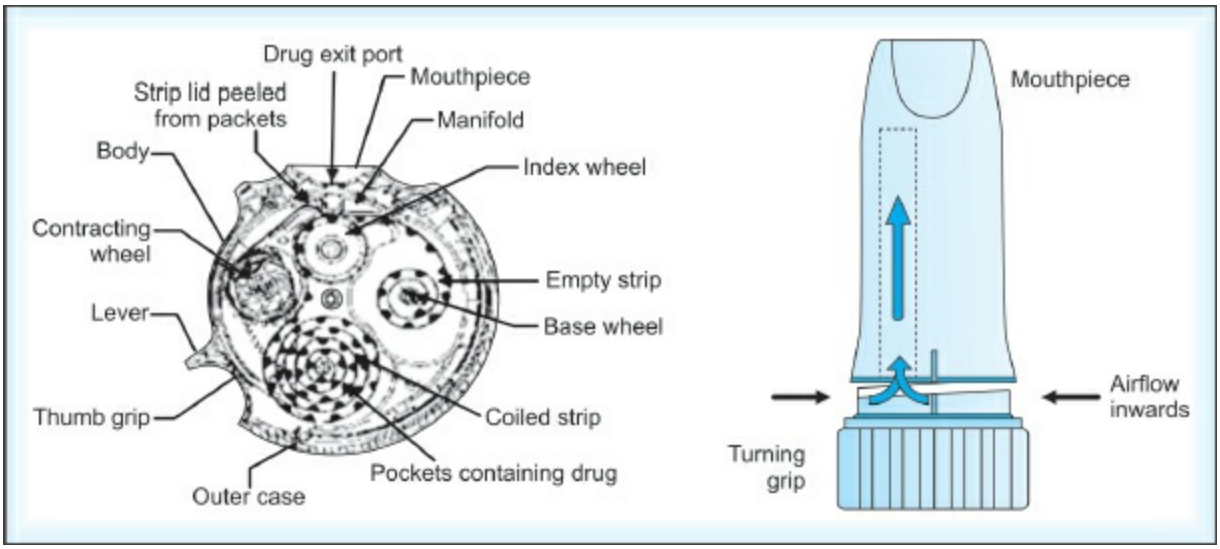



**Fig. 21.18:** A simplified view of the Diskus® and a Turbuhaler DPI

## AEROSOLS FOR INTRANASAL DRUG DELIVERY

Drug delivery systems intended for the deposition of medication into the nasal passageways have long been used as a most effective means of administering drugs intended to produce either a local or systemic effect. Until recently, the modes of administering intranasal preparations have been limited to nasal drops, nonpressurized nasal sprays (mists), inhalants, and intranasal gels (jellies), creams and ointments. A new alternative to these traditional intranasal preparations is gaining rapid popularity—the pressurized metered nasal aerosol.

The intranasal aerosol offers numerous advantages, including the delivery of a measured dose of drug, excellent depth of penetration into the nasal passageway with minimal inadvertent penetration into the lungs, reduced droplet or particle size, lower dosage than comparable systemic preparations, maintenance of sterility from dose to dose, greater patient compliance, decreased mucosal irritability and greater flexibility in product formulation.

| Trade name(s) | Dosage form | Active ingredient | Indication |
|---|---|---|---|
| Decadron Turbinaire* | Pressurized aerosol suspension | Dexamethasone sodium Phosphate | Allergic or inflammatory Nasal conditions |
| Beconase† | Pressurized aerosol suspension | Beclomethasone dipropionate | Seasonal and perennial Rhinitis |
| Vancenase†† | Pressurized aerosol suspension | Beclomethasone dipropionate | Seasonal and perennial Rhinitis |

* Merck Sharp and Dohme, West Point, PA

† Glaxo Inc., Research Triangle Park, NC

†† Schering Corporation, Kenilworth, NJ

The following aerosol products are available for intranasal administration:

Although the dexamethasone preparation (Decadron Turbinaire) was introduced into the market some 10 to 12 years ago, it was not until the introduction of beclomethasone dipropionate (Beconase and Vancerase) during the past few years that these nasal aerosol products have met with widespread success. This success has been responsible for the increased interest shown by pharmaceutical manufacturers in developing additional aerosols for administration by the nasal route. The two intranasal aerosols currently marketed in the United States are suspension formulations. The basic formulation for the intranasal aerosol suspension is as follows:

| | Weight % |
|---|---|
| Active ingredients (micronized) | up to 1.0% |
| Dispensing agent, additives, solvents, etc. | up to 1.0% |
| Propellant 12/11 (60:40) | up to 98.0% |
| | 100.0% |

In most cases, the intranasal formulation is almost identical to the comparable inhalation product (e.g. Vanceril, Vancenase; Decadron Respihaler, Decadron Turbinate). This is not surprising since the nasal passageways, like the respiratory airways, require a fine uniform distribution of material to promote therapeutic activity and minimize irritability. Various attempts have been made to formulate these active ingredients in a nonaerosol, squeeze-bottle type of product; however, because of the relatively large particle size produced by the squeeze bottle, the product has been found to be quite irritating to the nasal mucosa, with the particles being rapidly transferred to the back of the mouth, where they are swallowed or expectorated.

Probably the major difference between the inhalation aerosol product and the intranasal aerosol product is the design of the adaptor. The nasal adaptors are considerably shorter and narrower, minimizing propellant vapourization before contacting the mucosa. This results in lower percentages of smaller particles, which is desirable because this decreases the number of particles entering the respiratory airways.

Solution and emulsion types of nasal aerosols are also regarded as vehicles for the delivery of medication through the intranasal passageways, but they have proven to be intrinsically more complex to formulate. These forms of pressurized aerosol nasal preparations introduce additional problems, such as vehicle viscosity (cilia effects), irritation from additives, and leakage from nasal passageways. Many investigators are now assessing the biopharmaceutical aspects of nasal absorption and in the near future, one can optimistically expect to see a number of drugs administered intranasally for systemic action.

* The abbreviation psia is for pounds per square inch absolute, which can be converted to pounds per square inch gauge (psig) by subtracting atmospheric pressure (14.7) from psia.

# 22:
# Sterilization

Sterilization is the process designed to produce a sterile state. It is an essential concept in the production of sterile pharmaceutical products like ophthalmics and parenterals. The traditional concept of the sterile state is the absolute condition of total destruction or elimination of all living microorganisms. This concept has given way to the reality that sterile is a term that must be given relative connotation and that the probability of having achieved the absolute can only be predicted on the basis of kinetic projection of microbial death rates. Therefore, sterility in the absolute sense cannot be shown to have been achieved, but rather, can be approached with an increasing probability of success as a sterilization process is improved. With terminal methods of sterilization of a parenteral product, particularly steam under pressure, a probability of no more than one non-sterile unit in a million ($10^{-6}$) is readily achievable. Even greater levels of assurance can be achieved with current technology. In this chapter, sterile indicates a probable condition of complete freedom from viable microorganisms with the limitations just expressed; these limitations are developed more fully later in the chapter. The term aseptic indicates a controlled process or condition in which the level of microbial contamination is reduced to the degree that microorganisms can be excluded from a product during processing. It describes an “apparently” sterile state.

Persons responsible for carrying out sterilization procedures must be acutely aware of the degree of effectiveness as well as the limitations of each sterilization process. They must also understand that these processes may have a deleterious effect on the material to be sterilized. In the processing of pharmaceuticals, it is often necessary to reach a compromise between the most effective sterilization procedure and one that will not have a significant adverse effect upon the material to be sterilized. For example, it may be necessary to add an antibacterial agent to a thermally-sensitive product to enhance the effectiveness of a low-temperature sterilization process; thereby decomposition is prevented while the combined effect of the antibacterial and

the heat provides reasonable assurance that the product will be sterilized.

Microorganisms exhibit varying resistance to sterilization procedures. The degree of resistance varies with the specific organism. In addition, spores, the form that preserves certain organisms during adverse conditions, are more resistant than vegetative forms of the organism. The data given in Table 22.1 illustrate the varying resistance of different spores to moist and dry heat. Therefore, the conditions required for a sterilization process must be planned to be lethal to the most resistant spores of microorganisms normally encountered, with additional treatment designed to provide a margin of safety against a sterilization failure.

After consideration of the principles of microbial death the sterilization processes of interest in industrial pharmacy are studied in this chapter under the two main divisions of physical processes and chemical processes. Particular emphasis is placed on the principles involved and on applications of the processes to pharmaceuticals.

**Table 22.1:** Times required for lethal effect on bacterial spores by thermal exposure

| | Time (min) | | | | | |
|---|---|---|---|---|---|---|
| | Moist heat | | | Dry heat | | |
| **Organisms** | **100°C** | **110°C** | **121°C** | **120°C** | **140°C** | **170°C** |
| *B. anthracis* | 5–15 | — | — | — | 180 | — |
| *C. botulinum* | 330 | 90 | 10 | 120 | 60 | 15 |
| *C. welchii* | 5–10 | — | — | 50 | 5 | 7 |
| *C. tetani* | 5–15 | — | — | — | 15 | — |
| Soil bacilli | >1020 | 120 | 6 | — | — | 15 |

## MICROBIAL DEATH KINETICS AND TERMINOLOGY

An important term in expressing microbial death kinetics for heat, chemical and radiation sterilization is the *D*-value. The *D*-value is the time (for heat or chemical exposure) or the dose (for radiation exposure) required for the microbial population to decline by one decimal point (a 90%, or one logarithmic unit, reduction). The *D*-value may be estimated graphically, as shown in Fig. 22.1, or mathematically, as shown by Eq. (1):

$$D = \frac{U}{\log N_0 - \log N_u} \quad \ldots (1)$$

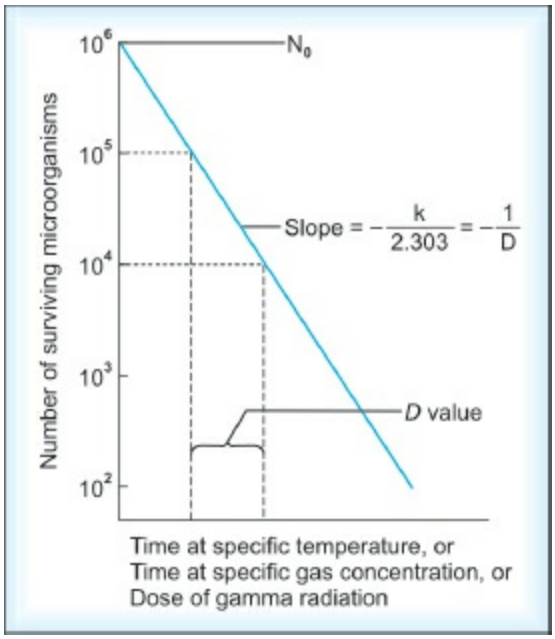



**Fig. 22.1:** Graphic representation of the semi-logarithmic microbial death rate

where, U is the exposure time or exposure dose, under specific conditions, $N_0$ is the initial microbial population (product bioburden) and $N_u$ is the microbial population after receiving U time or dose units of sterilant exposure. For example, after 5 min of product exposure to a temperature of 121°C, the microbial population was reduced from $2 \times 10^5$ to $6 \times 10^3$. Then, the *D*-value at 121°C is:

$$D_{121} = \frac{5 \text{ min}}{\log(2\times10^5) - \log(6\times10^3)} = 3.28 \text{ min}$$

Thus, at 121°C, the microbial population is decreased by 90% every 3.28 min.

*D*-values have been defined precisely for various microorganisms contained in certain environments (liquids and solid surfaces) at specific temperatures for heat sterilization, and at direct exposure to cobalt-60 irradiation. *D*-values cannot be defined precisely for microorganisms exposed to such gases as ethylene oxide because of the complex interaction of heat, concentration of gas, and relative humidity. *D*-values are estimated for gas sterilization when it is possible to keep heat and humidity values constant, varying only the concentration of gas.

Other key terms used in the determination of microbial death rates include *microbial load*, or *bioburden;* the *Z*-value; the *F-value*; the $F_0$ *value;* and the probability of non-sterility. These terms are defined in Table 22.2 and *Z* value plots are shown in Fig. 22.2, where

**Table 22.2:** Definition of key terms employed in microbial death kinetics

| Symbol | Term | Definition |
|---|---|---|
| $N_0$ | Bioburden | The population or number of living microorganisms per defined unit, surface, or system |
| $Z$ | Resistance value | The number of degrees (C or F)* required for a 1 log reduction in the *D*-value $Z = \frac{T_2 - T_2}{\log D_2 - \log D_1}$ |
| $F(T, Z)$ or $F_T^Z$ | Sterilization process equivalent time | The equivalent time at temperature, *T* delivered to a unit of product calculated using a specified value of *Z* |
| $F_0$ | Sterilization process equivalent time | The equivalent time at a temperature of 121°C delivered to a unit of product calculated using a *Z*-value of 10°C |
| $N_u$ | Probability of nonsterility $N_u = \text{anti}\log\left(\log N_0 \frac{U_T}{D}\right)$ | The number of non-sterile units per batch or the theoretic or extrapolated number of living micro-organisms per defined unit after a given equivalent heating time, *U* at a specific temperature, *T* |

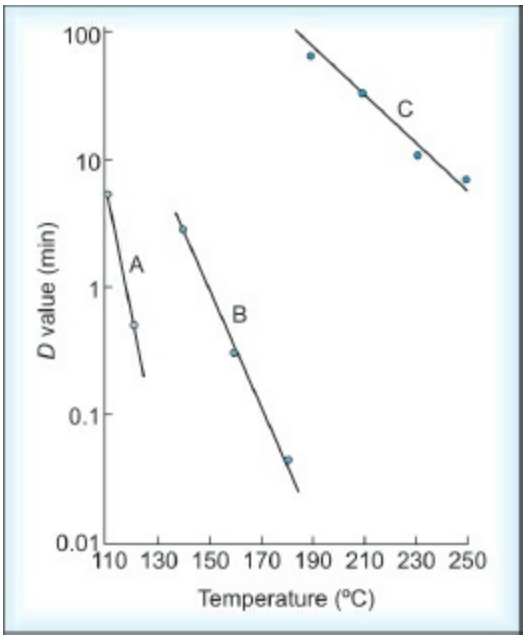



**Fig. 22.2:** *Z*-value plots of log *D* versus temperature

*A: Z = 10°C for B. stearothermophilus spores exposed to steam sterilization.*

*B: Z = 22°C for B. subtilis var niger spores exposed to dry heat sterilization.*

*C: Z = 54°C for E. coli endotoxin exposed to dry heat sterilization*

The $F_0$ value is a term widely used in sterilization cycle design and validation. Its current application is limited to steam sterilization although an *F*-value can be computed for any thermal method of sterilization. The $F_0$ value can be defined by the following two equations:

$$F_0 = \Delta t \sum 10^{\frac{T-121}{10}} \quad \text{... (2)}$$

where, *t* is the time interval between product temperature measurements *T*.

$$F_0 = D_{121} (\log N_0 - \log N_u) \text{ ... (3)}$$

where, $N_0$ and $N_u$ are those terms defined previously.

The $F_0$ value of Eq. (2) is obtained by physical measurement of product temperature and substitution of that temperature for *T* in the exponent. For example, if the product temperature was measured every 5 min from 0 to 30 min and found to be 25°C, 110°C, 118°C, 120°C, 121°C, and 100°C, the $F_0$

value would be:

$F_0$ = 5 min (0 + 0.079 + 0.501 + 0.794 + 5.000 + 0.0079)
$F_0$ = 5 min (6.382)
$F_0$ = 31.91 min

By definition, when the $F_0$ value is used, the *Z*-value is assumed to be 10°C. This means that for every 10°C increase in product temperature, the *D*-value is decreased by 90%, or 1 log unit.

Equation (3) is the biologic $F_0$ equation because the $F_0$ value is calculated after determining the $D_{121}$ value and the product bioburden, $N_0$. The probability of nonsterility is whatever level is desired, usually a minimum of $10^{-6}$. In general, Eq. (3) is applied under two circumstances. Given $D_{121}$, $N_0$ and $N_u$, the $F_0$ value is calculated. For example, if $D_{121}$ = 1 min, $N_0 = 10^2$ and $N_u = 10^{-6}$, then:

$F_0$ = 1 min (log $10^2$ – log $10^{6}$)
$F_0$ = 8 min

Given $D_{121}$, $N_0$ and $F_0$, the achieved level of nonsterility may be calculated. For example, if $D_{121}$ = 2 min, $N_0 = 10^2$ and $F_0$ = 8 min, then:

$$N_u = \text{antilog}\left(\log N_0 - \frac{F_0}{D_{121}}\right) \quad \text{... (4)}$$

$$N_u = \text{antilog}\left(\log 10^2 - \frac{8}{2}\right)$$

$$N_u = 10^{-2}$$

## F-Value Applications

The importance of $F_0$ values in steam sterilization cycle validation may be summarized as follows:

1. $F_0$ relates the killing efficiency of the process at any temperature with the killing effect produced at the desired sterilization temperature of 121°C.
2. $F_0$ provides a single quantitative value describing the thermal exposure time of the cycle to which the product was exposed equivalent to 121°C.
3. $F_0$ incorporates the contribution of the heating and cooling portions of the temperature-time profile during a cycle with the overall lethal effect of heat upon microorganisms.
4. $F_0$, if used to describe the lethal effect upon microorganisms at the coolest location in the sterilizer, represents the most conservative estimate of the degree of destruction of microorganisms, and thus the safest conditions for determining cycle time.

At least three factors affect the $F_0$ value. They are (1) the container characteristics: Size, geometry, and heat transfer coefficient, (2) the product volume and viscosity and (3) the size and configuration of the batch load in the sterilizer.

*F*-value equations can be applied to dry heat sterilization although most materials sterilized by dry heat can be subjected to overkill temperature-time cycles. The reference temperature $T_0$ would not be 121°C, of course, but it probably would be 170°C, since this temperature is specified in the USP/NF XXI. The Z-value would not be 10°C, but would be in the range of 22°C for the destruction of *B. subtilis var niger* spores on glass to 54°C for the destruction of endotoxin (*see* Fig. 22.2).

Aseptic processing also requires validation to assure batch to batch consistency in producing a given probability of product sterility. While *D* and $F_0$ values cannot be applied, a probability of nonsterility levels can be obtained by process—simulation testing using microbiologic growth medium, a suitable type arid number of challenge microorganisms, and a relevant number of containers. The percent contamination level (% C) is calculated as follows:

$$\%C = \frac{N_G}{N_T - N_D} \times 100 \qquad \text{... (5)}$$

where, $N_G$ is the number of undamaged containers with microbial growth, $N_T$ is the total number of containers filled and $N_D$ is the number of damaged contaminated containers. Procedures for validation of aseptic fill for solution drug products have been presented in a recent publication by the Parenteral Drug Association.

# STERILIZATION METHODS

## Physical Processes of Sterilization

### *Thermal Methods*

The lethal effectiveness of heat on microorganisms depends upon the degree of heat, the exposure period, and the moisture present. Within the range of sterilizing temperatures, the time required to produce a lethal effect is inversely proportional to the temperature employed. For example, sterilization may be accomplished in 1 h with dry heat at a temperature of 170°C, but may require as much as 3 h at a temperature of 140°C. While it is common practice to identify cycle times in terms of the maximum temperature hold time, total heat input may be computed in terms of F values, as explained previously; however, the lethal effect must be computed in terms of the time during which the entire mass of the material is heated.

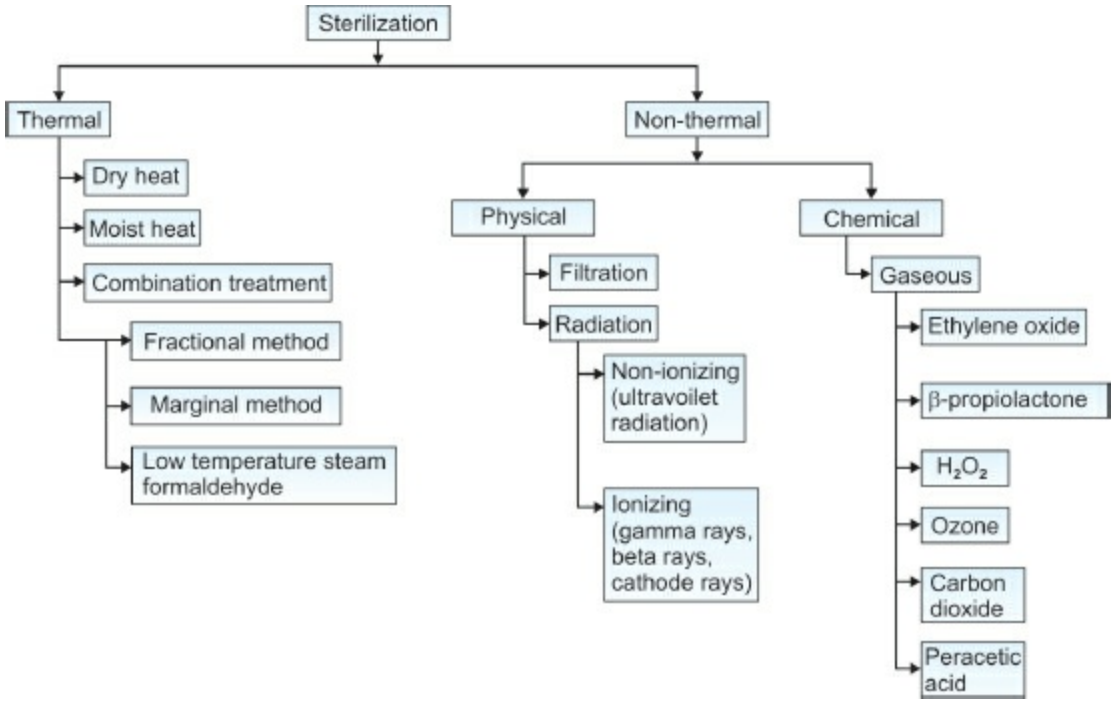



**Fig. 22.3:** Classification of sterilization methods

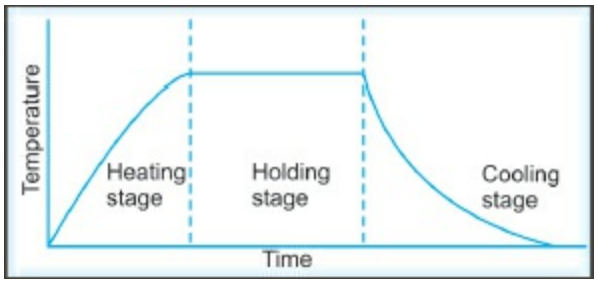



**Fig. 22.4:** Typical temperature profile of a heat sterilization process

Thermal methods of sterilization may conveniently be divided into those accomplished by dry heat and those by moist heat, both of which can be used for terminal sterilization of pharmaceutical preparations and medical devices.

## *Dry Heat*

Substances that resist degradation at temperatures above approximately 140°C (284°F) may be rendered sterile by means of dry heat. Two hours exposure at temperature of 180°C (356°F) or 45 min at 260°C (500°F) normally can be expected to kill spores as well as vegetative forms of all microorganisms. This total sterilizing cycle time normally includes a reasonable *lag time* for the substance to reach the sterilizing temperature of the oven chamber, an appropriate hold period to achieve sterilization and a cooling period for the material to return to room temperature.

**Mechanism of action:** Dry heat is believed to exert its lethal action upon microorganisms by oxidizing proteins, affecting particularly the reproductive process.

**Hot air ovens:** The ovens used to achieve hot air sterilization are of two types, natural convection and forced convection. Circulation within *natural-convection ovens* depends upon the currents produced by the rise of hot air and fall of cool air. This circulation can be easily blocked with containers, resulting in poor heat distribution efficiency. Differences in temperature of 20°C or more may be found in different shelf areas of even small laboratory ovens of the natural-convection type.

*Forced-convection ovens* provide a blower to circulate the heated air around the objects in the chamber. Efficiency is greatly improved over natural convection. Temperature differences at various locations on the shelves may be reduced to as low as ±1°C. The lag times of the load material therein also are greatly reduced because fresh hot air is circulated rapidly

around the objects. The curves shown in Fig. 22.5 illustrate the difference in lag time for some of the same containers of corn oil when heated in a natural convection oven as compared with the same oven equipped for forced circulation.

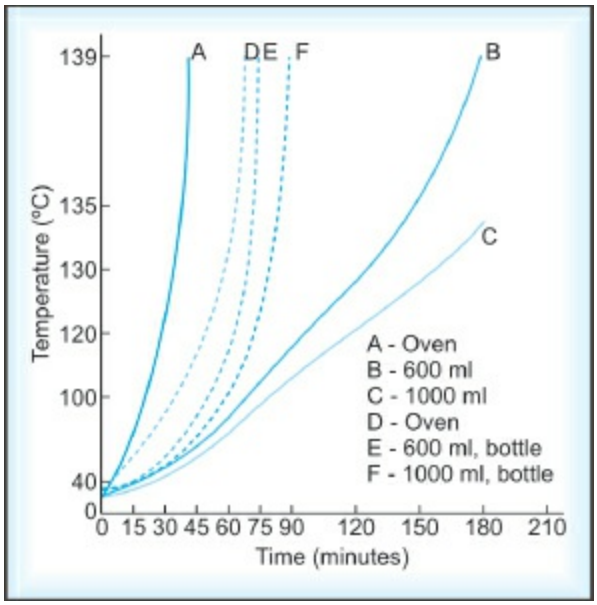



**Fig. 22.5:** Rate of heating corn oil in pyrex liter bottles in the same hot air oven with natural convection (—•—) and forced circulation (---•---)

**Dry heat tunnels:** Another type of sterilizer is the tunnel unit with a moving belt, designed to thermally sterilize glass bottles and similar items as they move through the tunnel. The items are cooled with clean air before they exit the tunnel, usually directly into an aseptic room and linked in a continuous line with a filling machine. Such units require careful validation.

**Factors in determining cycle time:** The cycle time is composed of three parts (1) the thermal increment time of both the chamber and the load of material to be sterilized, assuming that both start at room temperature, (2) the hold period at the maximum temperature and (3) the cooling time. The material lags behind the increasing temperature of the chamber. The time required for all of the material to "catch-up" with the temperature of the chamber is longer with larger quantites of material, poorer thermal conductance properties of the material, and lower heat capacity. The relationship of these factors must be carefully determined during validation studies so that effective cycle times can be planned.

The cycle time is most commonly prescribed in terms of the hold time, for example, 2 h at 180°C dry heat. The hold time may be shown by sensors detecting the temperature of the chamber at its coolest spot; however, a better indication of the actual thermal condition is obtained by sensing, usually with a thermocouple, the coolest spot in the load of the material to be sterilized. When such a location is used, and when this coolest spot is known from previous validation studies, the timing required for sterilization is correctly programmable. It should be remembered that other parts of the load of material may be heated for a longer period and if it is thermally unstable, degradation could occur. Therefore, the thermal stability of the material to be sterilized must be known and the optimum method of sterilization selected to achieve effective sterilization throughout the entire mass of the material while maintaining its stability and integrity.

**Effect on materials:** The elevated temperatures required for effective hot air sterilization in a reasonable length of time have an adverse effect on many substances. Cellulose materials, such as paper and cloth, begin to char at a temperature of about 160°C (320°F). At these temperatures, many chemicals are decomposed, rubber is rapidly oxidized, and thermoplastic materials melt. Therefore, this method of sterilization is reserved largely for glassware, metal-ware, and anhydrous oils and chemicals that can withstand the elevated temperature ranges without degradation. Expansion of materials is also appreciable, as they are heated from room to sterilizing temperatures. Therefore, glassware must not be wedged tightly in the oven chamber, containers for oils must be large enough to permit expansion of the oil, and provision must be made for the expansion of other substances.

Advantage may be taken of the anhydrous state achieved with this method of sterilization to provide dry glassware and metalware at the end of an adequate heating cycle. Dry equipment and containers are essential in the manufacture of an anhydrous product, but they are also desirable to prevent dilution of an aqueous product. Also, dry equipment can be kept sterile during storage more easily than wet equipment. Further, dry heat effectively destroys pyrogens, usually requiring about twice the hold time for sterilization.

To maintain a sterile condition after sterilization, environmental contamination must be excluded. The openings of equipment must be covered with a barrier material such as aluminum foil. As an alternative,

items to be sterilized may be placed in a covered stainless steel box or similar protective container.

**Applications:** Dry heat sterilization is used for powders, containers and equipment whenever possible because an adequate cycle results in sterile and dry equipment. Highspeed processing lines recently developed have included a hot-air tunnel for the continuous sterilization of glass containers, which are heated by infrared lamps or by electrically heated, filtered, circulating air. Glass and metal equipment usually withstand dry heat sterilization without difficulty, although uneven thermal expansion may cause breakage or distortion. However, rubber and cellulosic materials undergo degradation. Certain ingredients, such as chemicals and oleaginous vehicles, to be used in sterile pharmaceutical preparations are sometimes sterilized with dry heat at lower (usually, 140°C or less) temperatures. In such cases, it must be established that the heating cycle has no deleterious effects on the ingredients and that the cycle time is adequate to achieve sterilization. They must also be carefully protected after sterilization until incorporated aseptically in the product to prevent contamination from the environment.

## *Moist Heat (Autoclaving)*

Moist heat is more effective than dry heat for thermal sterilization. It should be remembered, however, that normal moist heat cycles do not destroy pyrogens.

**Mechanism of action:** The mechanism by which microorganisms are killed by moist heat is thought to involve the coagulation of proteins of living cells. The thermal capacity of steam is much greater than that of hot air. At the point of condensation (dew point), steam liberates thermal energy equal to its heat of vapourization. This amounts to approximately 540 calories per gram at 100°C (212°F) and 524 calories per gram at 121 °C (250°F). In contrast, the heat energy liberated by hot dry air is equivalent to approximately only 1 calorie per gram of air for each degree centrigrade of cooling. Therefore, when saturated steam strikes a cool object and is condensed, it liberates approximately 500 times the amount of heat energy liberated by an equal weight of hot air. Consequently, the object is heated much more rapidly by steam. In addition, when steam under pressure is employed, a rapidly changing fresh supply of heat-laden vapour is applied to the object being heated. This is due both to the pressure under which steam is

applied and to the partial vacuum produced at the site where steam is condensed, for it shrinks in volume by about 99% as it condenses. The data given in Table 22.3 illustrate this principle, using the effect of varying amounts of water on the temperature required to coagulate egg albumin. The temperature required is inversely related to the moisture present.

**Table 22.3:** Effect of moisture and heat on egg albumin

| Water (%) | Temperature (°C) | Effect |
|---|---|---|
| 50 | 56 | Coagulation |
| 25 | 80 | Coagulation |
| 6 | 145 | Coagulation |
| 0 | 170 | Coagulation and oxidation |

**Autoclave:** The density of steam is lower than that of air. Therefore, steam enters an autoclave chamber and rises to the top, displacing air downward, as illustrated by the gravity displacement autoclave shown in Fig. 22.6. Objects must be placed in the chamber with adequate circulation space around each object and so arranged that air can be displaced downwards and out of the exhaust line from the chamber. Any trapped air, e.g. air in containers with continuous sides and bottoms, or in tightly wrapped packs, prevents the penetration of steam to these areas, and thus prevents sterilization. The air trapped in this manner is heated to the temperature of the steam, but hot air at a temperature of 120°C (248°F) requires a cycle time of 60 h to ensure a lethal effect on spores. A 20-min exposure at this temperature with hot dry air, therefore, would be entirely inadequate.

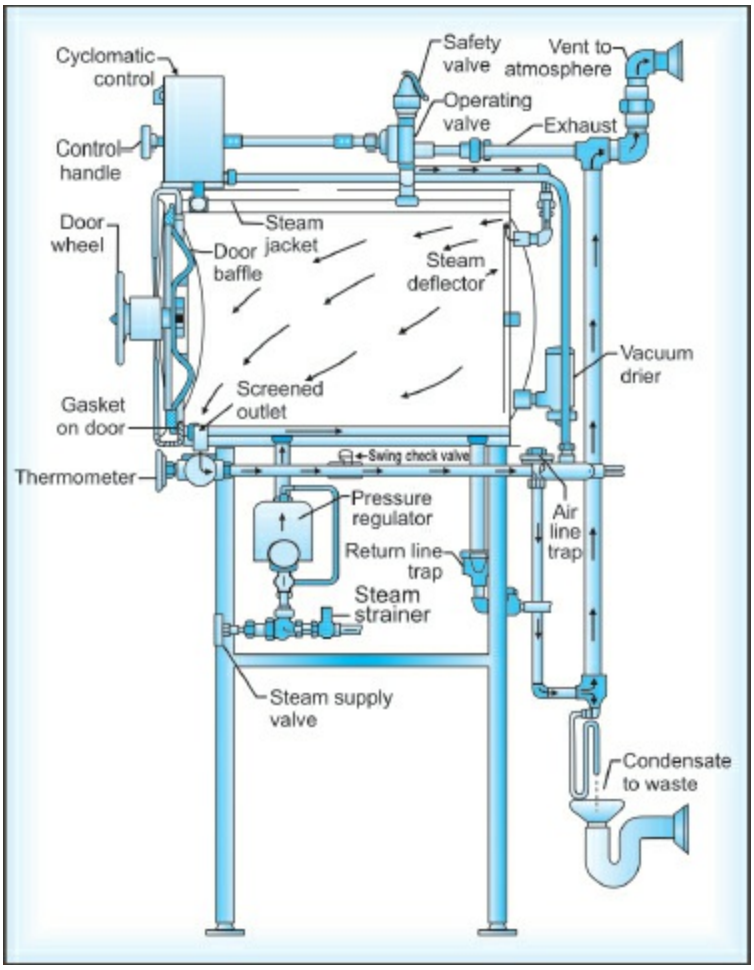



**Fig. 22.6:** Cross-sectional diagram of the functional parts of an autoclave

**Air-steam mixtures:** While air-steam mixtures have a lower temperature and lower thermal capacity than pure steam, the presence of air may be utilized to control the pressure in the chamber when flexible-walled containers of products are being sterilized. For example, plastic bags of large-volume parenterals (LVPs) or collapsible tubes of aqueous jellies would swell and burst in an autoclave utilizing steam only, particularly during the cooling phase. When air is mixed with the steam and the air pressure is independently controlled, the pressure applied to the outside of the containers can be adjusted to equal the internal pressure, so that the containers do not burst. Because of the tendency of steam and air to stratify, the mixture must be mixed continuously; this is usually accomplished by means of a blower.

**Factors determining cycle time:** Spores and vegetative forms of bacteria may be effectively destroyed in an autoclave employing steam under pressure during an exposure time of 20 min at a pressure of 15 psig 121°C (250°F) or

as little as 3 min at 27 psig 132°C (270°F). These time intervals are based on the assumption that the steam has reached the innermost recess of the material to be sterilized and that the temperature of the material is held for at least one half of that time interval. In the case of bottles of solution, the heat must be conducted through the wall of the container, raise the temperature of the solution to that of its environment, and generate steam within the container from the water therein. Therefore, a significant lag time is involved before the solution reaches the sterilizing temperature.

The determination of lag time and its inclusion in the planned total cycle time is no less important for moist-heat sterilization than hot-air sterilization, as discussed previously. By way of illustration, it has been found that 1200 ampoules, each containing 5 ml of a solution, can be effectively sterilized in an autoclave at 121°C (250°F) during an exposure time of 20 min. A single bottle containing the same total volume of solution (6 L) required an exposure of 60 min at 121°C (250°F).

**Approaches for reduction of cycle time:** Prolonged heating of most objects is detrimental to the material. For example, fabrics and rubber parts deteriorate with loss of tensile strength, solutions may undergo adverse chemical changes, and metal objects may become pitted. Therefore, the total cycle time should be controlled so that the heating period is not unnecessarily prolonged. Usually, this is best accomplished by shortening the cooling period. For non-sealed items of equipment or containers that do not contain solutions, the steam may be exhausted to the outside rapidly at the end of the sterilizing cycle. Objects are thereby cooled rapidly, particularly if removed from the autoclave chamber. Such a procedure cannot be employed for solutions, whether sealed or unsealed in containers, because the rapid release of chamber pressure would cause violent ebullition of the hot solution, with spattering of the contents of unsealed containers and explosion of sealed containers.

**Table 22.4:** Pressure-temperature relationships and antimicrobial efficacies of alternative steam sterilization cycles

| Temperature (°C) | Holding time (minutes) | Steam pressure (kPa) | (psi) | Inactivation factor* (decimal reductions) |
|---|---|---|---|---|
| 115 | 30 | 69 | 10 | 5 |
| 121 | 15 | 103 | 15 | 10 |
| 126 | 10 | 138 | 20 | 21 |
| 134 | 3 | 207 | 30 | 40 |

* Calculated for a spore suspension having a $D_{121}$ of 1.5 minutes an a Z-value of 10°C

One method for rapid extraction of heat from sealed containers of solutions is to spray the containers with gradually cooling water, while the pressure in the chamber is concurrently reduced. Another accelerated cooling method employs short pulses of high-pressure steam introduced into the loaded chamber. As the steam expands in the chamber, it extracts heat from the containers of solution. The steam is exhausted from the chamber at a rate that provides for a gradual reduction of the pressure concurrent with the temperature reduction. By these methods, it is sometimes necessary to introduce pulses of air into the chamber to replace all or part of the steam so that the pressure around the containers is not reduced too rapidly. By the spray cooling method, it has been reported that the cooling time for a load of 200 one-liter bottles of solution may be reduced from about 20 h to about 20 min.

A relatively new approach to a reduction in the total heating cycle time has been the introduction of a precycle vacuum. In a specially designed autoclave, a precycle vacuum of at least 20 mm Hg is drawn. More recent studies have shown that a double vacuum drawn in a sequence prior to the heating cycle removes the air more effectively from porous materials. The subsequent introduction of steam permits rapid penetration and load heating with complete elimination of air pockets. Since the total heating period is markedly reduced owing to the reduction in the temperature increment time, a higher temperature (usually 135°C (275°F) may be employed with less deleterious effects on materials. This method is particularly suited to operating room packs in hospitals, where the total cycle time for large packs has been reduced from about 78 min by the conventional method to about 14 min. Such a method cannot be used for solutions or other objects that cannot withstand the high vacuum employed.

**Combination treatment:** Moist heat is also used for lower temperature sterilization procedures. Temperatures of 100°C (212°F) or lower are used for

these so-called *marginal, or fractional,* methods. The term marginal originates from the questionable reliability of the processes. The term fractional is derived from the fact that these processes are normally performed by two or three exposures to moist heat, alternated with intervals during which the material is held at room or incubator temperatures.

Fractional methods of sterilization such as tyndallization, employing a temperature of 100°C (212°F), and inspissation, employing temperatures as low as 60°C (MOT), are relatively effective in reducing the number of vegetative forms of microorganisms, but are unreliable against spores. For certain preparations, the effectiveness of these processes may be improved by the inclusion of a bacteriostatic agent. These marginal methods of sterilization should be reserved for substances that must be processed by a thermal method but that cannot withstand higher temperatures without degradation. The assurance of sterility is comparatively low, however. In pharmaceutical industry, another type of combination treatment involves low-temperature steam formaldehyde (LTSF). The steam is below atmospheric pressure (70–80°C) and formaldehyde gas produces sporicidal effect. Formaldehyde has been used to sterilize medical devices but care must be taken because it is toxic to human and is also mutagenic and carcinogenic.

The amount of heat needed for sterilization can be reduced by using combined treatment of heat plus a low water potential and reduced pH (below 4.5). Although some spores may survive the initial treatment, they then germinate and die, reducing the number of organisms present as time passes, a process called *autosterilization.*

**Wrapping materials:** Wrappings for equipment and supplies subjected to moist heat sterilization must permit easy penetration of steam and escape of air. They must also possess sufficient wet strength so that they will not tear or burst during the process. After sterilization, the wrapping must provide an efficient bacterial barrier so that equipment remains sterile for a reasonable time until used. In addition, maintenance of sterility depends upon complete coverage of the contents of the pack, drying of the wrapping after the process, and a static air state within. Acceptable disposable process wrapping materials include 20-lb weight Kraft paper, special parchment paper and Tyvek. Reusable types include close-weave nylon and Dacron. Except for Kraft paper, all are low-lint materials.

**Application:** It is generally accepted that the most reliable thermal

method of sterilization is the use of moist heat under pressure. Therefore, this method of sterilization should be employed whenever possible. Aqueous pharmaceutical preparations in hermetically-sealed containers that can withstand the temperature of autoclaving can be rendered sterile and remain so indefinitely unless tampering with the seal occurs. Non-aqueous preparations in sealed containers cannot be sterilized in this manner during a normal cycle because no water is present within the container to generate steam, and thereby effect sterilization.

Moist heat sterilization is also applicable to equipment and supplies such as rubber closures, glassware and other equipment with rubber attachments; filters of various types; and uniforms. To be effective, however, air pockets must be eliminated. This normally requires that the items be wet when placed in the autoclave. They also will be wet at the end of the sterilizing cycle. When moisture can escape without damage to the package, part of the moisture can be removed by employing an evacuation step at the end of the cycle. Even this process does not usually completely dry the equipment. Therefore, when such an equipment is used in processing, allowance must be made for the diluting effect of this water, or preferably, a small portion of the product may be used to rinse or flush the water out of the equipment. In some instances, when dry equipment is required and it must be sterilized by autoclaving, the equipment may be dried in a vacuum oven before use.

## *Non-thermal Methods (Cold sterilization)*

## *Radiation Sterilization*

**Ultraviolet light:** Ultraviolet light is commonly employed to aid in the reduction of contamination in the air and on surfaces within the processing environment. The germicidal light produced by mercury vapour lamps is emitted almost exclusively at a wave length of 2537 Å (253.7 nm). It is subject to the laws for visible light, i.e. it travels in a straight line, its intensity is reduced in proportion to the square of the relative distance it travels and it penetrates materials poorly or selectively. Ultraviolet light penetrates clean air and pure water well, but an increase in the salt content and/or the suspended matter in water or air cause a rapid decrease in the degree of penetration. For most other applications, penetration is negligible, and any germicidal action is confined to the exposed surface.

**Mechanism of action:** When ultraviolet light passes through matter, 

energy is liberated to the orbital electrons within constituent atoms. This absorbed energy causes a highly energized state of the atoms and alters their reactivity. When such an excitation and alteration of activity of essential atoms occurs within the molecules of microorganisms or of their essential metabolites, the organism dies or is unable to reproduce. The principal effect may be on cellular nucleic acids, which have been shown to exhibit strong absorption bands within the ultraviolet wavelength range.

**Lethal dosage:** The lethality of ultraviolet radiations has been well established; however, it also has been shown that organisms exposed to ultraviolet radiations can sometimes recover a fact not surprising if the previously described theory of lethality is correct. Recovery has been increased by the addition of certain essential metabolites to the culture, adjustment of the pH of the medium, or exposure to visible light shortly after exposure to the ultraviolet radiations. Therefore, adequate exposure to the radiations must occur before reliance can be placed upon obtaining a sterilizing effect. The germicidal effectiveness of ultraviolet light is a function of the intensity of radiation and time of exposure. It also varies with the susceptibility of the organism. The data in Table 22.5 show some of this range of susceptibility. It can be seen that if the intensity of radiation on a surface was 20 microwatts per $cm^2$, the minimum intensity usually recommended, it would require approximately 1100-s exposure to kill *B. subtilis* spores, but only approximately 275 s to kill *S. hemolyticus*. The intensity of UV radiation can be measured by means of a light meter having a phototube sensitive to the 2537 Å wavelength.

**Table 22.5:** Intensity of radiation at 2537 A necessary to completely destroy certain microorganisms

| **Organism** | **Energy (mw-s/$cm^2$)** |
|---|---|
| *Bacillus subtilis* | 11,000 |
| *B. subtilis spores* | 22,000 |
| *Eberthella typhosa* | 4,100 |
| *Escherichia coli* | 6,600 |
| *Pseudomonas aeruginosa* | 10,500 |
| *Sarcina lutea* | 26,400 |

| | |
|---|---|
| *Staphylococcus aureus* | 6,600 |
| *Streptococcus hemolyticus* | 5,500 |
| *Saccharomyces cerevisiae* | 13,200 |
| *Penicillium roqueforti* | 26,400 |
| *Aspergillus niger* | 330,000 |
| *Rhizopus nigricans* | 220,000 |

**Maintenance and use:** To maintain maximum effectiveness, ultraviolet lamps must be kept free from dust, grease, and scratches because of the large reduction in emission intensity that will occur. Also, they must be replaced when emission levels decrease substantially (about 30 to 50%), owing to energy-induced changes in the glass that inhibits the emission.

Personnel present in areas where ultraviolet lights are on should be protected from the direct and reflected rays. These rays cause reddening of the skin and intensely painful irritation of the eyes. The American Medical Association has recommended that the maximum safe human exposure for 1 h be limited to 2.4 mw/cm$^2$.

Ultraviolet lamps are used primarily for their germicidal effect on surfaces or for their penetrating effect through clean air and water. Therefore, they are frequently installed in rooms, air ducts and large equipment in which the radiation can pass through and irradiate the air, and also reach exposed surfaces. Water supplies also have been sterilized when the limit of penetration has been carefully determined and controlled so that adequate irradiation throughout has been achieved.

**Ionizing radiations:** Ionizing radiations are high-energy radiations emitted from radioactive isotopes such as cobalt-60 (also caesium-137) (*gamma rays*) or produced by mechanical acceleration of electrons to very high velocities and energies (*cathode rays, beta rays).* Gamma rays have the advantage of being absolutely reliable, for there can be no mechanical breakdown; however, they have the disadvantages that their source (radioactive material) is relatively expensive and the emission cannot be shut off as it can from the mechanical source of accelerated electrons. Accelerated electrons also have the advantage of providing a higher and more uniform dose rate output.

**Mechanism of action:** Ionizing radiations destroy microorganisms by stopping reproduction as a result of lethal mutations. These mutations are brought about by a transfer of radiation beam energies to receptive molecules in their path, the direct-hit theory. Mutations also may be brought about by indirect action in which water molecules are transformed into highly-energized entities such as hydrogen and hydroxyl ions. These, in turn, bring about energy changes in nucleic acids and other molecules, thus eliminating their availability for the metabolism of the bacterial cell. Ionizing radiations differ from ultraviolet rays in their effects on matter, primarily in that the former are of a higher energy level, actually producing ionization of constituent atoms.

**Lethal dosage:** The dosage is determined by the energy released by the gamma rays or by the number of electrons that impinge on each square centimeter of absorbing substance (the target). The *rad* is the unit of absorbed radiation, the unit of dosage now most frequently employed. It is arbitrarily defined as the absorption of 100 ergs of energy per gram of substance. The depth of penetration within a target of a given dose is directly related to the electron voltage of the source, and indirectly related to the density of the material to be irradiated. Bacterial spores and viruses are generally four to five times more resistant than vegetative bacteria and molds. A dose of 2 to 2.5 megarads, however, is considered adequate to ensure sterility. Currently, there is no evidence of reactivation of microorganisms as has been found with ultraviolet light.

**Electron accelerators:** Electron accelerators are of two general types, the linear and the Van de Graaff accelerators. The principle of the linear accelerator may be followed from Fig. 22.7. Very high-frequency microwaves (radar) collect electrons from a cathode and accelerate them as they travel through the vacuum tube, reaching almost the speed of light. The electrons are emitted and directed to the target at an energy range of 3 to 15 million electron volts (MeV). Since energy potentials of 10 MeV or higher may produce radioactive materials, linear accelerators of more than 9 MeV are not normally used for sterilization.

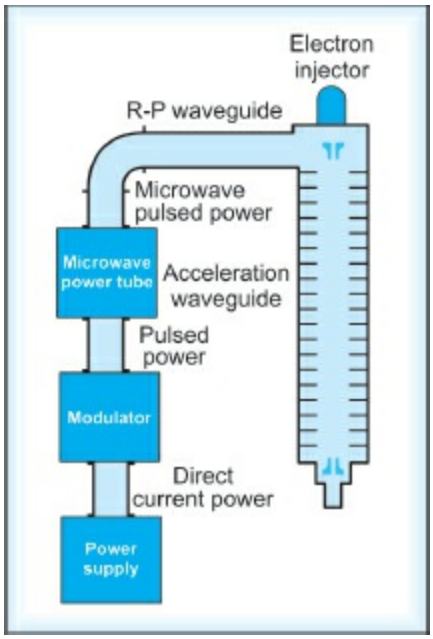



**Fig. 22.7:** Operating principle of a linear electron accelerator (*Courtesy*of High Voltage Engineering Corp)

The van de Graaff accelerators are capable of energy potentials up to 3 MeV. They utilize the force exerted on a charged particle by a high voltage potential in an electric field as a means of direct particle acceleration.

**Applications:** Accelerated electrons or gamma rays may be used to sterilize select products by a continuous process. Most other product sterilization procedures must be performed in batches. Continuous process sterilization requires exacting control so that there are no momentary lapses in sterilizing effectiveness. Assurance of adequate dose delivery, complete and uniform coverage of the product and adequate penetration have been achieved in the effective and routine sterilization of sutures, using a linear accelerator. Adequate dosage is usually determined by the effect of the absorbed energy, at the maximum determined depth of penetration, on photographic film and/or on the biologic indicator *B. pumilus.*

The use of radiation is increasing in frequency and extent as experience is gained with this method, particularly for the sterilization of medical plastic devices. It has been given new impetus by the question raised by the Occupational Safety and Health Administration (OSHA) on the safety of ethylene oxide and the low environmental level now being permitted. Availability of facilities for this method, using both energy sources, is

increasing. An individual medical device or pharmaceutical manufacturer may not justify the high cost of a facility for radiation sterilization, but the increasing availability of centers performing contract services is making this method a more viable option.

A number of vitamins, antibiotics and hormones in the dry state have been successfully sterilized by radiation. Liquid pharmaceuticals are more difficult to sterilize because of the potential effect of the radiations on the vehicle system as well as the drug.

## *Filtration Sterilization*

Filtration is an absolute process which assures the removal of particles, including microorganisms above a definite size, from solutions and gases without the application of heat. Ideally, filters should not alter the solution or gas in any way, neither removing the desired constituents nor imparting undesired components. This requirement essentially limits the types of filters currently employed to the polymer types listed in Table 22.6. Furthermore, almost all of those currently in use with parenteral solutions and gases are of the membrane-type, that is, tissue-thin materials removing particles primarily by sieving. When a filter removes constituents from a solution, such a removal is usually due to the phenomenon of adsorption, which being a surface phenomenon, occurs during only the first portion of the filtration, that is, until the surface of the filter is saturated with the adsorbed molecule or ion. The most common attack on the filter itself is due to the solvent properties of the vehicle of certain parenteral products. Since the most common solvent for parenteral solutions is water, and the use of other types of solvents is limited, this usually is not a problem. Moreover, the development of membrane filters composed of materials having high resistance to most pharmaceutical solvents has further reduced this problem.

**Table 22.6:** Comparative characteristics of the membrane filters (0.2 μm porosity) used for sterilization

| Filter type | Particular uses | Pharmaceutical solvent to be avoided | Trade name | Sterilization—Autoclavable | Company |
|---|---|---|---|---|---|
| **HYDROPHILIC** | | | | | |
| Acrylic with nylon substrate | Aqueous solutions and glycols | Dimethylformamide | Versapor<br>AN Hydrophilic<br>Acropor | No, ETO<br>No, ETO | Gelman<br>Gelman |
| Cellulose acetate/nitrate | Aqueous solutions | Benzyl alcohol, ethanol, propylene glycol, dimethyl-formamide | MF-Millipore<br>SM 11307<br>Membrafil<br>BA 83 | No, ETO<br>Yes<br>Yes<br>No, ETO | Millipore<br>Sartorius<br>Nuclepore<br>S and S |
| Cellulose acetate | Aqueous solutions with alcohol resis-tance | Benzyl alcohol, ethanol (selective), propylene glycol, dimethylformamide | GA Metricel<br>Celotate<br><br>OE66<br>SM 11107 | Yes<br>No, ETO<br><br>No, ETO<br>Yes, and dry heat | Gelman<br>Millipore<br>S and S<br>Sartorius |
| Cellulose | Aqueous solutions and phar maceutical solvents | Solvent resistant | SM 11607<br><br>RC58 | Yes, and dry heat<br><br>Yes | Sartorius<br><br>S and S Pall |
| Nylon 66 (polyamide) | Aqueous solutions | Solvent resistant | Ultipor $N_{66}$ | Yes | Nuclepore |
| Polycarbonate | Aqueous solutions | Benzyl alcohol, dimethylformamide | Nuclepore<br>Uni-Pore | Yes<br>Yes | Nuclepore<br>Bio-Rad |
| Polysulfone | Aqueous solutions | Benzyl alcohol, dimethylformamide | HT Tuffryn | Yes, or limited dry heat | Gelman |
| Polyvinylidene ide | Aqueous solutions with 35% solvents | Acetone, dimethyl-formamide | Durapore Hydrophilic | Yes | Millipore |
| **HYDROPHOBIC** | | | | | |
| Polytetra-fluorethylene with PE substrate | Air and non-aqueous solvents, aqueous solutions if pre-wet with ethanol | Solvent resistant | TF Teflon<br><br>Fluoropore<br><br>TE35<br>Filinert | Yes and dry heat<br><br>Yes and dry heat<br><br>Yes and dry heat | Gelman<br><br>Millipore<br>S and S<br>Nuclepore<br>Pall |
| Polyvinylidene difluoride | Air and aqueous solutions if pre-wet with ethanol | Acetone, dimethyl-formamide | Durapore | Yes | Millipore |

As noted in Table 22.6, membrane filters are usually composed of plastic polymers, including cellulose acetate and nitrate, nylon, polyvinyl chloride, polycarbonate, polysulfone and Teflon. Occasionally, sintered metals such as stainless steel and silver are used when highly durable characteristics are required.

Since most of the membrane filters are disposable, the problem of

cleaning after use is limited to the reusable filter housing and support screen. These are usually made of stainless steel or tough plastic polymers that are cleaned rather easily. Careful attention must be given, however, to disassembly of the housing and scrubbing to remove any residues that might introduce contamination on subsequent use.

The membranes are usually rendered hydrophilic by treatment with a surface active agent at the time of manufacture. If this is not done, particularly at the lower porosities, an aqueous solution cannot be forced through the filter except under very high pressure. When non-wetting with water is desired, however, as with such non-aqueous solvents as ethanol and inert gases, the polymer is left in its hydrophobic form.

**Mechanism of action:** Membrane filters function primarily by sieving, or by screening particles from a solution or gas, thus retaining them on the filter surface. Because of the nature of membrane filters and their limited thickness, there is little entrapment within the filter medium, this being a mechanism applicable to the function of depth filters, such as those made of glass and paper. Membrane filters also function in some instances by electrostatic attraction. This would apply particularly to the filtration of dry gases, in which electrostatic charges tend to increase because of the frictional effect of the flowing gas.

The pores, or holes, through any filter medium consist of a range of sizes. For example, if a filter is designated as 0.2 μm, porosity, the porosity most commonly used to effect sterilization, the maximum mean pore diameter is 0.2 μm, with many pores much smaller than this and a few larger. The latter may have diameters as large as 0.5 μm, but they are so few in number that the probability of a microbial spore (commonly rated as being 0.5 μm in diameter) finding those few pores is highly remote. However, it must be recognized that there is a probability of this happening, even though remote. Therefore, it is no longer acceptable to consider such filters an absolute means of sterilizing a solution. To increase the probability of achieving a sterile filtrate, some researchers are proposing that the solution be passed through a series of two 0.2 μm porosity filters. Others have suggested that a 0.1 μm porosity filter be used, but this would greatly reduce the flow rate.

Since membrane filters function primarily by sieving, particles of any kind in a solution are retained on the surface. If the content is relatively high, particles may accumulate on the surface and plug the filter so that the flow of

solution decreases and perhaps stops. To avoid this problem, when solutions have a high content of solids, particularly when the solids are deformable macromolecules, the solution can best be processed by passing it through one or more pre-filters, the first usually being a relatively porous depth filter. With depth filters, particles may gradually migrate through the filter if filtration time is prolonged, if there is a high pressure differential, or if there is frequent fluctuation of the pressure.

**Liquid flow through a filter:** The flow rate of a liquid through a filter is affected by the size of the pores through the filter, the pore volume (the proportion of open space to solid matrix), the surface area of the filter, the pressure differential across the filter, and the viscosity of the liquid. Of these factors, the two most practical ways to increase flow rate are to increase the surface area of the filter or the pressure differential across the filter. There is a practical limit to increasing the diameter of a disc filter; thus, if larger surface areas are required, a pleated filter in a cartridge form is often used. In this way, a large increase in surface area may be achieved within a relatively small overall dimension of the filter unit. Within the limits of the physical strength of the filter and its housing, the pressure differential can be increased to several hundred pounds per square inch. In pharmaceutical practice, however, the pressure differential used is rarely more than 25 to 30 pounds per square inch. Usually, a positive pressure is applied on the liquid upstream of the filter, but a vacuum may be drawn downstream of the filter. In the case of a vacuum, the maximum differential achievable is one atmosphere, or approximately 15 psig. Furthermore, the negative pressure in the filtrate chamber makes it difficult to prevent the ingress of contamination from the environment. Therefore, for filtrations designed to render solutions sterile, it is preferable to apply pressure upstream of the filter using a gas filtered to be free from microorganisms. Any leakage that may occur in such a system causes loss to the outside without contamination of the sterile filtrate.

Solutions having a high viscosity normally have a slow flow rate. In most instances, the rate can be increased by warming the solution, thereby reducing its viscosity, provided the warming does not have an adverse affect on the solution.

As previously mentioned, the flow rate through a filter also depends on the relative pore volume of the filter. All filters must have a solid matrix that forms the framework for the pores. The lower the amount of solid matrix is in

proportion to the pore spaces, the higher are the pore volume and the flow rate.

**Types of filters:** Since the filter membranes are designed to be used once and then discarded, they are disposable; further, filter housings composed of plastic polymers, which are also intended to be disposable, are becoming increasingly available. Thus, all after-use cleaning is eliminated. In addition, the membrane filter is sealed into the housing by the manufacturer, so that the risk of leakage is minimal.

Membrane filters are usually in the form of discs or pleated cylinders (cartridges). They range from 13-mm discs (approximately 0.8 $cm^2$) to 20-in or longer cartridges (approximately 0.84 $m^2$). The housings are usually of stainless steel or of various plastic polymers.

A few years ago, it was a rather common practice to use filters that were reusable, such as diatomaceous earth, sintered glass and unglazed porcelain. Because of the problems of adequate cleaning between uses and testing, current applications of these filters are limited.

**Aseptic processing:** Sterilization of a solution by filtration provides an extremely clean solution, removing dirt particles as well as microorganisms in the micron size range. After sterilization, however, the filtrate must be transferred from the receiver and subdivided into the individual final containers. The objective of this process, known as aseptic processing, is to exclude every microorganism from all steps of the process subsequent to filtration. Accomplishing this requires a rigidly-controlled aseptic environment and technique. The difficulty in maintaining such an aseptic condition is the greatest problem associated with sterilization by filtration; however, for solutions that are adversely affected by heat, this may be the only way in which sterilization can be accomplished. Aseptic processing is technically not a sterilization process, but is mentioned here because of its close involvement with sterilization by filtration. It is used for products that cannot be terminally sterilized, that is, sterilized after they have been sealed in the final container (*see* Chapter 23).

**Applications:** Filtration is used for nonterminal sterilization and has to be employed under strict aseptic conditions. It is employed for those pharmaceuticals which can not be sterilized by terminal processes, or to which agents like additives, heparin and vitamins etc. are added post-sterilization. It is used to sterilize the thermolabile pharmaceuticals, aqueous

liquids, oils, organic solutions, and air and other gases.

## Chemical Processes of Sterilization

### *Gas Sterilization*

Gas sterilization is not new. Such gases as formaldehyde and sulfur dioxide have been used for sterilization for many years. These gases are highly reactive chemicals, however, and are difficult to remove from many materials after exposure. Therefore, their usefulness is limited. Two newer gases, ethylene oxide and β-propiolactone, have fewer disadvantages than the older agents and therefore have assumed importance in sterilization. Undoubtedly, the advent of plastic materials and the need for a practical method of sterilizing them have spurred the development of the newer gaseous sterilizing agents, particularly ethylene oxide. The chemical biocides are generally classified into alkylating and oxidizing agents.

### *Alkylating Gases*

These include ethylene oxide, β-propiolactone (BPL), propylene oxide, glutaraldehyde and formaldehyde. Out of these, ethylene oxide is the most widely used and discussed in detail here. Propylene oxide is not widely used as it requires a long sterilization cycle. Glutaraldehyde is used in solution form for the sterilization of surgical instruments. Formaldehyde is a surface sterilant, and can not be used to sterilize occluded areas. The uses of low temperature steam formaldehyde (LTSF) sterilization have been discussed earlier.

**Mechanism of action:** Alkylating gases are believed to exert their lethal effect upon microorganisms by alkylating essential metabolites, affecting particularly the reproductive process. Alkylation probably occurs by replacing active hydrogen on sulfhydryl-, amino-, carboxyl-, or hydroxyl-groups with a hydroxyediyl radical. The altered metabolites are not available to the microorganism and so it dies without reproducing.

**Ethylene oxide:** Ethylene oxide (EtO) is a cyclic ether ($[CH_2]_2O$) and is a gas at room temperature. Alone, it is highly flammable, and when mixed with air, explosive. Admixed with inert gases such as carbon dioxide, or one or more of the fluorinated hydrocarbons (Freons) in certain proportions, ethylene oxide is rendered non-flammable and safe to handle. As a gas, it penetrates readily such materials as plastic, paperboard and powder. EtO

dissipates from the materials simply by exposure to the air. It is chemically inert towards most solid materials. On the other hand, in the liquid state, as compressed in cylinders, EtO dissolves certain plastic and rubber materials and requires particular care in handling.

**Sterilization process:** Sterilization with EtO involves a carefully-validated procedure using a pressure chamber. The material to be sterilized is placed in a room or chamber and exposed to a relative humidity of up to 98% for a period of 60 min or longer. It is then placed in the chamber, previously heated to about 55°C (131°F), and an initial vacuum of approximately 27 in. Hg is drawn. The EtO is then introduced, along with moisture, to achieve a relative humidity of 50 to 60%, to the pressure required to give the desired concentration of EtO (Table 22.7), which is maintained throughout the exposure period. Following the exposure period of 6 to 24 h, depending on the degree of contamination, the penetrability of the material, and the concentration of EtO, the gas is exhausted, and a vacuum of approximately 25 inches Hg is drawn. Filtered air is then introduced into the chamber until atmospheric pressure is attained.

A heated chamber is used to decrease the time required for this sterilization process. A temperature of 55°C (131°F) has no adverse effect on most substances. It has been suggested that a rise in temperature of 17°C permits the shortening of the exposure period by about one-half.

Moisture has been found to exert a significant effect on the sterilization process, as well although reports have varied greatly with respect to the conditions and amount of moisture that are essential. It appears that a relative humidity (RH) of 30% or more is essential for effective antibacterial activity. Studies have shown that microorganisms must first be hydrated if they are to be killed by ethylene oxide within the usual cycle time. If significantly dehydrated previously, their rehydration may require several days exposure to relative humidities of 75% or more. Moisture introduced into the sterilizing chamber along with the gas may not adequately hydrate the microorganism as the moisture must first be absorbed by the surrounding material and then penetrate the microorganism. Therefore, a moisturizing dwell period at up to 95% RH should be the first step in every sterilizing cycle as an aid in the distribution and absorption of moisture by the material to be sterilized. The dwell period also aids in establishing a moisture equilibrium within the chamber load, particularly for materials that

preferentially absorb moisture.

The exposure conditions most frequently used with EtO are shown in Table 22.7. Note that concentrations higher than the minimum effective concentration of 450 mg per liter of chamber volume reduce the exposure period. The concentrations employed are directly related to the pressure of the various mixtures required to attain that concentration. Available equipment may limit the pressure, and thereby the concentration attainable.

**Table 22.7:** Exposure conditions used with ethylene oxide mixtures at a temperature of 55°C (131°F)*

| Commercial name | Mixture content (%) | Ethylene oxide concentration (mg/L) | Chamber pressure (Psig) | Minimum exposure period (h) |
|---|---|---|---|---|
| Carboxide | 10 Ethylene oxide<br>90 Carbon dioxide | 450 | 28 | 6 |
| Oxyfume-20 | 20 Ethylene oxide<br>80 Carbon dioxide | 670<br>920 | 18<br>30 | 4<br>3 |
| Cry-Oxcide (Benvicide) | 11 Ethylene oxide<br>54 Trichlorofluorornethane<br>35 Dichlorodifluoromethane | 450<br>850 | 5<br>18 | 5<br>3 |
| Pennoxide | 12 Ethylene oxide<br>88 Dichlorodifluoromethane | 650 | 7 | 4 |

* Following a humidifying (60% relative humidity) dwell period of 60 min.

In addition, note that liquid EtO is frequently used instead of the mixtures with inert gases, thereby eliminating the necessity for high-pressure handling equipment. The liquid is usually vapourized into the sterilizing chamber previously evacuated to at least 720 mm (28 in.) Hg. In the absence of oxygen (in a vacuum) and a spark, there is no danger of explosion with EtO.

Normally, the dissipation of EtO from materials is accomplished readily at the end of a sterilizing cycle by evacuation followed by a short period of aeration, that is, exposure to the normal atmosphere. It has been found, however, that certain materials-notably rubber, certain plastics, and leather have a strong affinity for ethylene oxide and may require prolonged aeration, as long as 12 to 24 h, before items made from these materials may safely be used. Tissue irritation may result if the ethylene oxide is not entirely dissipated. Concern also exists for the carcinogenic and mutagenic properties of EtO and residues in materials for human use. Both the Occupational Safety and Health Administration (OSHA) and the FDA have been studying this

matter. As a consequence, OSHA has recently established an occupational exposure standard of 1 ppm of EtO as an 8-hour time-weighted average concentration.

**β-propiolactone:** β-propiolactone ($[CH_2]_2O$ CO) (BPL) is a cyclic lactone and is a nonflammable liquid at room temperature. It has a low vapour pressure, but since it is bactericidal against a wide variety of microorganisms at relatively low concentrations, no difficulty is experienced in obtaining bactericidal concentrations of the vapour. Studies have indicated that vapour concentrations of approximately 2 to 4 mg per liter of space are effective at a temperature not below 24°C and a relative humidity of at least 70%, with an exposure period of at least 2 h.

**Application:** Alkylation may also occur with drug molecules in pharmaceutical preparations, particularly in the liquid state. Therefore, EtO sterilization of pharmaceuticals is limited essentially to dry powders of substances shown to be unaffected. It has an extensive application, however, to plastic materials, rubber goods, and delicate optical instruments. It has also been found that stainless steel equipment has a longer useful life when sterilized with EtO instead of steam. The effective penetrability of EtO makes it possible to sterilize parenteral administration sets, hypodermic needles, plastic syringes, and numerous other related materials enclosed in distribution packages of paperboard or plastic. Although the cycle time for sterilization with EtO is quite long and certain problems contributing to sterilization failures have yet to be elucidated, this method of sterilization has made it possible to sterilize many materials that would be virtually impossible to sterilize with other known methods. The penetrability of BPL vapour has been found to be poor, therefore, its principal use appears to be the sterilization of surfaces in large spaces, such as entire rooms.

## *Oxidizing Gases*

These include hydrogen peroxide, ozone, chlorine dioxide, and peracetic acid. They are relatively unstable and their decomposition can lead to microenvironments within the load that are not exposed to the full concentration of the agent.

**Hydrogen peroxide ($H_2O_2$):** It has been shown to be effective against spores at a range of temperatures. Its action is greatest when used at near-saturation levels on clean dry surfaces and it does not leave a toxic residue.

The vapour is usually obtained by evaporation of heated stock solution. Water condensation on surfaces that are being sterilized can reduce local concentrations of hydrogen peroxide, and hence its activity and decomposition by catalytic activity and absorption by cellulosic materials (i.e. paper) can also reduce its efficacy. The combination of $H_2O_2$ with cold plasma, cupric or ferric ions, ozone or UV has been shown to enhance activity.

**Ozone ($O_3$):** It is an effective sterilant, but the complex control of humidity required and corrosion problems have limited its applications and it is not routinely used. Ozone, as a disinfectant for water, currently shows more promise and commercial systems are available for several applications.

**Chlorine dioxide:** It has sporicidal activity and is currently under commercial development, although its applications are limited due to its effect on materials such as uncoated aluminium foil, uncoated copper, polycarbonate and polyurethane.

**Peracetic acid:** It exists as a liquid at room temperature and is used for high level disinfection. However, vapourized peracetic acid is used to sterilize isolators, although a long contact time is required. Peracetic acid can cause corrosion of certain metals and rubbers and has a low penetrating power.

## *Surface Disinfection*

The use of chemical disinfectants in the pharmaceutical industry is designed primarily to reduce the microbial population so that asepsis can be maintained in a limited, controlled environment. Most disinfectants do not destroy spores during any reasonable contact period; therefore, they do not sterilize a surface. However, as adjuncts to thorough cleaning of surfaces, disinfectants when used properly may be expected to provide an aseptic condition of the surfaces involved.

The effectiveness of a disinfectant depends on the nature of the surface, the nature and degree of contamination, and the microbicidal activity of the agent employed. Hard, smooth surfaces are much easier to disinfect than rough porous ones. Since most disinfectants are not effective against spores, only vegetative forms of microorganisms can be expected to be killed. The effectiveness of the agent will depend on the number of organisms present

and their sensitivity to the agent. Therefore, it is essential to select an agent that has been proven effective against the common contaminants.

Hundreds of disinfectants have been made available commercially. Methods of evaluation differ significantly; therefore, comparisons of effectiveness are often quite difficult. In one comparison of 16 commercial disinfectants (identified by active constituents only) at least five were found to be ineffective by the screening method proposed. The two most effective contained a combination of four phenolic compounds.

Spaulding has provided a valuable summary of disinfectants. He recommends a 2% solution of one of the phenolic germicide cleaners for floors and walls, and 1:1000 concentrations of quaternary ammonium solutions or 1 to 2% solutions of phenolic germicides for smooth, hard surfaces. If the object is metallic, he recommends that 0.2% sodium nitrite be added to the quaternary ammonium solutions and 0.5% sodium bicarbonate to the phenolic germicides to prevent rusting. Higher concentrations of disinfectants normally would be expected to be more effectively bactericidal; however, concentrations are limited by the detrimental effect that most of these solutions have on the surfaces to which they are applied. Chemical attack may be evidenced by such effects as pitting and rusting of metal surfaces and cracking and discolouration of painted surfaces.

**Table 22.8:** Examples of chemical indicators for monitoring sterilization processes

| Sterilization method | Principle | Device | Parameter(s) monitored |
|---|---|---|---|
| Heat Autoclaving dry heat | Temperature-sensitive coloured solution | Sealed tubes partly filled with a solution which changes colour at elevated temperatures rate of colour change is proportional to temperature, e.g. Browne's tubes | Temperature, time |
| Dry heat only | Temperature-sensitive chemical | Usually a temperature-sensitive white wax concealing a black marked printed (paper) surface; at a predetermined temperature the wax rapidly melts exposing the background mark(s) | Temperature |
| Heating in an autoclave only | Steam-sensitive chemical | Usually an organic chemical in a printing ink base impregnated into a carrier material A combination of moisture and heat produces a darkening of the ink, e.g. autoclave tape. Devices of this sort can be used within dressing packs to confirm adequate removal of air and penetration of saturated steam Bowie-Dick test) | Saturated steam |
| | Capillary principle | Consists of a blue dye in a waxy pellet, the melting-point of which is depressed in the presence of saturated steam. At autoclaving temperatures and in the continued presence of steam, the pellet melts and travels along a paper wick forming a blue band the length of which is dependent upon both exposure time and temperature | Temperature saturated steam time |
| Gaseous sterilization | | | |
| Ethylene oxide | Reactive chemical | Indicator paper impregnated with a reactive chemical which undergoes a distinct colour change on reaction with EO in the presence of heat and moisture. With some devices rate of colour development varies with temperature and EO concentration | Gas concentration temperature time (selected device). NB a minimum relative humidity (RH) is required for device to function |
| | Capillary principle | Based on the same migration along wick principle as Thermalog S. Optimum response in a cycle of 600 mg/L EO, temperature 54°C, rh 40–80%. Lower EO levels and/or temperature will slow response time | Gas concentration temperature, time (selected cycles) |
| Low temperature steam and formaldehyde | Reactive chemical | Indicator paper impregnated with a formaldehyde, steam and temperature-sensitive reactive chemical which changes colour during the sterilization process | Gas concentration temperature, time (selected cycles) |
| Radiation sterilization | Radiochromic | Plastic devices impregnated with radiosensitive chemicals which undergo colour changes at relatively low radiation doses | Only indicate exposure to radiation |
| | Dosimeter device | Acidified ferric ammonium sulphate or ceric sulphate solutions respond to irradiation by dose-related changes in their optical density | Accurately measure radiation doses |

Selected disinfectants, particularly glutaraldehyde and BPL have been found to improve the effectiveness of such physical methods of sterilization as ultraviolet light and ultrasonics. Other select combinations of chemical and physical sterilizing agents have shown increased reliability as compared to either agent alone.

## Evaluation of Sterilization Methods

The effectiveness of each sterilization process must be demonstrated prior to its use under processing conditions. A thorough evaluation must be carried out, both of the functional capabilities of the equipment and of the process methods under the most demanding conditions of operation. Only when proved to be consistently effective, can a particular procedure be considered a valid sterilization process. In addition, at frequent intervals during use, the equipment and methods should be re-evaluated to ensure that they are functioning properly.

### *Thermal Methods*

The duplication of proven thermal methods of sterilization cannot be taken for granted. Mechanical equipment as well as personnel are subject to failure. Therefore, indicators should be used as a check on the duplication of the conditions of a proven (validated) process, locating the indicator where there is greatest impediment to the penetration of heat.

Among the indicators available, the most widely used is the *thermocouple.* These indicators are often connected to recorders so that a continuous record of the actual temperature at the location of the thermocouple can be obtained.

For autoclave sterilization, a variety of other indicators are also used. These include *wax* or *chemical pellets* that melt at 121°C and *paper strips* that are impregnated with chemicals that change colour under the influence of moisture and heat. All of these have limited reliability in indicating the length of time for which a temperature of 121°C has been maintained.

Resistant bacterial spores in sealed ampoules or impregnated in dry paper strips are used as biologic indicators. Their destruction is an evidence of the intended effect of a sterilization process. Their use to prove the effectiveness of new sterilizing equipment or processes is widely accepted, but their use as indicators for routine process control is questioned by some. Among the concerns are (1) lot-to-lot variability in the resistance of spores, (2) lot-to-lot variability in the number of viable spores, (3) difficulty in obtaining pure cultures and (4) inherent danger of placing viable spores in a sterilizer load of materials for human use.

## *Non-thermal Methods*

### *Filtration Sterilization*

Although membrane filters are tested and labeled by the manufacturer, the pore size and integrity of the filter should be checked before use. The least complicated method for doing this is the bubble-point test. This test is performed by applying air pressure, or other gas pressure, to the upstream side of a hydrophilic filter in which the pores are filled with water. The pressure is gradually increased until bubbles pass through the filter and are detected in a liquid downstream. This bubble-point pressure is inversely proportional to the diameter of the pores, and thus is a measure of the largest pores. The filter manufacturer identifies the appropriate test pressure for each pore size, for example, 55 Psig for a hydrophilic membrane filter of 0.2 μm porosity as given in Table 22.6. If there is even a pinhole or similar defect in the filter, bubbling occurs at a much lower pressure than expected. For hydrophobic membranes, the filter is usually wetted with ethanol or methanol prior to application of the air pressure.

For cartridge-type filters, it is practical to measure the diffusion of air, or other gas, through the water-filled pores of the filter medium because of the large surface area. Pressure is applied to the upstream side of the filter as specified by the manufacturer, at approximately 10% of the bubble test pressure. The air dissolves in the water in the pores of the membrane and is released from the downstream side of the membrane at a rate that is directly related to the pore size. This rate is measured by the volume of the air collected downstream or by the loss of pressure from the upstream side as the air diffuses.

### *Biological Indicators*

USP defines a biological indicator (BI) as a characterized preparation of a specific microorganism that provides a defined and stable resistance to a specific sterilization process. Microorganisms widely recognized as suitable for biological indicators are spore-forming bacteria, because, with the exception of ionizing-radiation processes, these microorganisms are significantly more resistant than normal microflora. A BI can be used to assist in the performance qualification of the sterilization equipment and in the development and establishment of a validated sterilization process for a

particular article. BIs are used in processes that render a product sterile in its final package or container, as well as for the sterilization of equipment, materials, and packaging components used in aseptic processing. They may also be used to monitor established sterilization cycles and in periodic revalidation of sterilization processes. They may also be used to evaluate the capability of processes used to decontaminate isolators or aseptic clean-room environments. BIs used in sterilization method are listed in Table 22.9.

**Table 22.9:** Biological indicators used in sterilization methods

| **Sterilization methods** | **Biological indicators** |
|---|---|
| Moist heat | Spores of *Bacillus stearothermophilus*, *Bacillus subtilis*, *Bacillus coagulans* and *Clostridium sporogenes* |
| Dry heat | Spores of *Bacillus subtilis* var *niger* |
| Ionizing radiation Filtration | Spores of *Bacillus pumilus Pseudomonas diminuta* also known as *Brevundimonas diminuta* |
| Ethylene oxide | Spores of *Bacillus subtilis* var *niger* |
| Hydrogen peroxide and peracetic acid | Spores of *Bacillus stearothermophilus*, *Bacillus subtilis*, *Bacillus coagulans* and *Clostridium sporogenes* |
| Formaldehyde | *Bacillus subtilis* var *niger* |

**Types of biological indicators:** According to USP, there are at least three types of BIs. Each type of indicator incorporates a known species of a microorganism of known sterilization resistance to the sterilization mode. Some BIs may also contain two different species and concentrations of micro organisms. One form of BI includes spores that are added to a carrier (a disk or strip of filter paper, glass, plastic, or other materials) and packaged to maintain the integrity and viability of the inoculated carrier. Carriers and primary packaging shall not contain any contamination (physical, chemical, or microbial) that would adversely affect the performance or the stability characteristics of the BI. The carrier and primary packaging shall not be degraded by the specific sterilization process, which is used in a manner that will affect the performance of the BI. The carrier should withstand transport

in the primary and secondary packaging and handling at the point of use. The design of the carrier and primary packaging should minimize the loss of the original inoculum during transport, handling, and shelf-life storage. Another form of BI is a spore suspension that is inoculated on or into representative units of the product to be sterilized. This represents an inoculated product; however, a simulated inoculated product may be used if it is not practical to inoculate the actual product. A simulated product is a preparation that differs in one or more ways from the actual product, but performs as the actual product using test conditions or during actual production sterilization processing. Spore suspensions with a known *D*-value should be used to inoculate the actual or simulated product. If a simulated inoculated product is used, it must be demonstrated that it will not degrade the sterilization resistance of the bioindicator. The physical design of actual or simulated product can affect the resistance of spore suspensions that are inoculated on or into the products. In the case of liquid inoculated products, it is often advisable to determine both the *D*-value and z value of the specific biological indicator microorganism in the specific liquid product. The population, *D*-value, *Z*-value where applicable, and endpoint kill time of the inoculated actual or simulated product should be determined. A third form of biological indicator is a self-contained indicator. A self-contained biological indicator is designed so that the primary package, intended for incubation following sterilization processing, contains the growth medium for recovery of the process-exposed microorganisms. This form of biological indicator, together with the self-contained growth medium can be considered a system. In the case of self-contained biological indicators, the entire system provides resistance to the sterilization process.

## Limitations of Sterilization Methods

Sterilization processes can involve extreme conditions such as high temperature, and the use of toxic substances, that can damage the product and/or its packaging. Therefore, there needs to be a balance between acceptable sterility assurance and acceptable damage to the product and container. Knowledge of the preparation and packaging design, and the choice of the sterilization techniques help in making the appropriate selection to achieve maximum kill while decreasing the risk of product and packaging deterioration. Nevertheless, each sterilization processes have their own limitation (Table 22.10). Limitations associated with established recommended procedures are usually linked to the nature of the process (e.g. heat, radiation).

**Table 22.10:** Limitations of sterilization methods

| Sterilization Methods | Limitations |
|---|---|
| **Thermal** | Damage to preparation and containers,<br>Potential longer exposure time needed in dry heat method<br>High pressure may causes air ballasting in moist heat method |
| **Radiation** | Risk to the operator<br>Product damage due to water radiolysis<br>Change in potency<br>Discoloration of some glasses and plastics (e.g. PVC) occurs with *X* radiation<br>Hardness and brittle property of metals may also change with *X* radiation<br>Poor penetration and significant product heating at high dose with *X* radiation |
| **Filtration** | Not-effective with small particles (e.g. viruses, prions)<br>In process strict aseptic tech-nique is required<br>Integrity of membrane filter may damage<br>Microbial contaminant growth may occur in depth |

| | |
|---|---|
| | filters<br>Shedding of material from depth filters |
| **Chemical** | High toxicity risk to the operator |
| **Method** | Gases are explosive in nature<br>Decontamination required post process<br>Slow process<br>Damage to some cellulosic material with formaldehyde<br>Microbial resistance is reported with gluteraldehyde and peracetic acid<br>Corrosion occurs with peracetic acid |

## New Technologies

### *Ultrahigh Pressure*

Vegetative microorganisms and bacterial spores are found to be inactivated at pressures above 100 MPa and 1200 MPa, respectively. The use of high pressure for food preservation has been combined with the chemical effect of preservative such as low pH. The low pH ensures the prevention of growth of bacterial spores. The use of ultrahigh pressure has now been applied to food industry and such technology has been evaluated for the production of bacterial vaccines. High pressure tends to denature preferably macromolecules, i.e. enzymes, membranes, genomic materials of microbial cells, and lead to death of microorganism. The advantages of this process is that the quality and taste of product tend not to be affected by the system.

### *High-intensity Light Pulse (HILP)*

The application of intense light such as high-intensity laser has been known to kill microorganisms, principally through a combination of UV radiation and heat treatment. Such processes have been shown to inactivate vegetative microorganisms and bacterial spores. HILP has a potential application in terminal sterilization of clean solutions, such as water, saline, dextrose and opthalmic pharmaceuticals. The type of container used as a packaging material is of prime importance since it must not shield the transmission of light.

### *Gas Plasma*

Gas plasma is generated by the application of a strong magnetic field to a compound in gaseous phase (e.g. $H_2O_2$). This process create free radicals, that can damage cellular components (e.g. membrane, nucleic acid), a mechanism of action similar to that of oxidizing agents. Gas plasma, which is dry sterilization process, is effective against vegetative microorganisms and bacterial spores. As an additional advantage, the method being non thermal is applicable to thermolabile pharmaceuticals.

# 23:
# Sterile Products

Sterile products are dosage forms of therapeutic agents that are free of viable microorganisms. Principally, these include parenteral, ophthalmic, and irrigating preparations. Of these, parenteral products are unique among dosage forms of drugs because they are injected through the skin or mucous membranes into internal body compartments. Thus, because they have circumvented the highly efficient first line of body defense, i.e. the skin and mucous membranes, they must be free from microbial contamination and toxic components, as well as possess an exceptionally high level of purity. All components and processes involved in the preparation of these products must be selected and designed to eliminate, as much as possible, contamination of all types, whether of physical, chemical, or microbiologic origin.

Parenteral preparations may be given by various routes: intravenous, intramuscular, subcutaneous, intradermal and intraperitoneal (Fig. 23.1).

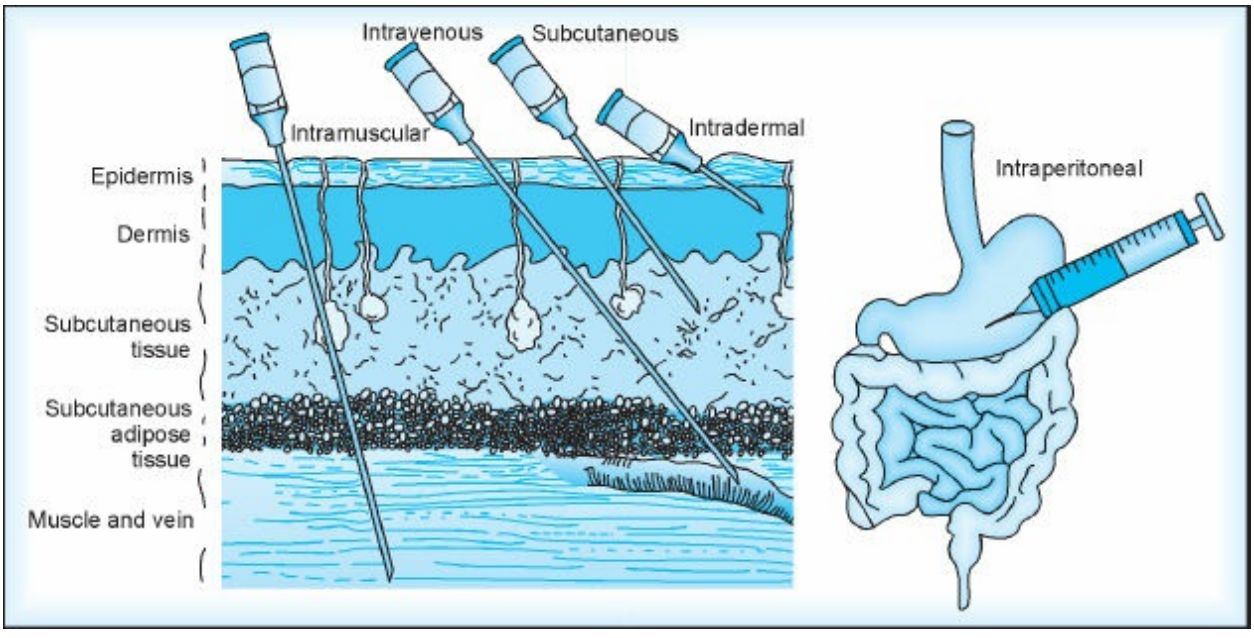



**Fig. 23.1:** Parenteral routes of drug administration

When injection occurs via an intravascular route, complete drug availability occurs immediately; no absorption is necessary. For all other routes, at least a blood vessel wall, and usually one or more tissue cell walls, must be permeated before the drug can enter the circulation. Most often, this occurs by passive diffusion and is most favourable when the drug has both lipophilic and hydrophilic properties, with the former being predominant. With non-vascular injections, absorption is also affected by such factors as the size and number of blood vessels supplying the tissue, the movement (exercise) of the tissue following injection, the physical and chemical properties of the drug and such characteristics of the dosage form as whether it is a solution, suspension, or emulsion, the nature of the vehicle, and its pH. Once in the circulating blood, the physiologic effect of a therapeutic agent is affected by the extent to which it distributes throughout the body, by the degree of binding to plasma proteins and by its rate of elimination by hepatic metabolism and/or renal excretion.

Intravenous and intraspinal preparations are rarely given in a form other than aqueous solutions. The danger of blockage of fine capillaries, particularly in the brain, precludes the use of forms other than solutions for IV administration, although emulsions have been given in which the particle size of the dispersed phase is carefully controlled. The sensitivity of the nerve tissues generally precludes the use of anything but the purest of solutions for intraspinal medication. Preparations given intramuscularly, subcutaneously, or intradermally can be administered as solutions, suspensions, or emulsions. Even solid pellets may be implanted subcutaneously or intramuscularly. The vehicles can range from Water for Injection, to glycols, to fixed oils. Although care must be exercised to avoid undue tissue irritation, mild local irritation is permissible at these injection sites. Details of specialized parenteral routes, volume, types of medication administered and formulation constrains are summarized in Table 23.1.

**Table 23.1:** Parenteral routes of drug administration

| Routes | Usual volume (ml) | Types of medication administered | Formulation constraints |
|---|---|---|---|
| **Primary parental routes** | | | |
| Small-volume parenterals | | | |
| Subcutaneous | 0.52–2 | Insulin, vaccines | Need to be isotonic |
| Intramuscular | 0.5–2 | Nearly all drug classes | Can be solutions, emulsions preferably isotonic |
| Large-volume intravenous parenterals: | 1–1000<br>100 and large (infusion unit) | Nearly all drug classes (*see* precautionary notes in text) | Solutions and some emulsions |
| **Other parenteral routes** | | | |
| Intra-arterial: Directly into an artery (immediate action sought in peripheral area) | 2–20 | Radiopaque media antineoplastics, antibiotics | Solutions and some emulsions |
| Intrathecal, Intraspinal (into spinal canal) | 1–4 | Local anesthetics, analgesics; neurolytic | Must be isotonic |
| Intraepidural (into epidural space near spinal column)<br>Intracisternal (directly into caudal region of the brain between the cerebellum and the medula oblongata) | 6–30 | Local anesthetics, narcotics, $\alpha_2$ agonists, steroids | Must be isotonic |
| Intra-articular (directly into a joint. Usually for a local delivery there, as for steroid anti-inflammatory action in arthritis) | 2–20 | Morphine, local anesthetics, steroids, NSAIDs | Must be isotonic |
| Intracardial directly into the heart, when life is threatened (epinephrinic stimulation in severe heart attack) | 0.2–1 | Cardiotonic drugs, calcium | |
| Intrapleural: Directly into the pleural cavity or a lung (also used for fluid withdrawl) | 2–30 | Local anesthetics, narcotics, chemotherapeutic agents | |
| **Diagnostic testing** | | | |
| Intradermal | 0.05 | Diagnostic agents | Should be isotonic |

The nature of a preparation can influence significantly the rapidity of onset of a therapeutic effect from a drug, the duration of the effect, and the form of the absorption pattern achieved. Therefore, the development of the formulation for a parenteral product must be integrated carefully with its intended administration in a patient.

The chemical and physical properties of a drug must be determined, its interaction with any desired excipients must be studied, and the effect of each

step of the process on its stability must be studied and understood.

Preparations for the eye, though not introduced into internal body cavities, are placed in contact with tissues that are very sensitive to contamination. Therefore, similar standards are required for ophthalmic preparations.

Irrigating solutions are now also required to meet the same standards as parenteral solutions because during an irrigation procedure, substantial amounts of these solutions can enter the bloodstream directly through open blood vessels of wounds or abraded mucous membranes. Therefore, the characteristics and standards presented in this chapter for the production of large-volume parenteral solutions apply equally to irrigating solutions.

## EFFECT OF ROUTE OF ADMINISTRATION

The intended route of administration has a marked effect on the formulation of a parenteral product. The volume in which a dose of the drug must be encompassed is one factor to consider. For intracutaneous injections a volume of more than 0.2 ml rarely is used because tissue volume is small and compact; also, absorption is quite slow owing to the lack of blood vessels. Volumes of 1 ml or less may be injected subcutaneously and only occasionally are volumes of more than 2 ml given intramuscularly. Volumes of 10 ml or less may be given intraspinally, but only by the IV route may large volumes be given safely, provided careful control of the rate of administration is undertaken. It is not convenient to administer a volume of more than 20 ml by a syringe, and usually it is not practical to set up an infusion unit for less than 250 ml.

Isotonicity is a characteristic that is probably of greatest importance for intraspinal injections because the circulation of the cerebrospinal fluid is slow, and disturbances of osmotic pressure quickly cause headache and vomiting. Since intracutaneous injections are given mostly for diagnostic purposes, nonisotonic solutions may cause false signs of irritation. Isotonicity is preferable for the comfort of the patient, but is not essential for SC and IM injections. For the rapid absorption of drugs given intramuscularly, a slightly hypertonic solution may increase the rate by causing local effusion of tissue fluids. Usually, IV fluids should be isotonic, although slow administration of a paratonic solution may be performed safely if rapid dilution with the blood occurs.

In general, only solutions of drugs in water may be given intravenously. Suspensions may not be given because of the danger of blockage of the small blood vessels. Aqueous or oleaginous suspensions and oleaginous solutions cannot normally be given subcutaneously because of the pain and irritation caused. Muscle tissue tolerates oils and suspended particles fairly well and is therefore the only route normally suitable for their administration.

The administration of a drug deep into the muscle tissue results in a pool of the product at the site of injection. From this depot, the drug is released at a rate determined to a large extent by the characteristics of the formulation. Whether the solvent is aqueous or oleaginous affects the rate of absorption; oleaginous solutions are usually more slowly absorbed. Increasing the

viscosity of solutions slows the absorption, just as gelatin or polyvinylpyrrolidone in water and aluminum monostearate in oils. Utilizing modifications of the drug molecule to render it less soluble (for instance, the formation of various esters or salts) permits the production of stable suspensions, causing a marked reduction in the rate of absorption of the drug from the depot. Thus, utilizing various modifications in formulation of the product makes it possible to retard the rate at which a drug is released from a depot.

Ophthalmic preparations are formulated in much the same way as parenteral solutions. The eye is particularly sensitive to irritation; therefore, formulation should be directed towards minimizing irritation. Normally, clean aqueous solutions are preferable for ophthalmic use. Suspensions of solids have been used in the eye when the therapeutic need superseded the need to avoid irritating effects, as for the suspensions of corticosteroids used occasionally. It has been found that a foreign body sensation increases as the concentration of suspended particles, regardless of size, approaches 5%.

This brief discussion of some of the factors involved in the formulation of parenteral dosage forms has been intended simply to introduce the student to this important area. This area is changing steadily as the ingenuity of research pharmacists spurs the development of new and improved formulation aids and techniques.

Sterile products are most frequently solutions or suspensions, but may even be solid pellets for tissue implantation. The control of a process to minimize contamination for a small quantity of such a product can be achieved with relative ease. As the quantity of product increases, the problems of controlling the process to prevent contamination multiply. Therefore, the preparation of sterile products has become a highly specialized area in pharmaceutical processing. The standards established, the attitude of personnel and the process control must be of a superior level.

The organizational divisions normally responsible for the preparation of sterile products in the pharmaceutical industry are product development, production, packaging and control. The treatment of the subject matter in this chapter is in accordance with these four divisions of responsibility, with emphasis on the distinctive aspects of sterile product manufacturing. Particular emphasis is placed on the production division, for it is in this division that the distinctiveness of sterile product processing is particularly

evident.

Requirements for components related to the product formula and its stability are considered in the product development section. Many of the principles of product development, control, and packaging are identical for sterile and non-sterile products. Since some of these principles are treated elsewhere in this text, only those that are distinctive for sterile products are covered in this chapter.

# FORMULATIONS

## Ophthalmic Preparations

Products to be instilled into the eye, while not parenterals by definition, have many similar, and often identical, characteristics. The formulation of stable, therapeutically-active ophthalmic preparations requires high purity of ingredients as well as freedom from chemical, physical (particles), and microbial contaminants. These preparations usually require buffers to stabilize the pH of the product, additives to render it isotonic or nearly so, and stabilizers such as antioxidants when appropriate for the particular ingredients. Those ophthalmics used in larger quantities, such as eye irrigants, or in the case of devices such as contact lenses, are usually relatively uncomplicated solutions similar to large-volume parenterals.

One characteristic not as critical for ophthalmics is freedom from pyrogens since pyrogens are not absorbed systemically from the eye; however, insofar as pyrogens are indicative of a microbiologically clean process, they should not be present.

### *Freeze-dried Products*

Solutions intended to be freeze-dried must be aqueous, for the drying process involves the removal of water by sublimation. Since the solution is in existence for only a brief period during processing, stability problems related to the aqueous system are practically nonexistent. However, the formulation must reflect the characteristics to be imparted to the solid residue (cake) after drying, and those required of the solution after reconstitution at the time of use. Often, the drug alone does not give sufficient solid residue or the characteristics appropriate for the product; therefore, substances often must be added to provide the characteristics desired. Among the characteristics required of a good cake are (1) a uniform colour and texture, (2) a supporting matrix of solids sufficient to maintain essentially the original volume after drying and (3) sufficient strength to prevent crumbling during storage. In addition, the nature and amount of solids in the solution largely determine (1) the eutectic temperature of the frozen solution, the subzero temperature at which the frozen material will melt, which determines the temperature below which the product must be held during freeze-drying, (2) the rate of thermal and vapour transfer through the product during the process of drying and (3)

the rate of solution of the product during reconstitution.

The percentage of solids in the frozen plug should be between approximately 2 and 25%. Among the best salts for providing uniform crystal size, uniform colour and texture, physical strength, and rapid reconstitution are the monobasic and dibasic sodium phosphates. Sodium chloride is often used, but when used alone, the cake tends to shrink markedly in volume and to appear crusty and crumbly. When organic substances, such as mannitol, sorbitol, sucrose, and gelatin are used to provide solids for the cake, care must be taken during the heating, particularly during the terminal stages of drying, to avoid discolouration of the cake by charring. Added substances required in the formulation must not be volatile under the conditions of drying; therefore, antibacterial agents such as phenol, chlorobutanol, and benzyl alcohol should not be used.

## *Long-acting Formulations*

Long-acting parenteral drug formulations are designed, ideally, to provide slow, constant, and sustained release of a drug over a prolonged period of time, essentially to simulate and replace the more hazardous, continuous i.v. infusion of a drug. Rarely, if ever, is the ideal achieved, but extensive research has resulted in depot dosage forms that approach the desired goal.

In one type of depot formulation, which is referred to as “dissolution-controlled,” the rate of drug absorption is controlled by the slow dissolution of drug particles, with subsequent release to tissue fluid surrounding the bolus of product in the tissue. The formation of drug salts with very low aqueous solubility is one of the most common approaches to this type of formulation. Control of the particle size also can contribute to slow dissolution in that larger particles or crystals dissolve more slowly than small crystals with proportionately more surface area. Further, the suspension of the drug particles in vegetable oils, and especially if gelled with substances such as aluminum monostearate, produces prolonged absorption rates.

Another type of depot formulation is produced by the binding of drug molecules to adsorbents. Only the free portion, in equilibrium with that which is bound, can be absorbed. As drug is absorbed, a shift in equilibrium is established, and the drug is slowly released from the bound state to the free state. This is particularly exemplified by the binding of vaccines to aluminum hydroxide gel to provide a sustained release. A third type of depot

preparation is the encapsulation type, in which biodegradable or bioabsorbable macromolecules such as gelatin, phospholipids, and long-chain fatty acids become a diffusion matrix for the drug. The drug is encapsulated within the matrix, and release of drug molecules is controlled by the rate of permeation out of the diffusion barrier and by the rate of biodegradation of the barrier macromolecules. A fourth type is the esterification type depot preparation, in which esters of a drug that are bioerodible are synthesized. The esterified drug is deposited in tissue at the site of injection to form a reservoir of drug. The rate of drug absorption is controlled by the partitioning of the drug esters from the reservoir to tissue fluid and by the rate at which the drug ester regenerates the active drug molecule. Often, these esters are dissolved or suspended in oleaginous vehicles, which further slow the release.

**Suspensions:** The solids content of parenteral suspensions usually ranges between 0.5 and 5%, but may go as high as 30% in some antibiotic preparations. The amount of solids and the nature of the vehicle determine the viscosity of the product, an important factor because of syringeability, the facility with which the product is passed in and out of a syringe. The property of thixotropy is sometimes utilized, particularly with oleaginous suspensions, to provide the sedimentation stability of a gelled preparation during storage and the syringeability of a fluid at the time of administration.

Probably the most important requirement for parenteral suspensions is a small and uniform particle size. Various techniques are available for the reduction of particles, including dry or wet ball milling, micropulverization, fluid energy grinding, ultrasonic insonation of shock-cooled saturated solutions, and spray drying. Small, uniform particles are required to give slow, uniform rates of sedimentation and predictable rates of dissolution and drug release. Also, uniform particle size reduces the tendency for larger crystal growth during storage, since it has been found that relatively small crystals often tend to disappear and large crystals grow larger in a mixture. Such a change can cause caking of a suspension, difficult syringeability because of the large particles and changes in the dissolution and drug release rate following injection.

The stabilization of a suspension for the period between manufacture and use presents a number of problems. As indicated, solids gradually settle and may cake, causing difficulty in redispersion prior to use. Surface active

agents may aid in the preparation and stabilization of a suspension by reducing the interfacial tension between the particles and the vehicle. Polysorbate 80, lecithin, Emulphor EL-620 and Pluronic F-68 are among the surface active agents that have been used in parenteral suspensions. The concurrent addition of a hydrocolloid, such as sodium carboxymethylcellulose, may enhance the effect of the surfactant and cause loss of surface charge of the dispersed particles, water repellency, and the tendency to agglomerate. The following is an example of such a formulation:

| | |
|---|---|
| Cortisone acetate, microfine | 25 mg |
| Polysorbate 80 (surface active agent) | 4 mg |
| Sodium CMC (protective colloid) | 5 mg |
| Sodium chloride (for tonicity effect) | 9 mg |
| Benzyl alcohol (antibacterial) | 9 mg |
| Water for Injection, to make | 1 ml |

Among other protective colloids that have been employed are acacia, gelatin, methylcellulose, and polyvinylpyrrolidone.

Occasionally, parenteral suspensions may be improved by a slight increase in viscosity, either by increasing the amount of protective colloid or by adding a compound such as sorbitol. In other formulations, it has been found that flocculation of the suspended particles has been necessary to prevent packing to a dense cake. The addition of selected ions that increase the surface charge of the solid particles may cause them to form fluffy aggregates. These settle rapidly, but to a large sedimentation volume, which can easily be re-dispersed. Monosodium citrate has been used effectively for such a purpose.

**Emulsions:** The principal problem in the formulation of parenteral emulsions is the attainment and maintenance of uniform oil droplets of 1 to 5 μm in size as the internal phase. With emulsions, separation of the phase does not occur as readily as with suspensions because the difference in density between the oil and water is relatively small. One such product, an emulsion of a natural vitamin $K_1$, has been stabilized with lecithin.

Intravenous nutrient emulsions that have been made contain, for example, 15% cottonseed oil, 4% dextrose, 1.2% lecithin, and 0.3% of an

oxyethyleneoxypropylene polymer, the latter two ingredients being the emulsifiers. The dispersed phase should have droplet sizes of less than 1 μm. The emulsion must be stable to autoclaving. Elevated temperatures, however, tend to produce coalescence of the dispersed phase, and excessive shaking cause acceleration of the rate of creaming. Small amounts of gelatin, dextran and methylcellulose have been found to aid in stabilizing the emulsions, but they are also adversely affected by elevated temperatures.

The preparation of a parenteral emulsion is troublesome. It is made more difficult by the rigid requirement for particle size control to prevent emboli in blood vessels, by the limited choice of emulsifiers and stabilizers of low toxicity, and by the preservation of the oil phase against the development of rancidity.

## FORMULATION DEVELOPMENT

The final objective in the development of a sterile product is the elicitation of a therapeutic effect in a patient. The formulation of a sterile product involves the combination of one or more ingredients with a medicinal agent to enhance the convenience, acceptability, or effectiveness of the product. Rarely is it preferable to dispense a drug singly as a sterile dry powder unless the formulation of a stable liquid preparation is not possible. Emphasis on the details of physical or chemical factors must not be allowed to take precedence over the prime consideration, the use of the product in a patient.

## Therapeutic Agent

A therapeutic agent is a chemical compound subject to the physical and chemical reactions characteristic of the class of compounds to which it belongs. Therefore, a careful evaluation must be made of every combination of two or more ingredients to ascertain whether or not adverse interactions occur, and if they do, of ways to modify the formulation so that the reactions are eliminated or minimized. The formulation of sterile products is challenging, with respect to the knowledge and ingenuity of the persons responsible.

The amount of information available to the formulator concerning the physical and chemical properties of a therapeutic agent, particularly if it is a new compound, is often quite meager. Information concerning basic properties must be obtained, including molecular weight, solubility, purity, colligative properties, and chemical reactivity, before an intelligent approach to formulation can begin. Improvements in formulation are a continuing process, since important properties of a drug or of the total formulation may not become evident until the product has been stored or used for a prolonged time. However, because of the extensive test documentation required by the US Food and Drug Administration (FDA), only outstanding formulations can be justified for continuance to the state of a marketed product.

## Vehicles or Solvent System

### *Aqueous Systems*

By far, the most frequently employed vehicle for sterile products is water, since it is the vehicle for all natural body fluids. The superior quality required for such use is described in the monograph on Water for Injection in the USP. Requirements may be even more stringent for some products, however.

One of the most inclusive tests for the quality of water is the total solids content, a gravimetric evaluation of the dissociated and undissociated organic and inorganic substances present in water. However, a less timeconsuming test, the electrolytic measurement of conductivity of water, is the one most frequently used. Instantaneous measurements can be obtained by immersing electrodes in water and measuring the specific conductance, a measurement that depends on the ionic content in water. The conductance may be expressed by the meter scale as conductivity in micromhos, resistance in megohms, or ionic content as parts per million (ppm) of sodium chloride. The validity of this measurement as an indication of purity of water is inferential in that methods of producing high-purity water, such as distillation and reverse osmosis, can be expected to remove undissociated substances along with those that are dissociated. Undissociated substances such as pyrogens, however, could be present in the absence of ions and not be disclosed by the test. Therefore, for contaminants other than ions, additional tests should be performed.

Additional tests for quality of water for Injection with permitted limits are described in the USP monographs. When comparing the total solids permitted for Water for Injection with that for Sterile Water for Injection, one will note that considerably higher values are permitted for Sterile Water for Injection. This is necessary because the latter product has been sterilized, usually by a thermal method, in a container that had dissolved to some extent in the water. Therefore, the solids content will be greater than for the non-sterilized product. On the other hand, the 10 ppm total solids officially permitted for Water for Injection may be much too high when used as the vehicle for many products. Water shall contain a minimal amount of organic compounds. Such compounds are undesirable for two main reasons: they may be toxic, and/or they may serve as sources of nutrition for microorganisms. In practice, Water for Injection normally should not have a conductivity of more than 1 

micromho (1 megohm, approximately 0.1 ppm NaCl) and total organic carbon (TOC) not more than 500 ppm.

## *Non-aqueous and Mixed Solvents*

In the formulation of sterile pharmaceutical products, it is sometimes necessary to eliminate water entirely or in part from the vehicle, primarily because of solubility factors or hydrolytic reactions. A non-aqueous solvent must be selected with great care for it must not be irritating, toxic, or sensitizing, and it must not exert an adverse effect on the ingredients of the formulation. The screening of such a solvent must therefore include an evaluation of its physical properties, such as density, viscosity, miscibility and polarity, as well as its stability, solvent activity, and toxicity. Solvents that are miscible with water, and that are usually used in combination with water as the vehicle, include dioxolanes, dimethylacetamide, N-(*β*-hydroxyediyl)-lactamide, butylene glycol, polyethylene glycol 400 and 600, propylene glycol, glycerin, and ethyl alcohol. Water-immiscible solvents include fixed oils, ediyl oleate, isopropyl myristate, and benzyl benzoate. The most frequently used non-aqueous solvents are polyediylene glycol, propylene glycol, and fixed oils.

**Solvent selection:** A parenteral therapeutic agent is given by preference as a solution. If aqueous, the solution is physiologically compatible with body tissues, and the biologic response elicited should be reasonably predictable.

The high dielectric constant of water makes it possible to dissolve ionizable electrolytes, and its hydrogen bonding potential brings about the solution of such organic substances as alcohols, aldehydes, ketones, and amines. Conversely, water is a poor solvent for nonpolar compounds, such as alkaloidal bases, which require non-polar solvents. Since therapeutically active compounds given by injection range in property from highly polar to non-polar, solvents having complementary properties must be employed if a solution is to be achieved.

Adding to the complexity of solvent selection, is the requirement that solvents to be injected must be of low toxicity to body tissues. Ether is a solvent for testosterone, but is highly irritating to body tissues and cannot be used alone as a solvent for an injectable preparation. Frequently, the desired solubility can be achieved with mixed solvents, e.g. the use of approximately 40% ethanol in water to solubilize the digitalis glycosides.

Compounds that are dissolved in water are often subject to degradative reactions, such as hydrolysis, oxidation, decarboxylation, and racemization. Formulation must be designed, in such cases, to minimize the degradative effects. Often, these reactions are markedly affected by the pH of the solution. Epinephrine in solution undergoes racemization and oxidation, but if the pH is maintained at 3.0 or less, little reaction occurs. The oxidation reaction can be further reduced by displacing atmospheric oxygen with an inert head space gas and adding 0.1% (w/v) sodium metabisulfite as an antioxidant. Atropine sulfate rapidly hydrolyzes in solution, but if the pH is maintained with a buffer system at about 3.5 to 4.0, hydrolysis does not occur at a significant rate.

The use of a mixed solvent system often reduces degradative reactions. Barbituric acid derivatives hydrolyze readily in water, particularly at a low pH. It has been shown, however, that pentobarbital sodium is soluble and stable in a vehicle containing 60% polyethylene glycol 400 and 10% ethanol in water at a pH of 8.

The aforementioned reactions do not occur in an anhydrous, non-polar vehicle, such as fixed oil, although the presence of a small amount of water may permit slight reactions. Oleaginous injections are subjected, however, to the disadvantages of being viscous (thus difficult to administer, particularly in cold weather) and of involving frequent incidence of pain upon injection.

## Solutes

The physical and chemical purity of solutes used for sterile preparations must also be exceptional. Obviously, the contaminants entering a product with a solute have the same effect as if they entered via the vehicle. Even small traces of contaminants may be detrimental to products, necessitating purification of the solute. For a few substances (for example, ascorbic acid and calcium gluconate), special parenteral grades are commercially available.

In addition, solutes should be free from microbial and pyrogenic contamination. This entails not only proper quality of the chemical as procured, but also storage conditions designed to prevent contamination, particularly after a container has been opened. Preferably, production lots should be designed to use the entire contents of packages of chemicals whenever possible.

## *Added Substances*

Substances added to a product to enhance its stability are essential for almost every product. Such substances include solubilizers, antioxidants, chelating agents, buffers, tonicity contributors, antibacterial agents, antifungal agents, hydrolysis inhibitors, antifoaming agents, and numerous other substances for specialized purposes. At the same time, these agents must be prevented from adversely affecting the product. In general, added substances must be non-toxic in the quantity administered to the patient. They should not interfere with the therapeutic efficacy or with the assay of the active therapeutic compound. They must also be present and active when needed throughout the useful life of the product. Therefore, these agents must be selected with great care, and must be evaluated as to their effect upon the entire formulation. An extensive review of excipients used in parenteral products and the means for adjusting pH of these products has recently been published and should be referred to for more detailed information. Table 23.2 provides a list of excipients commonly used in commercial parenteral products.

**Table 23.2:** Excipients used for commercial parenteral products

| **Excipients** | **Concentration range(%)** |
|---|---|
| **Antimicrobial preservatives** | |

| | |
|---|---|
| Benzyl alcohol | 0.5–10.0 |
| Benzethonium chloride | 0.01 |
| Butylparaben | 0.015 |
| Chlorobutanol | 0.25–0.5 |
| Metacresol | 0.1–0.25 |
| Methylparaben | 0.01–0.18 |
| Myristyl gamma picolinium chloride | 0.17 |
| Phenol | 0.065–0.5 |
| Phenylmercuric nitrate | 0.001 |
| Propylparaben | 0.005–0.035 |
| Thimerosal | 0.001–0.02 |
| **Solubilizers, wetting agents, or emulsifiers** | |
| Dimethylacetamide | 0.01 |
| Dioctyl sodium sulfosuccinate | 0.015 |
| Egg yolk phospholipid | 1.2 |
| Ethyl alcohol | 0.61–49.0 |
| Ethyl lactate | 0.1 |
| Glycerin | 14.6–25.0 |
| Lecithin | 0.5–2.3 |
| PEG-40 castor oil | 7.0–11.5 |
| Polyethylene glycol 300 | 0.01–50.0 |
| Polysorbate 20 | 0.01 |
| Polysorbate 40 | 0.05 |
| Polysorbate 80 | 0.04–4.0 |
| Povidone | 0.2–1.0 |
| Propylene glycol | 0.2–50.0 |
| Sodium deoxycholate | 0.21 |
| Sorbitan monopalmitate | 0.05 |

| | |
|---|---|
| Theophylline | 5.0 |
| **Buffers** | |
| Acetic acid | 0.22 |
| Adipic acid | 1.0 |
| Benzoic acid and sodium benzoate | 5.0 |
| Citric acid | 0.5 |
| Lactic acid | 0.1 |
| Maleic acid | 1.6 |
| Potassium phosphate | 0.1 |
| Sodium phosphate monobasic | 1.7 |
| Sodium phosphate dibasic | 0.71 |
| Sodium acetate | 0.8 |
| Sodium bicarbonate | 0.005 |
| Sodium carbonate | 0.06 |
| Sodium citrate | 4.0 |
| Sodium tartrate | 1.2 |
| Tartaric acid | 0.65 |
| **Bulking substances or tonicity modifiers** | |
| Glycerin | 1.6–2.25 |
| Lactose | 0.14–5.0 |
| Mannitol | 0.4–2.5 |
| Dextrose | 3.75–5.0 |
| Sodium chloride | Varies |
| Sodium sulfate | 1.1 |
| Sorbitol | 2.0 |
| **Suspending agents** | |
| Gelatin | 2.0 |
| Methylcellulose | 0.03–1.05 |
| Pectin | 0.2 |
| Polyethylene glycol 4000 | 2.7–3.0 |

| | |
|---|---|
| Sodium carboxymethylcellulose | 0.05–0.75 |
| Sorbitol solution | 50.0 |
| **Chelating agents** | |
| Edetate disodium | 0.00368–0.05 |
| Edetate calcium disodium | 0.04 |
| Edetate tetrasodium | 0.01 |
| **Local anesthetics** | |
| Procaine HCl | 1.0 |
| Benzyl alcohol | 5 |
| **Stabilizers** | |
| Creatinine | 0.5–0.8 |
| Glycine | 1.5–2.25 |
| Niacinamide | 1.25–2.5 |
| Sodium acetyltryptophanate | 0.53 |
| Sodium caprylate | 0.4 |
| Sodium saccharin | 0.03 |

**Antibacterial agents:** Antibacterial agents in bacteriostatic concentration must be included in the formulation of products packaged in multiple-dose vials, and are often included in formulations to be sterilized by marginal processes or made by aseptic manipulation. The requirements of activity, stability, and effectiveness of antibacterial agents in parenterals have been reviewed in published papers.

**Antioxidants:** Antioxidants, included in many formulations to protect a therapeutic agent susceptible to oxidation, particularly under the accelerated conditions of thermal sterilization, may function in at least two ways., i.e. (1) by being preferentially oxidized (reducing agents), and thereby gradually used up, or (2) by blocking an oxidative chain reaction in which they are not usually consumed. In addition, certain compounds have been found to act as synergists, increasing the effectiveness of antioxidants, particularly those blocking oxidative reactions. A fourth group of compounds are useful in this connection in that they complex with catalysts that otherwise would

accelerate the oxidative reaction. Because of the differences in action, combinations of these agents are sometimes used. In Table 23.3, the more commonly employed antioxidants are listed according to the above four groupings.

**Table 23.3:** Antioxidants used in sterile products

| Compound | Usual Concentration (%) |
|---|---|
| **Antioxidants (reducing agents)** | |
| Ascorbic acid | 0.02–0.1 |
| Sodium bisulfite | 0.1–0.15 |
| Sodium metabisulfite | 0.1–0.15 |
| Sodium formaldehyde sulfoxylate | 0.1–0.15 |
| Thiourea | 0.005 |
| **Antioxidants (blocking agents)** | |
| Ascorbic acid esters | 0.01–0.015 |
| Butylated hydroxytoluene (BHT) | 0.005–0.02 |
| Tocopherols | 0.05–0.075 |
| **Synergists** | |
| Ascorbic acid | 0.01–0.05 |
| Citric acid | 0.005–0.01 |
| Citraconic acid | 0.03–0.45 |
| Phosphoric acid | 0.005–0.01 |
| Tartaric acid | 0.01–0.02 |
| **Chelating agents** | |
| Ethylenediaminetetraacetic acid salts | 0.01–0.075 |

It should also be mentioned that for those products in which oxygen enters into a degradative reaction, an antioxidant effect can be achieved by displacing oxygen (air) from contact with the product. Usually, this is accomplished by saturating the liquid with either nitrogen or carbon dioxide

and sealing the final container after displacing the air above the product with the gas.

**Buffers:** Buffers are added to maintain the required pH for many products, as change in pH may cause significant alterations in the rate of degradative reactions. Changes in pH may occur during storage as a result of the dissolution of glass constituents in the product, release of constituents from rubber closures or plastic components in contact with the product, dissolution of gases and vapours from the airspace in the container and diffusion through the rubber or plastic component, or reactions within the product. Buffers must have the capacity to maintain the pH of the product against these influences, but not enough to prevent the body fluids from overwhelming the buffer following administration. In most cases, the biologic effectiveness of the drug is maximum at or near the biologic fluid pH rather than at the stabilizing pH of the injected product.

Acetates, citrates and phosphates are the principal buffer systems used, but buffer systems making use of other ingredients in the formulation are often used to reduce the total number of ingredients in the product. Buffer systems must be selected with consideration of their effective range, concentration, and chemical effect on the total product.

**Tonicity contributors:** Compounds contributing to the isotonicity of a product reduce the pain of injection in areas with nerve endings. Various agents are used in sterile products to adjust tonicity. Simple electrolytes such as sodium chloride or other sodium salts and non-electrolytes such as glycerin and lactose are most commonly used for this purpose. Tonicity adjusters are usually the last ingredients added to the formulation after other ingredients in the formulation are established and the osmolality of the formulation measured. Although the freezing point depression of the solution is most frequently used to determine whether a solution is isotonic, isotonicity actually depends on the permeability of a living semipermeable membrane that separates the solution from a biologic cell system. Most frequently, for sterile pharmaceutical preparations, the membrane concerned is the one enclosing the red blood cells. Therefore, a preparation cannot be considered to be isotonic until it has been tested in a biologic system. A hemolytic method, using red blood cells, has been described. If the formulation is still hypotonic (i.e. <280 mOsm/kg as measured by an osmometer), tonicity adjusting agents are added until the formulation is

isotonic. If the formulation is hypertonic, the degree of hypertonicity and the intended route of drug administration need to be considered. For i.v. administration, hypertonicity values up to 360 mOsm/kg are not considered harmful. However, for other routes of administration, efforts should be made to make the final preparation isotonic before administration. This can be accomplished either by reducing concentrations of ingredients, if acceptable, or by diluting the preparation before administration.

**Chelating agents:** Chelating agents may be added to bind, in non-ionizable form, trace amounts of heavy metals, which if free, would catalyze degradative changes. The chelating agent most commonly used is the trisodium or calcium disodium salt of ethylenediamine tetraacetic acid in a concentration of about 0.05% (w/v). An interesting example of the use of this chelating agent is the stabilization of thimerosal in poliomyelitis vaccine. Thimerosal is present as a bacteriostatic agent, but it is unstable in the presence of cupric ions, the breakdown products of which destroy the antigenicity of the vaccine. The chelating agent stabilizes the thimerosal, and thereby stabilizes the vaccine. The heavy metals extracted from rubber closures also may be bound by the presence of a chelating agent, reducing the possibility of reactions with ingredients in the formulation.

**Inert gases:** These have been used to displace oxygen from a solution and reduce the possibility of oxidative changes in the formulation. Inert gases may be used to stabilize solutions in other ways. For example, sodium bicarbonate injection decomposes, particularly during autoclaving, to produce sodium carbonate, carbon dioxide, and water. Saturation of the solution with carbon dioxide inhibits this reaction and stabilizes the solution.

**Protein stabilizers:** A number of ingredients have been shown to stabilize proteins, both in the dry and solution state. Serum albumin competes with therapeutic proteins for binding sites in glass and other surfaces and minimizes the loss of the protein caused by surface binding. With concern about viral contamination in natural substances like albumin, other competitive binding agents are being investigated (e.g. hetastarch). A number of different types of substances are used as ***cryoprotectants*** and ***lyoprotectants*** to minimize protein denaturation during freeze-drying. Primary examples include polyhydric alcohols (sorbitol, glycerol, polyethylene glycol); amino acids (glycine, lysine, glutamine); non-reducing sugars (trehalose, sucrose); and polymers such as dextran,

polyvinylpyrolidone, and methylcellulose. Surface active agents, such as Poloxamer 188 (Pluronic 68), polysorbate 80, and polysorbate 20 are widely used to minimize protein aggregation at air/water and water/solid interfaces. Antioxidants, buffers and chelating agents are also used to stabilize proteins in solution when necessary.

## CONTAINERS

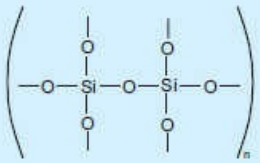
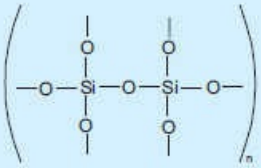
Containers are in intimate contact with the product. No container presently available is totally non-reactive, particularly with aqueous solutions. Both the chemical and physical characteristics affect the stability of the product, but the physical characteristics are given primary consideration in the selection of a protective container.

Glass containers traditionally have been used for sterile products, many of which are closed with rubber stoppers. Interest in plastic containers for parenterals is increasing, and such containers are being used for commercial ophthalmic preparations and IV solutions.

## Plastic Containers

The principal ingredient of the various plastic materials used for containers is the thermoplastic polymer. The basic organic structural unit for each type commonly encountered in the medical field is given in Table 23.4. Although most of the plastic materials used in the medical field have a relatively low amount of added ingredients, some contain a substantial amount of plasticizers, fillers, antistatic agents, antioxidants, and other ingredients added for special purposes. These ingredients are not usually chemically bound in the formulation and, therefore, may migrate out of the plastic and into the product under the conditions of production and storage. Considerable variability also has been encountered in the purity of the commercially available polymers.

**Table 23.4:** Comparative properties of container materials

| Material | Structural unit | Average density (g/ml) | Autoclavable (physical stability) |
|---|---|---|---|
| **Thermoplastic polymers** | | | |
| Polyethylene low density | $(—CH_2—CH_2—)_n$ | 0.92 | No |
| High density | $(—CH_2—CH_2—)_n$ | 0.96 | Yes |
| Polypropylene | $(—CHCH_3—CH_2—)_n$ | 0.90 | Yes |
| Polyvinyl chloride | | | |
| Flexible | $(—CHCl—CH_2—)_n$ | 1.2 | Yes (cautiously) |
| Rigid | $(—CHCl—CH_2—)_n$ | 1.2 | Yes (cautiously) |
| Polycarbonate | $(—O—C_6H_4—C(CH_3)_2—C_6H_4—CO_2—)_n$ | 1.2 | Yes |
| Polyamide (nylon) | $(—CH_2—(CH_2)_4—NHCO—)_n$ | 1.1 | Yes (repeatedly) |
| Polystyrene | $(—CH(C_6H_5)—CH_2—)_n$ | 1.05 | No |
| Polytetrafluoroethylene (Teflon) | $(—CF_2—CF_2—)_n$ | 2.25 | Yes (repeatedly) |
| **Glass** | | | |
| Soda-lime |  | 2.48 | Yes |
| Borosilicate |  | 2.23 | Yes |
| **Rubber compounds** | | | |
| Butyl | $(—CH_2—C(CH_3) = CH—(CH_2)_2—C(CH_3)_2—)_n$ | 1.3 | Yes |
| Natural | $(—CH_2—C(CH_3) = CH—CH_2—)_n$ | 1.5 | Yes |
| Neoprene | $(—CH_2—CH = CCl—CH_2—)_n$ | 1.4 | Yes |
| Polyisoprene | $(—CH_2—C(CH_3) = CH—CH_2—)_n$ | 1.3 | Yes |
| Silicone | $(—Si(CH_3)_2—O—Si(CH_3)_2—O—)_n$ | 1.4 | Yes |

| Additives present | Leachable constituents | Water vapour permeation | Gas permeation ($O_2$) with product | Potential reaction | Physical properties |
|---|---|---|---|---|---|
| Low | Additives (low) | High | Low | Low | Translucent, flexible |
| Low | Additives (low) | Low | Low | Low | Translucent, semi-rigid |
| Low | Additives (low) | Moderate | Low | Low | Translucent, semi-rigid |
| High | Additives (high) | High | Low | Moderate | Transparent, flexible |
| Low | Additives (low) | High | Low | Low | Transparent, rigid |
| Low | Additives (low) | High | Low | Low | Transparent, rigid |
| Low | Additives (low) | High | Low | High | Translucent, rigid, tough |
| Low | Additives (low) | High | High | Moderate | Transparent, rigid |
| Low | Additives (nil) | Low | Low | Nil | Translucent, tough, rigid, temperature-resistant |
| High | $K_2O$, $Na_2O$, MgO, CaO (High) | None | None | High | Optically clear, rigid |
| Low | $B_2O_3$, $Na_2O$, CaO, | Low | None | Low | Optically clear, rigid |
| Moderate | Additives (Moderate) | Low | Moderate | Moderate | Opaque, flexible |
| High | Additives (high) | Moderate | Moderate | High | Opaque, flexible |
| High | Additives (high) | Moderate | Moderate | High | Opaque, flexible |
| High | Additives (high) | Moderate | Moderate | Moderate | Opaque, flexible |
| Moderate | Additives (Moderate) | Very high | Very high | Low | Translucent, flexible |

As the name indicates, thermoplastic polymers melt at elevated temperatures. All of the polymeric materials listed in Table 23.4 except low-density polyethylene and polystyrene can be autoclaved if they have been formulated with a low amount of plasticizers, although most of them soften at autoclaving temperatures and care must be exercised to avoid fusing adjacent surfaces or otherwise deforming them. Also listed in Table 23.4 are certain properties of the plastic materials most commonly used for containers and components in drug packaging. These properties vary considerably with the type and amount of additive combined with the polymer.

Plastic containers are used mainly because they are light in weight, are non-breakable, and, when low in additives, have low toxicity and low reactivity with products. Tissue toxicity can occur from certain polymers, but additives are a more common cause. Reactivity due to sorption (absorption and/or adsorption) has been found to occur most frequently with the polyamide polymers, but additives leached from any of the plastic materials may interact with ingredients of the product.

Most polymers are adversely affected by the elevated temperatures required for thermal sterilization and have a relatively high permeability for water vapour (Table 23.4). Significant permeation of gases, including oxygen, may occur with some materials, polystyrene having by far the highest level of permeation of those listed.

A relatively new group of plastics, the polyolefins, deserve special mention. The two of interest today in the parenteral field are polypropylene and the copolymer polyethylene-polypropylene. Polypropylene is the most widely used. It is a linear polymer that can be produced to be highly crystalline. Because of its crystallinity, it has a high tensile strength, a high melting point of 165°C, and a relatively low permeability to gases and water vapours. It is translucent, abrasion-resistant, and has a high surface gloss. It withstands normal autoclaving temperatures. It must be stabilized with an antioxidant, however, the type and concentration of which must be carefully controlled to avoid leaching on one hand and degradation of the plastic on the other.

Interest in the use of plastic materials as containers for sterile products is increasing, but careful evaluation of their potential interaction with products with which they are in contact is essential. Flexible polyethylene containers

are used for ophthalmic solutions to be administered in drops and flexible polyvinyl chloride bags for IV solutions. The latter have a particular advantage over glass bottles in that no air from the patient's bedside need enter the container as the liquid flows out; the bag simply collapses. The new group of polymers, the polyolefins, have made possible the development of bottles that are rigid enough to hold their shape during processing but can collapse under atmospheric pressure as outflow of a solution occurs during IV administration to a patient. Thus, the characteristics of a rigid container are utilized during processing and handling, but the advantage of collapsibility of a flexible container is achieved for aseptic administration.

The USP has provided test procedures for evaluating the toxicity of plastic materials. Essentially, the tests consist of three phases (1) implanting small pieces of the plastic material intramuscularly in rabbits, (2) injecting eluates using sodium chloride injection, with and without alcohol, intravenously in mice, and injecting eluates using polyethylene glycol 400 and sesame oil intraperitoneally in mice and (3) injecting all four eluates subcutaneously in rabbits. The reaction from the test samples must not be significantly greater than non-reactive control samples.

## Glass Containers

Glass is still the preferred material for containers for injectable products. Glass is composed principally of the silicon dioxide tetrahedron, modified physicochemically by such oxides as those of sodium, potassium, calcium, magnesium, aluminum, boron, and iron. The two general types of glass are soda-lime and borosilicate (*see* Table 23.4). The glass that is most resistant chemically is composed almost entirely of silicon dioxide, but it is relatively brittle and can only be melted and molded at high temperatures. Boric oxide somewhat modifies the above characteristics as it enters the structural configuration, but most of the other oxides apparently enter the spaces within the structure and reduce the strength of the interatomic forces between the silicon and oxygen. Therefore, the latter oxides lower the melting point of the glass and are comparatively free to migrate. Consequently, they also lower the chemical resistance of the glass; i.e. they may migrate into a product over a prolonged period of contact, particularly with aqueous solutions. In solution, the oxides may hydrolyze to raise the pH, catalyze reactions, or otherwise enter into chemical reactions. Glass flakes are also sometimes produced as a result of the action of the solution. These interactions are markedly accelerated during the elevated temperature required for auto-claving.

### *Chemical Resistance*

The USP provides the Powdered Glass and the Water Attack tests for evaluating chemical resistance of glass. The test results are measures of the amount of alkaline constituents leached from the glass by purified water under controlled elevated temperature conditions; the Powdered Glass test is performed on ground, sized glass particles, and the Water Attack test is performed on whole containers. The conditions of the test must be rigidly controlled to obtain reproducible results since the quantity of alkaline constituents leached is small. The Water Attack test is used only with containers that have been exposed to sulfur dioxide fumes under controlled humidity conditions. Such treatment neutralizes the surface alkaline oxides, thereby rendering the glass more resistant chemically. This increased resistance is lost, however, if the container is subjected to repeated autoclaving, hot air sterilization, or hot detergent treatment.

On the basis of the results from the official tests, glass compounds are classified into four types, (*see* Table 23.4). The greatest chemical resistance is provided by Type I, and the least by NP (non-parenteral) glass. It should be noted, however, that within these types, as well as Types II and III, a range of compositions are available. The chemical resistance of the glass influences the selection of the type to be used for various products. Table 23.5 provides a brief summary of the general classes of products used with the four glass types. Type I glass is preferred for most sterile products, but Types II and III may be used when the product has a non-aqueous vehicle or the period of contact with the aqueous vehicle is brief, as with dry powders reconstituted just prior to use, or if the nonreactivity between the glass and product has been established.

**Table 23.5:** USP glass types, test limits and selection guide

| Type | General description | Type of test | Test limits: Size (ml) | Test limits: ml of 0.02 N $H_2SO_4$ | General use |
|---|---|---|---|---|---|
| I | Highly-resistant borosilicate glass | Powdered glass | All | 1.0 | Buffered and unbuffered aqueous solutions. All other uses |
| II | Treated soda-lime glass | Water attack | 100 or less<br>over 100 | 0.7<br>0.2 | Buffered aqueous solutions with pH below 7.0. Dry powders, oleaginous solutions |
| III | Soda-lime glass | Powdered glass | All | 8.5 | Dry powders, oleaginous solutions |
| NP | General-purpose soda-lime glass | Powdered glass | All | 15.0 | Not for parenterals. For tablets, oral solutions and suspensions, ointments, and external liquids |

## *Physical Characteristics*

The protection of light-sensitive products from the degradative effect of ultraviolet rays may be one of the important physical characteristics of a glass container. Ultraviolet rays can be completely filtered out by the use of amber glass; however, the colour of amber glass is produced largely by the presence of iron oxide, traces of which may subsequently be leached into the product. If the product contains ingredients subject to iron catalyzed chemical reactions, amber glass cannot be used. The product must then be protected from ultraviolet rays by means of an opaque carton surrounding a flint

(colourless) glass container.

In addition to other physical characteristics, glass containers should have sufficient physical strength to withstand the high pressure differentials that develop during autoclaving and the abuse that occurs during processing, shipping, and storage; a low coefficient of thermal expansion to withstand the thermal shocks that occur during washing and sterilization procedures; transparency to facilitate inspection of the contents; and uniform physical dimensions to facilitate handling by the mechanical machinery used for automatic production operations.

Glass containers may be manufactured by drawing from glass tubing or by blow molding. Ampoules, cartridges, and vials drawn from tubing have a thinner, more uniform wall thickness with less distortion than containers made by blow molding. The greater strength of blown vials and bottles, however, may be essential for handling by mechanical processing equipment. Large vials and bottles are made only by blow molding.

The physical dimensions of glass containers can readily be varied to meet design needs, especially those made by blow molding. An example is the double-chambered vial designed to contain a freeze-dried product in the lower chamber and the solvent in the upper chamber, separated by a rubber disc. The modifications in cartridge shapes for use with various disposable dosage units, and the wide-mouthed ampoules with flat or rounded bottoms to facilitate filling dry materials or viscous liquids, also illustrate the variations in physical dimensions possible with glass containers. The development of the easy-open ampoule several years ago, permitting opening without a file, was an important modification in the physical structure of these containers, marketed under the name “Colour-Break” and “Score-Break” ampoules.

Glass containers are sometimes coated internally with silicone fluid to produce a hydrophobic surface. To achieve permanency, the silicone must be baked at a temperature of approximately 150°C (300°F). This additional operation is justified for such applications as to reduce the adherence of heavy, costly suspensions or emulsions, or to increase the slippage of a plunger in a syringe barrel.

## Container use Considerations

The size of single-dose containers is limited to 1000 ml by the USP and multiple-dose containers to 30 ml, unless permitted otherwise in a particular monograph. The size limitation for multiple-dose vials is intended to limit the number of entries for withdrawing a portion of the contents of the vial with the accompanying risk of microbial contamination of the remaining contents. The particular advantage of these containers is flexibility of dosage offered to the physician. Single-dose containers are intended to provide sufficient drug for just one dose, the integrity of the container being destroyed when opened so that it cannot be reclosed and used again. Single-dose containers may range from liter bottles of IV solutions to 1 ml, or smaller, cartridges. The desire for further reduction in the risk of contamination, both bacterial and viral, and an increased control over the administration of drugs, particularly in a hospital, have led to the recent development of single-dose, disposable administration units. For most of these units, the product container is a glass cartridge with plastic and metal fitments separated from immediate contact with the product.

## Rubber Closures

Rubber closures are used to seal the openings of cartridges, vials, and bottles, providing a material soft and elastic enough to permit entry and withdrawal of a hypodermic needle without loss of the integrity of the sealed container. Figure 23.2 illustrates some of the styles of rubber closures and their relationship with typical containers.

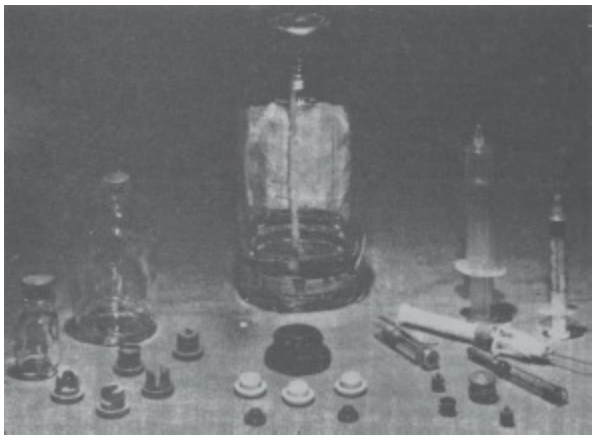


**Fig. 23.2:** Rubber closures associated with vials, bottles, and cartridges

### *Composition and Reactivity*

Rubber closures are compounded of several ingredients, principally, natural rubber (latex) and/or a synthetic polymer; a vulcanizing agent, usually sulfur; an accelerator, one of the several active organic compounds such as 2-mercaptobenzothiazole; an activator, usually zinc oxide; fillers, such as carbon black or limestone; and a variety of other ingredients such as antioxidants and lubricants. These ingredients are combined by kneading them into a homogeneous plastic mass on a roller mill. The homogeneous compound is rendered fluid and then vulcanized in the desired shape by means of molds under high pressure and temperature.

In Table 23.4 are listed the elastomeric polymers most commonly used in compounding rubber closures, as well as some of their comparative properties.

Ideally, closures should be completely nonreactive with the product with which they are in contact. No such ideal compound exists; therefore, each rubber compound should be tested for compatibility with each preparation

with which it is to be used. Two general compatibility problems exist, namely, the leaching of ingredients from the rubber compound with subsequent reaction with ingredients of the product, and the removal of ingredients from the product by sorption by the rubber compound or by vapour transfer through the closure.

Although compatibility problems are encountered with relative frequency, it should not be construed that rubber compounds invariably introduce such problems. Preliminary compatibility usually is assessed by placing the rubber closure in intimate contact with the product and maintaining the samples at elevated temperature levels for planned periods of time. At prescribed intervals, samples are examined for qualitative and quantitative evidence of chemical or physical change either in the closure or in the product.

## *Physical Characteristics*

Several properties of rubber closures are significant, particularly elasticity, hardness, and porosity. Rubber closures must be sufficiently elastic to provide a snug fit between the closure and the neck and lip of the glass container. They must also spring back to close the hole made by the needle immediately after withdrawal. Rubber closures must not be so hard that they require an excessive pressure to insert the hypodermic needle, and in doing so, must not produce a large number of fragments as the hollow needle cuts through the closure (coring). Although porous, they should not permit the easy transfer of water vapour and gases in either direction. *See* Table 23.4 for general comparisons. Minimal water vapour transfer is important, for example, to prevent the absorption of water by freeze-dried products.

Plastic or lacquer coatings are sometimes applied to the surfaces that will be in contact with the product. These coatings sometimes reduce vapour transfer, sorption, and leaching, but they do not usually provide the complete barrier desired. Teflon liners have been shown to provide an effective barrier against sorption and leaching.

The physical shapes of closures vary with their intended use. Several common shapes are shown in Fig. 23.2. The common flanged closure (center), slotted for freeze-dried products (left) or punctured for attachment of adapters for infusion sets (center, rear), and the plunger type for use with cartridges (right). Disc closures pre-assembled with aluminum caps are being

used to increase the speed of the operation in the high-speed packaging of antibiotics and other drugs. Other designs are used as needed for a particular application.

## *Testing*

Methods of testing for lot to lot uniformity of rubber closures have been studied for many years, but because of the nature of rubber compounds, consistent test results have been difficult to obtain. Progress has been made, however, and the LISP now describes physicochemical and biologic tests, but without test limits. The physicochemical tests on aqueous extracts include pH, turbidity (nephelos), residue on drying, iodine number, and heavy metals content. The biologic tests on saline, polyethylene glycol 400, and cottonseed oil extracts include acute and chronic toxicity in mice and rabbits. Further discussion of the purpose of these tests may be found in the literature.

## Devices

Devices, as considered here, are the various items of equipment used to convey the product from its container into the body of the patient or from one container to another; or the term may refer to the containers themselves. Devices associated with sterile products include the following:

Administration sets for large volume parenterals (LVPs)

Filter needles

Hypodermic needles

Hypodermic syringes

In-line filters

Plastic irrigating solution bottles

Plastic LVPs containers

Plastic ophthalmic dropping bottles

Transfer needles

Transfer sets

Although the contact time of the product with the device is usually brief, it is intimate; therefore, compatibility between the device and the product must be evaluated. For example, it has been shown that insulin can be adsorbed by PVC tubing during the time of contact for administrative of an IV solution, approximately 6 h.

The materials used for devices are mostly the same as those used for containers, but may include others if short-term contact has been shown to be acceptable. For example, nylon and silicone rubber are used for i.v. catheters, and stainless steel is used for hypodermic needles. Even aluminum is sometimes used for the hub and cannula of needles, but aluminum is much more reactive with some products than stainless steel. Parts of a device that do not come into contact with the product, such as the clamp on an i.v. administration set, need not pass product stability evaluation.

All device components must be visibly clean, but the fluid path through the device should meet the same rigid standards for cleanliness as the product. Usually, this must be achieved during manufacture and assembly of the device since final wet rinse or cleaning may be difficult or impossible,

owing to configuration of the device. Further, moisture residues may be detrimental to stability, causing leaching or interference with sterilization. Plastic particles from molding or metal dust from the sharpening of needles are examples of particulate matter that must be eliminated. If solvents are used to assemble components of the device, care must be taken to eliminate any excess, especially in the fluid path.

Tests performed by representative sampling of a lot of a finished device include those for toxicity as specified by the USP and functional tests appropriate for the specific device. The latter must be adequate to assure that a given lot performs as intended in use, although the criticality of a particular defect differs. Defects are usually expressed in terms of Acceptable Quality Levels (AQLs); the more critical AQLs are less than 1.0. For example, split hubs of needles would be more critical than slight discolouration of tubing, and may have AQLs of 0.065% and 2.5%, respectively. Therefore, great care must be exercised by quality control to prevent critical defects from being passed along to the user. For example, the integrity of the seal of the permeation section within a kidney dialysis unit is so critical to its use that every unit must be tested before it is released.

## PRODUCTION

The production process includes all of the steps from the accumulation and combining of the ingredients of the formula to the enclosing of the product in the individual container for distribution. Intimately associated with these processes are the personnel who carry them out and the facilities in which they are performed. The most ideally planned processes can be rendered ineffective by personnel who do not have the right attitude or training, or by facilities that do not provide an efficiently-controlled environment.

To enhance the assurance of successful manufacturing operations, all process steps must be carefully reduced to writing after being shown to be effective. These written process steps are often called standard operating procedures (SOPs). No extemporaneous changes are permitted to be made in these procedures; any change must go through the same approval steps as the original written SOP. Further, extensive records must be kept to give assurance at the end of the production process that all steps have been performed as prescribed, an aspect emphasized in the FDA's Good Manufacturing Practices. Such inprocess control is essential to assuring the quality of the product, since these assurances are even more significant than those from product-release testing. The production of a quality product is a result of the continuous, dedicated effort of the quality assurance, production, and quality control personnel within the plant in developing, performing, and confirming effective SOPs.

To differentiate quality assurance from quality control, the former function is usually one of pre-planning those factors that bear upon the quality of a product and is thus a preventative development process. Quality control may include this aspect, particularly if there is only one organizational group directly responsible for quality in a plant, but it more likely concentrates on those operations and tests that have been designed to evaluate the quality actually achieved in a product.

To enhance the visualization of the passage of materials through the various steps of the production process, a flow diagram is provided in Fig. 23.3. In the initial step, the formula ingredients, container components, and processing equipment that have been released for use are drawn from their respective storage areas. The ingredients are compounded according to the master formula in an environment designed to maintain a high level of

cleanliness. If the product is a solution, it is filtered during transfer to the aseptic filling room.

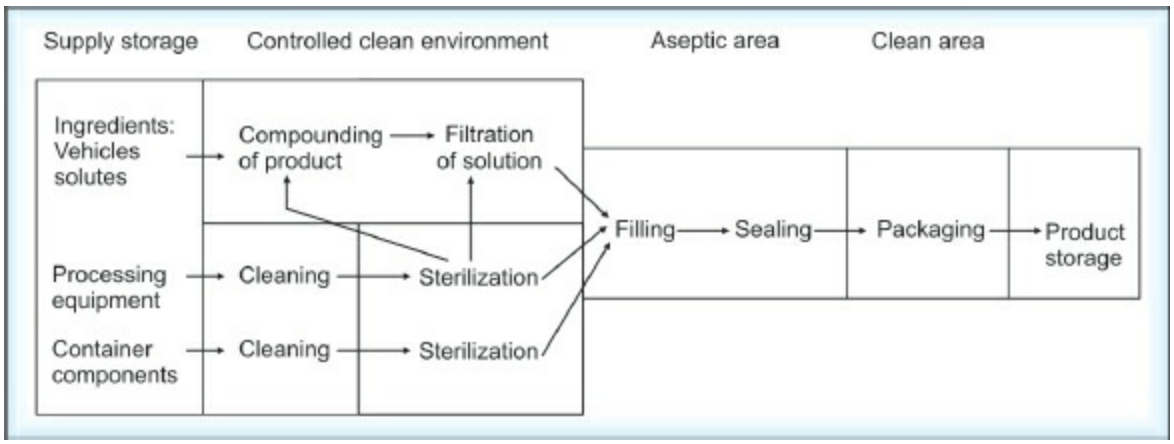



**Fig. 23.3:** Diagram of flow of materials through the production department

Process equipment and container components are cleaned thoroughly according to the required specifications, are assembled in a clean environment, and preferably, are sterilized and depyrogenated prior to use.

All equipment and supplies introduced into the aseptic filling area should be sterile (Fig. 23.3), having come directly from the sterilization process, preferably through double-ended sterilizers (Fig. 23.4). When this is not possible, packages, hose lines from equipment, and supplies should be passed through openings (ports) of minimal size that can be reclosed promptly, under aseptic conditions. Outer wrappings of packages should be loosened and the contents received, that is, the inner wrapping grasped, by personnel already in the aseptic filling room. When double wrappings are not feasible, the outer surfaces of boxes, packages, or equipment should be wiped with a disinfectant solution as they are transferred into the aseptic room. All supplies must be introduced into the aseptic filling rooms in such a manner that the aseptic state of these rooms is maintained, thereby preventing the introduction of environmental contamination into the product while it is being subdivided into individual containers. After these containers are sealed, contamination cannot enter the container and product.

As shown in Fig. 23.4, the product is sealed in its final container within the aseptic room. It is then transferred to the packaging area. This area is maintained clean but need not meet the standards imposed for the aseptic rooms or for the compounding room. Packaged products are placed in

quarantine storage until all tests have been completed and inprocess control records have been evaluated; then the product may be released for distribution.

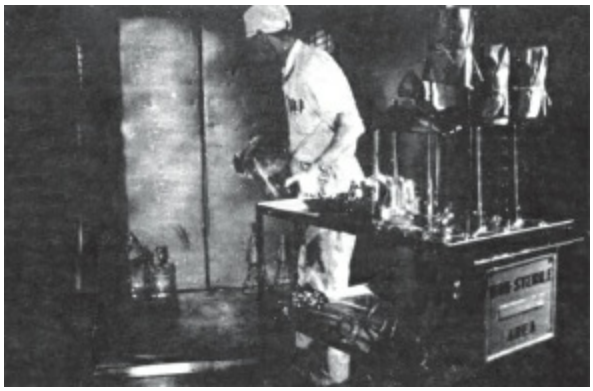


**Fig. 23.4:** Double-door sterilizing oven being loaded with clean equipment. The equipment will be removed sterile from the other door in the aseptic area after the sterilization cycle (*Courtesy* of Schering Corp)

## Facilities

The facilities for the manufacture of sterile products should be designed for control of cleanliness appropriate for each step. Nearperfect cleanliness must be achieved in the aseptic filling rooms. The surrounding areas should provide a buffer area in which standards of cleanliness are only slightly lower than those for the aseptic rooms. The prevention of contamination must be the primary objective in the design of these facilities.

To achieve such an exceptional design and construction standards, a knowledge of the purpose of the facility must be coupled with the utilization of the best construction materials and design. The ceiling, walls, and floors should be constructed of material that is easy to clean and non-porous, to prevent the accumulation of debris and moisture. Probably one of the best finishes for rigid surfaces is the “spray-on-tile”. This is a ceramic epoxy finish applied by spraying or painting to form a continuous, smooth, seal coating on the ceiling and walls. The rigorous effects from continuous washing with detergents and disinfectants, however, can cause even this epoxy finish to degrade and wear or peel. One of the best materials for floors is a ceramic-plastic cement applied as a thick coat over existing rigid flooring to form a continuous, sealed surface. Another flooring material used increasingly in areas of less heavy traffic is sheet vinyl with heat-welded seams, covered to the side walls and applied by adhesives on underlying surfaces. Movable metal partitions are sometimes used to provide flexibility of room arrangement, but they have the disadvantage of seams and joints, which are very difficult to seal.

Glass is often used in partitions to permit supervisory view of the operation, but more importantly, to provide more pleasant, better-lighted, less confining surroundings for the operators. Lighting fixtures should be recessed, and exposed piping or other dirt-collecting surfaces should not be tolerated. Furniture should be of non-porous, hard-surfaced materials, preferably stainless steel. Counters should be suspended from the wall.

Items of equipment that are difficult or impossible to sterilize should be kept out of the aseptic areas, if possible. If they must be used in the aseptic area, they should remain there and be continuously exposed to disinfecting processes. Whenever possible, operating machinery parts should be enclosed in stainless steel housing.

The mechanical servicing of electrical, gas, water, air ventilation, and other utility lines into these areas requires careful planning. One of the most effective plans for this is to provide a floor above, space beneath, or a corridor alongside of the production area where all service connections can be accessible and properly maintained. This prevents interruption of production, and most importantly, contamination of the production area by maintenance operations and personnel.

These basic design and construction features have been continued with the advent of HEPA-filtered laminar airflow capabilities (to be discussed in the following section). Laminar airflow is most frequently added to a clean room to achieve greater environmental control in a local area, such as in a workbench enclosure or over a filling line.

The reader is referred to the literature for a further discussion of the design, construction, and operation of facilities for the preparation of sterile products.

## *Environmental Control*

Effective environmental control, both physical and biologic, is essential, but the level achievable is related to the characteristics of the facility, as discussed previously. Further, rigid standards from plant to plant and from one geographic location to another are not appropriate. Allowance also must be made for variations in control associated with seasonal conditions.

The standards of environmental control vary, depending on the area involved (cleanup, packaging, compounding, or filling) and the type of product being prepared. Unquestionably, the entire area used for the preparation of a product prepared aseptically (without terminal sterilization) must be maintained under the most rigid control that existing technology permits. If the product is to be terminally sterilized, somewhat less rigid biologic control of the compounding and filling areas may be acceptable, but rigid standards of cleanliness must be maintained. High standards of cleanliness, excluding daily use of the disinfecting procedures, are usually acceptable for the clean-up and packaging areas.

## *Traffic Control*

Excellence in environmental control would be relatively easy to achieve were it not for the necessity of personnel and supplies to move from one area to

another. Therefore, a carefully designed arrangement to control and minimize traffic, particularly in and out of the aseptic areas, is essential. A floor-plan for a sterile products facility is shown in Fig. 23.5. Note that the only access directly from the outside is to the personnel wash rooms, the equipment wash rooms, the non-sterile manufacturing area, and the capping (packaging) area. Access by personnel to the aseptic corridor and aseptic compounding and filling rooms is only through an airlock. Passthrough openings and double-ended sterilizers are provided to permit controlled passage of supplies from non-aseptic to aseptic areas.

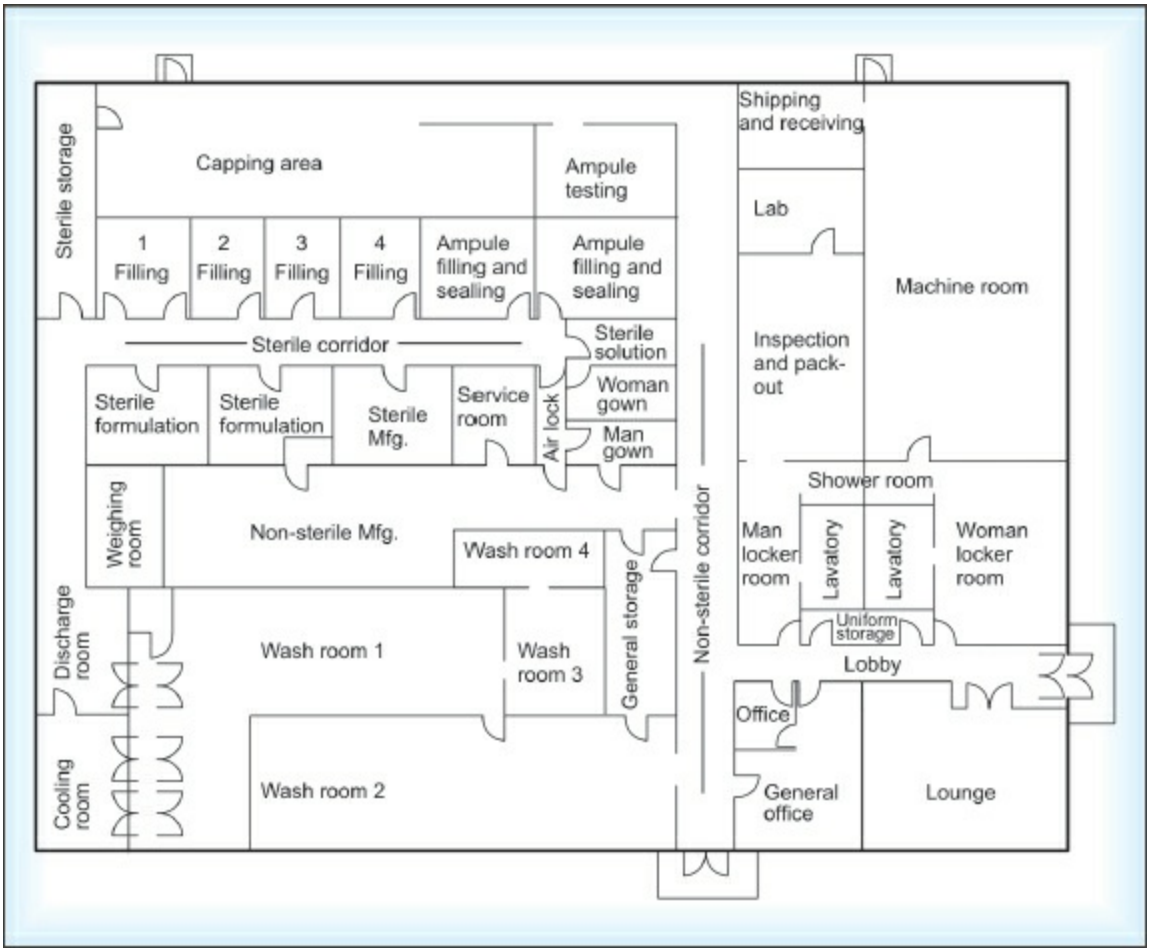



**Fig. 23.5:** Floor plans for industrial sterile products production area (*Courtesy* of Sobering Corp)

Personnel should be permitted to enter aseptic areas only after following rigidly prescribed procedures for removing their street clothing, washing their hands, and donning gowns, hats, shoes, facemasks, gloves, and other

prescribed attire. Once they have entered the aseptic area, they should not be permitted to move in and out of the area without regowning. Personnel assigned to cleaning and packaging should be restricted to these areas. Unauthorized personnel should never be permitted to enter the aseptic area.

## *Housekeeping*

Cleaning personnel must be imbued with the philosophy that not one remaining particle of debris is acceptable. Only with such an approach will the conditions be provided for achieving and maintaining proper environmental control. It must also be recognized that many, if not most, critical contaminating particles are subvisual in size.

All equipment and surrounding work area must be cleaned thoroughly at the end of the working day. No contaminating residues from the concluded process may remain. The ceiling, walls, and other structural surfaces must be cleaned with a frequency commensurate with the design of the facility, i.e. less frequently in a laminar airflow facility than in one that is not bathed with a constant clean airflow. All cleaning equipment should be selected for its effectiveness and freedom from lint-producing tendencies. It should be reserved for use in the aseptic areas only.

## *Surface Disinfection*

After thorough cleaning, all surfaces should be disinfected, at least in the aseptic areas. An effective liquid disinfectant should be sprayed or wiped on all surfaces (*see* Chapter 22).

Irradiation from ultraviolet lamps that are located to provide adequate radiation intensity on the maximum extent of surfaces in a room and that are maintained free from dust and films further reduces the viable microorganisms present on surfaces and in the air. Ultraviolet rays may be particularly useful to irradiate the inside, exposed surfaces of processing tanks, surfaces under hoods, surface of conveyor belts, and similar confined surfaces that are otherwise difficult to render aseptic; however, they cannot reach unexposed surfaces such as pipe connections to tanks, the underside of conveyors, and the inside of containers.

Ultraviolet lamps must be kept clean, and care must be taken to check for a decrease in effective emission, a natural occurrence due to a change in the glass structure with aging.

## *Air Control*

In any area occupied by personnel, the air must be exchanged at frequent intervals. Fresh outside or recycled air must first be filtered to remove gross particulate matter. A spun glass, cloth, or shredded polyethylene filter may be used for this preliminary cleaning operation. At times, more than one pre-filter may be used in series, the first of quite large and the next of somewhat smaller pore size, to provide a gradation of particle size removal from heavily contaminated air. To remove finer debris down to the sub-micron range, including microorganisms, a High Efficiency Particulate Air (HEPA) filter, defined as at least 99.97% efficient in removing particles of 0.3 μm size and larger, and composed of glass fibers and fillers or electrostatic precipitators, may be employed. Air passing through these units can be rendered virtually free from foreign matter. Another air cleaning system washes the air with a disinfectant and controls the humidity at the same time.

Blowers should be installed in the air ventilation system upstream to the filters so that all dirt-producing devices are ahead of the filters. The clean air is normally distributed to the required areas by means of metal (preferably stainless steel) ducts. Since it is practically impossible to keep these ducts as clean as required, it is normally preferable to install HEPA filters at the point where the clean air enters the controlled room. Alternatively, the ducts may be replaced with a room (a plenum), usually above the production area, into which clean air is blown and then distributed through openings into each of the process rooms. The entire plenum can be kept clean and aseptic.

The clean, aseptic air is distributed in such a manner that it flows into the maximum security rooms at the greatest volume flow rate, thereby producing a positive pressure in these areas. This prevents unclean air from rushing into the aseptic area through cracks, temporarily opened doors, or other openings. The pressure is reduced successively so that the air flows from the maximum security area to the hallway or other less critical areas for return to the filtration system. At the intake end of the system, fresh air, usually about 25%, is continually introduced for the comfort and needs of the personnel. Further, the air is usually conditioned with respect to temperature and humidity for the comfort of the personnel, and sometimes, to meet the special requirements of a product.

A relatively new air control system, based on laminar flow principles, has

greatly improved the potential for environmental control of aseptic areas. Currently, it is the only means available for achieving a Class 100 clean room. A Class 100 clean room is defined as a room in which the particle count in the air is not more than 100 per cubic foot of 0.5 μm and larger in size. HEPA-filtered air is blown evenly out of the entire back or top of a workbench (Fig. 23.6), or entire side or ceiling of a room. The airflow must be uniform in velocity and direction throughout any given cross-section of the area, being exhausted from the opposite side. The air velocity employed should be 100 ± 20 ft/min. Contamination is controlled because it is swept away with the airflow. Any contamination introduced downstream from the filter, however, may be carried to a critical working area located farther downstream. This may be caused by the improper placement of supplies, the manipulation of personnel, or discharge from operating equipment. Because the risk of introducing contamination in such a manner is generally considered to be less with vertical flow from ceiling-mounted HEPA filter units, vertical flow is most frequently utilized to protect critical sections of processing lines and similar activities. Horizontal flow, on the other hand, is used to protect processing lines and used most frequently for workbenches. Numerous reports have shown the marked benefit of laminar airflow for controlling working environments, from small workbenches to entire rooms.

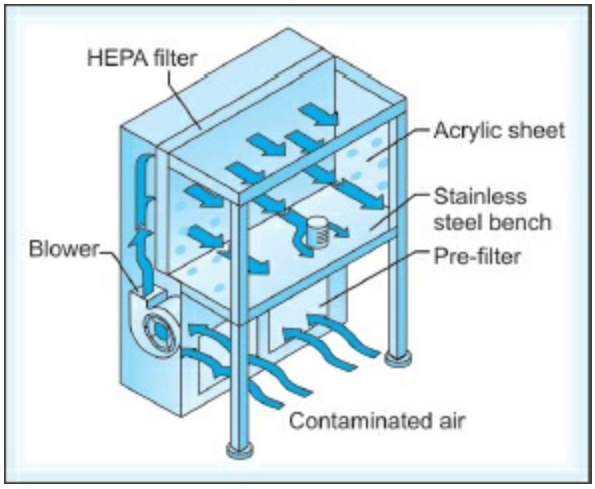



**Fig. 23.6:** Horizontal laminar airflow clean bench

Although class 100 work environments are normally specified for the most critical aseptic and/or clean operations, achieving such levels of cleanliness is expensive and requires effective maintenance and monitoring.

It should be recognized that not all operations associated with parenteral medications require such an environment. To such an end, other classes are defined. For example, a Class 10,000 room is the one in which the particle count is no more than 10,000 per cubic foot of 0.5 μm and larger in size. Such a cleanliness level is usually considered suitable for buffer areas around Class 100 worksites in which operations such as handling pre-cleaned containers, process filtration, and aseptic gowning of personnel may be performed. Still less stringent requirements would be applied to laboratories, stock staging areas, and finish packaging, where a Class 100,000 or similar cleanliness level would usually be considered suitable.

The effectiveness of the environmental control system is normally monitored on the basis of deviation, usually upward, from baseline counts determined from extensive testing by the environmental control procedures utilized. Biological evaluation methods most frequently utilized are summarized in Table 23.6 and include settling and surface contact nutrient agar plates, air impingement on nutrient media, and membrane filtration. Particulate matter evaluation methods, also summarized in Table 23.6, include membrane filtration and electronic particle counters. A combination of methods is usually utilized, involving two or more that best identify the control achieved of the particular circumstances of a given environment.

**Table 23.6:** Environmental monitoring methods

| Method | Principle of operation | Advantages | Disadvantages |
|---|---|---|---|
| **Biologic evaluation** | | | |
| Settling plates | Gravitational fallout in a given time on a given area | Uncomplicated. Low cost | Only heavier particles settle and are collected Irregularities in counts due to wild air currents, physical movements of personnel, etc. |
| Slit sampler | Measured volume of air drawn through slit and impacted on nutrient agar as plate turns | Measured volume of air sampled. Sampling related to time as plate turns | Velocity of impaction likely to have lethal effect on vegetative forms. Drying effect may be lethal to vegetative cells. Must be used with access to electricity and vacuum |
| Centrifugal sampler | Measured volume of air centrifugally blown on nutrient agar strip | Measured volume of air sampled. Unit can be easily carried by hand and is battery-operated. Unit head is sterilizable and body sanitizable | Velocity of impaction likely to have lethal effect on vegetative forms. Drying effect may be lethal to vegetative cells |
| Cascade sieve sampler | Measured volume of air cascaded through up to six plates of decreasing pore size and impacted on nutrient agar plates, with the smallest particles collected on the last plate | Measured volume of air sampled. Permits gradation of particles by size | Velocity of impaction likely to have lethal effect on vegetative forms. Must be used with access to vacuum. Affected by high RH; best used with dry conditions |
| Liquid impinger | Measured volume of air bubbled through liquid nutrient medium with impingement in liquid | Measured volume of air sampled. Less lethal action on vegetative forms since impingement is in "soft" liquid. Accepted as reference method | Liquid must be filtered or plated to isolate microorganisms. More complicated procedure Time-consuming procedure |
| Membrane filter sampler | Measured volume of air drawn through membrane filter with particles retained on surface and filter then incubated on nutrient agar plate | Measured volume of air sampled. May be used also for microscopic particle counting | Additional step of membrane being placed on nutrient agar in plate. Air pockets under membrane prevent growth. Drying effect of microorganisms on membrane is lethal to most vegetative forms. Vacuum source required |
| **Particulate matter evaluation** | | | |
| Membrane filter | Measured volume of air drawn through membrane filter with particles retained on surface | Measured volume of air sampled. Particles microscopically visible and identifiable. Particles can be sized and dimensionally described | Counting and sizing require experienced and trained microscopist. Identification of particles requires experienced and trained microscopist. Time-consuming procedure |
| Right angle light-scattering instrumental counter | Particle in viewing cell scatters light at right angles from incident light to photodetector tube | Quantitative count of particles in measured volume of air obtained. Instant results given. Range of sizes of particles measured | Sizing affected by light-scattering characteristics of particle surface. No differentiation between viable and non-viable particles costly |
| Near forward light-scattering instrumental counter | Particle in viewing cell scatters light forward from incident light to photodetector tube | Quantitative count of particles in measured volume of air obtained. Instant results given. Range of sizes of particles measured. Greater intensity than other methods; therefore, smaller particles can be counted | Sizing affected by light-scattering characteristics of particle surface. No differentiation between viable and non-viable particles. More costly than other methods |

## Personnel

The people who produce sterile products are usually non-professional persons, supervised by those with professional training. To be effective operators, they must be inherently neat, orderly, reliable, and alert, and have good manual dexterity. They should be appreciative of the vital role that every movement has in determining the quality of final product, i.e. its freedom from contaminants.

All employees should be in good health and should be subjected to periodic physical examinations. They should understand their responsibility to report the developing symptoms of a head cold, sore throat, or other infectious diseases so that they can be assigned to a less—critical area until they have fully recovered.

The attire worn by personnel in the aseptic areas usually consists of sterile coveralls, hoods, face masks, and shoe covers. Sterile rubber gloves also may be required.

Personnel entering the aseptic areas should be required to follow a definite preparatory procedure. This should include removing at least outside street clothing, scrubbing the hands and arms thoroughly with a disinfectant soap, and donning the prescribed uniform. A full-body water and soap shower would be essential in most biologic product processing plants—usually, both when entering and leaving the area to control contamination in both directions between personnel and the product. It must be recognized, however, that removing natural oils from the skin temporarily increases particle shedding. An air shower for the fully attired worker may be used at times to blow away loose lint, although the disruptive air currents generated may be detrimental to overall air control. A vibrating foot mat or a disinfectant foot bath also may reduce the transfer of contamination.

Because people are continually shedding viable and non-viable particulate matter from body surfaces, uniforms are worn to help control this emission. Preferably, they are of the coverall type and made of synthetic fibers such as Dacron. Dacron cloth is made of a continuous fiber, which makes it essentially lint-free and in air conditioned rooms, is acceptably comfortable. Hats and masks are sometimes made of special parchment paper and are discarded after use. Spun polyethylene has recently found favour as a material for uniforms.

Personnel working in equipment wash rooms, sterilizing rooms, and packaging areas are normally required to don clean uniforms daily and to be conscious of cleanliness, but are not required to meet the special requirements for personnel entering the aseptic areas.

## Processing

The initial processing step is the procurement of acceptable components (*see* Fig. 23.3). In a plant, the majority of components are requisitioned from tested and approved stock, and are then subjected to whatever processing steps are required to prepare them for use. A few components, such as Water for Injection, are manufactured to specifications as needed.

### *Water for Injection*

Water for Injection (WFI), usually is prepared by distillation in a still specifically designed to produce the high-quality water required. Reverse osmosis, however, is a method that is now approved by the USP, and is receiving increasing attention and use. The specifications for the quality of the water required have been discussed under the heading “Vehicles,” earlier in this chapter.

The specifications for a still should include (1) pre-purification of feed water by chemical softening, deionization, or filtration to improve the quality of the distillate and reduce the frequency of required cleaning due to insoluble scale in the boiler, (2) removal of entrained contaminants from the vapour before it is condensed by passage through an efficient baffle system, (3) ejection of volatile constituents from the top of the system before the vapour is cooled so that they will not redissolve and appear in the condensate and (4) construction of all surfaces that will come in contact with the vapour and condensate of a material that will not dissolve in even trace amounts, preferably pure tin, 304 stainless steel, or borosilicate glass.

In addition to conventional stills, two types of stills frequently used for the production of large volumes of water are the vapour compression stills and the multiple effect stills. While they operate on somewhat different principles, both utilize initially-heated feed water and steam to conserve on energy consumption and cooling water. Both types are capable of producing high-purity water at rates of 50 to 1000 or more gallons per hour.

A reverse osmosis system functions by applying pressure (usually 200 to 400 psi) to raw water sufficient to force the permeation of water through a select semipermeable membrane in the opposite direction to natural osmosis. The membranes most commonly used are composed of cellulose esters or polyamides (nylon) and are effective in retaining all macromolecules and

85% or more of small ions such as $Na^+$ and $Cl^-$. Since pyrogens are macromolecules, they should be retained as well as such viable particles as microorganisms. Greater efficiency and reliability are achieved by passing the water through two membranes in series. The acceptance of reverse osmosis for the preparation of WFI is increasing as experience is gained with the system and its characteristics are understood more fully.

## *Storage and Distribution*

The storage and distribution of WFI are as important as its production. A closed system is desirable, with air exchange through a filter that removes microorganisms, dirt, and vapours from the air as the tank is filled and emptied. Should microorganisms gain entrance to the tank, they may be prevented from multiplying by holding the temperature of the water at 80°C by means of a steam coil at the bottom of the tank. Normally, WFI should not be held for more than 24 h at room temperature before it is used, but if held at 80°C, continuous addition of fresh WFI as usage occurs is a common practice. The constant danger of microbial contamination, in spite of precautions, and subsequent development of pyrogenic substances in the water demand careful storage requirements.

The distribution of WFI from the storage tank to the point of use may be by direct withdrawal from the tank, or in large plants, through a pipe system. When a pipe system is used, special precautions must be followed to prevent contamination, including construction with welded stainless steel pipe, a closed system preferably with continuous circulation to avoid stagnation, maintenance at elevated temperature, complete isolation from all other piping systems, elimination of elbows or other pockets in which water can stagnate for long periods, and a means of thorough cleaning and sanitation at frequent intervals, as with clean steam or hot alkali.

## *Cleaning Equipment and Containers*

Equipments and containers to be used in the processing of a sterile product must be scrupulously cleaned. New, unused containers and equipments are contaminated principally with dust, fibers, and chemical films, which usually are relatively easy to remove, often by rinsing only. Debris that is more dangerous and more difficult to remove may be present as a residue from a previous use. Such debris usually must be removed by vigorous treatment

with hot detergents.

In general, equipment used previously should be scrubbed by hand immediately after use with an effective detergent that does not leave a residue of its own. Whenever possible, the equipment should be disassembled so that each part can be thoroughly scrubbed and cleaned, with particular attention to screw threads, joints, and other dirt-collecting structures. Live steam can sometimes be used to loosen debris effectively, particularly in areas that are not easily accessible. After cleaning, the equipment should be rinsed several times, with a final rinse with WFI. Just prior to reuse, large clean tanks and similar equipment should be rinsed thoroughly with WFI. Reserving equipment for use with only one type of product reduces cleaning problems.

A new method for large tanks, pipelines, and associated equipment that can be isolated and contained within a process unit has been developed and identified as a CIP (Clean in Place) system. Cleaning is accomplished primarily with high-pressure rinsing treatments delivered automatically within the equipment. This is usually followed by steam sanitization through the same system, although actual sterilization of the entire system with attached components, such as filters, is being investigated.

For glass or metal equipment small enough to be transported by hand, machine washing is possible. The glassware or metalware is automatically conveyed, usually in an inverted position, through a series of rigorous, high-pressure treatments, including hot detergent, hot tap water, and final rinses with distilled water. Because many containers have restricted openings, it is essential that the treatments in any washer be introduced through tubes into each container, with a smooth flow out. All parts of the machine coming in contact with the treatments must be noncorrosive so that metallic contaminants from the machine are not deposited on the glassware.

When high processing rates are not required, a cabinet-type washer embodying these principles may be used. In such a machine, the glass ware to be washed is held on a rack within a cabinet while the machine automatically goes through the sequence of treatments of the cycle.

## *Rinsing New Containers*

In cleaning new glassware, the detergent treatment is usually eliminated, and with it, the risk of a detergent residue. Without the detergent treatment, the cycle is essentially a rinsing process. To loosen debris by rinsing, alternating

hot (preferably clean steam) and cold treatments should be used. Final rinses should be done with filtered WFI. This sequence of treatments may be performed on the machines described above or on a rotary rinser. The containers are inverted on spindles in the front of the machine and carried through a series of rinses in one rotation. For ampoules or containers with a markedly constricted opening that makes water drainage incomplete, the final treatment is usually a blast of clean air to blow out the remaining water.

After cleaning, it is essential that the clean containers be protected from dust and other particulates that might be present in the environment. Therefore, the clean containers are often removed from the rinser and placed in clean stainless steel boxes for sterilization under the protection of HEPA-filtered laminar airflow (Fig. 23.7).

**Fig. 23.7:** Rinser designed to clean vials by racks sized to shipping cartons or by individual handling. Note laminar airflow to protect clean, wet vials

Conveyor-type rinsers, as shown in Fig. 23.8, have relatively high production rates. They have an advantage over the above rinsers in that they deliver the clean container at the opposite end from that at which the unrinsed containers are loaded. The delivery end can, therefore, be in an adjoining room (Fig. 23.7), away from the dust and dirt associated with packing cartons.

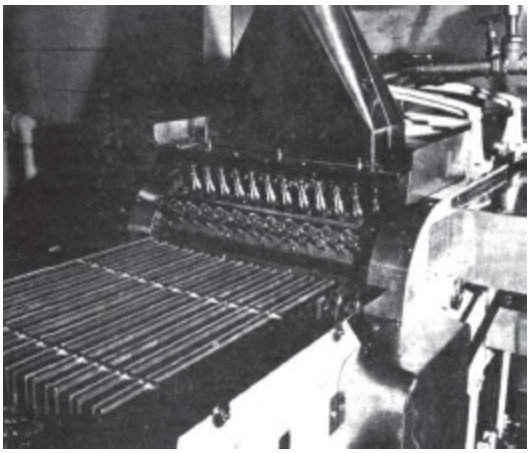


**Fig. 23.8:** Conveyor rinser delivering clean vials to an adjoining clean room (*Courtesy* of Schering Corp and Cozzoli Machine Co)

To eliminate the tedious handling of each individual container, rack loading rinsers have been designed with rack sizes adapted to container packing cartons. Such a rinser is shown in Fig. 23.7.

Minute quantities of oils, proteinaceous materials, or other debris retained in the rinser from a previously used container could be carried through many rinsing operations and be deposited on glassware subsequently being processed.

High-speed processing lines often are designed to clean containers by rinsing with clean water, or sometimes simply with clean air, under high pressure. The containers are shipped from the manufacturer on a support tray enclosed within a shrunken, tightly-fitted, polyethylene sheet to minimize the accumulation of dirt during shipment. The cleaned containers are usually then fed by conveyor to hot-air sterilizing tunnels, thereby minimizing handling of the containers, increasing the speed of processing, and allowing the clean sterile containers to be delivered through-the-wall into the aseptic filling room.

## *Cleaning Rubber and Plastic Components*

Rubber closures are usually washed by mechanical agitation in a tank of hot detergent solution (such as 0.5% sodium pyrophosphate) followed by a series of thorough water rinses, the final rinse being WFI. The objective is to remove the surface debris accumulated from the molding operation and from handling, and leachable constituents at or near the surface. Part of the debris

is attracted and held on the surface by electrostatic forces. Similarly, plastic materials accumulate surface debris.

Abrasion may occur during agitation, resulting in small, loose pieces of the rubber or plastic material. This problem is more acute with two-component closures, i.e. a rubber disc inserted in an aluminum cap, because agitation usually causes more abrasion of the aluminum and produces small aluminum fragments, which adhere tenaciously to the rubber disc or wedge between the disc and the cap.

Therefore, the multiple objectives for washing closures and other parts include loosening debris, minimizing abrasion, and sweeping away the loosened debris. Household cloth-washers of the horizontal rotating basket type or the center post agitator type have been used, but neither meets the requirements for an ideal washer for closures. More commonly today, the closures are subjected to gentle agitation with air bubbles, basket rotation accompanied by spray rinsing, or simple water movement followed by extensive rinsing with WFI. Handling after cleaning must be done carefully to prevent particle generation from abrasion and to prevent pickup of dust and other particulate matter from the environment. Therefore, the clean closures are often handled under the protection of HEPA-filtered laminar airflow.

Sometimes, the closures are subjected to an autoclave cycle as a part of the cleaning process. Such a treatment aids in loosening the surface debris and also leaching some of the extractives, from the closures thus reducing the subsequent contamination of the product. It should be remembered, however, that excessive heating is detrimental to the life of rubber and thermoplastic materials.

Particular attention should be given to the cleaning of rubber and plastic tubing. When tubing is reused, adequate cleaning of the lumen, which is not readily accessible, may be virtually impossible. When cleaning is necessary, special brushes may be helpful in reaching the lumen. The safest approach is to retain the tubing for reuse with the same product or to discard after one use.

## *Sterilization of Equipment*

In general, equipment, containers, closures, and all other components should be sterilized after cleaning and prior to use. The principles and practices of

sterilization are discussed in Chapter 22.

## Compounding the Product

The product should be compounded under clean environmental conditions (*see* Fig. 23.3). Aseptic conditions usually are not required, since it may not be possible or feasible to sterilize some of the ingredients or the equipment, e.g. large tanks. Whenever possible, however, equipment and ingredients should be sterile to reduce the microbial load.

The accuracy of compounding should meet the rigid standards accepted in pharmaceutical procedures, regardless of the batch size, recognizing that small multiple errors may be additive. In large batches, particular attention must be given to achieving and maintaining homogeneity of solutions, suspensions, and mixtures, maintaining a given temperature, and accelerating cooling. The order of mixing ingredients may become highly significant, for example, owing to the physical problem of distributing a pH-adjusting ingredient throughout a large tank of liquid. Compounding problems for large batches of product are often different from those for small batches.

Good planning requires anticipation of reasonable stock needs for products so that a single large lot, instead of several small lots, may be prepared. If made in divided portions, each portion is “a lot” requiring separate testing, thus multiplying the time and cost required.

### *Filtration of Solutions*

Solutions must be filtered. The primary objectives of filtration are clarification or sterilization of a solution. The two objectives differ principally in degree. Clarification is termed “polishing”, and a highly polished solution requires the removal of particulate matter down to at least 3 μm size. Further reduction in the size of the particulate matter removed, to approximately 0.3 μm, results in sterilization, the removal of viable microorganisms and spores. Where the objective of filtration is sterilization, a highly polished solution is concurrently produced. A solution having a high polish conveys the impression of exceptional quality and purity, a highly desirable characteristic for a sterile solution.

The various types of filters employed for sterile products, their selection, and use have been discussed in Chapter 5.

After filtration, the solution must be protected from environmental

contamination until it is sealed in the final container. Normally, this is best accomplished by collecting the filtrate in a container that is a part of a closed system, with air exchange through a bacteria-retentive filter. The filtrate is fed directly from the collecting vessel to the filling machine through sterile hose connections. A secondary “in-line” filter is often included as close to the outlet of the filler as possible to collect any lint or other particulate matter picked up from the lines or equipment. It is at the moment of filtration that a solution must pass from a clean environment to an aseptic environment (*see* Fig. 23.3), particularly if it has been sterilized by the filtration process.

## *Filling Procedures*

A liquid may be subdivided from a bulk container to individual dose containers more easily and uniformly than a solid. Mechanical subdivision of a mobile, low-density liquid can be achieved with light-duty machinery, but viscous, sticky, or high-density liquids require much more rugged machines to withstand the pressure required to dispense them.

## *Filling Equipment for Liquids*

Certain fundamental features are found on all machines used for filling containers with liquids. A means is provided for repetitively forcing a measured volume of the liquid through the orifice of a delivery tube designed to enter the constricted opening of a container. The size of the delivery tube is governed by the opening in the container to be used, the viscosity and density of the liquid, and the speed of delivery desired. The tube must freely enter the neck of the container and deliver the liquid deep enough to permit air to escape without sweeping the entering liquid into the neck or out of the container. To reduce the resistance to the flow of the liquid, the tube should have the maximum possible diameter. Excessive delivery force causes splashing of the liquid and troublesome foaming, if the liquid has a low surface tension.

The delivery of relatively small volumes of liquids is usually obtained from the stroke of the plunger of a syringe. The stroke of the syringe forces the liquid through a two-way valve that provides for an alternate filling of the syringe from a reservoir and delivery to a container. For heavy, viscous liquids, a sliding piston valve provides more positive action.

A drop of liquid normally hangs at the tip of the tube after a delivery.

When the container to be filled is an ampoule, withdrawal of the tube without wetting the long restricted neck is almost impossible, unless the hanging drop of liquid is retracted. Thus, a retraction device is designed as a part of most filling machines.

Filling machines should be designed so that the parts through which the liquid flows can be easily demounted for cleaning and sterilization. These parts also should be constructed of non-reactive materials such as borosilicate glass or stainless steel. Syringes are usually made of stainless steel when the pressures required for delivery of viscous liquids or large volumes would be unsafe for glass syringes.

Filling machines, such as the one shown in operation in Fig. 23.9, can be designed to provide high delivery volume precision. The stroke of the syringe can be repeated precisely; therefore, once a particular setting has been calibrated for a delivery, high precision is possible. The precision can be affected by certain operating factors, however, such as the speed of delivery, the uniformity of speed, the expansion of rubber tubing connecting the valve with the delivery tube, and the rapidity of action of the valves.

**Fig. 23.9:** Aseptic filling of vials with liquid under vertical laminar airflow, followed by stoppering with forceps by hand

Sterile solutions of relatively low potency dispensed in large volume (up to one liter) do not normally require the precision of filling that is required for small volumes of potent injectables. Therefore, bottles of solution are usually filled by gravity, pressure, or vacuum filling devices. Gravity filling is

relatively slow, but is accomplished in a simple manner. The liquid reservoir is positioned above the filling line with a hose connection from the reservoir to a shut-off device at the filling line. The shut-off device is usually hand-operated, and the bottles are filled to graduations on the bottles.

The pressure pump filler often is operated semi-automatically and differs from the gravity filler principally in that the liquid is under pressure. It is usually equipped with an overflow tube connected to a receiver to prevent excess filling of the container.

Vacuum filling is commonly used in faster filling lines for large liquid volumes because it is more adaptable to automation. A vacuum is produced in a bottle when a nozzle gasket makes a seal against the lip of the bottle to be filled. The vacuum draws the liquid from a reservoir through the delivery tube into the bottle. When the liquid level reaches the level of an adjustable overflow tube, the seal is mechanically loosened and the vacuum released. Any liquid that had been drawn into the vacuum line is collected in a trap receiver and then returned to the reservoir.

It is obvious that the accuracy and precision of machine filling of sterile liquids vary with the method. Therefore, a method is selected to provide the degree of accuracy and precision required by the nature of the product. A slight excess is required in each container to provide for the loss that occurs at the time of administration due to adherence to the wall of the container and retention in the syringe and hypodermic needle lumen. A table of suggested excess volumes is found in the USP.

The danger of overdosage as well as economic factors limit the amount of excess desirable in a given container. A reduction of only 0.01 ml of unnecessary excess in each 1 ml ampoule of a lot of 10,000 would yield approximately 100 more containers of the product.

Emulsions and suspensions often require specially-designed filling equipment because of their high viscosity. To obtain a reasonable flow rate, high pressures must be applied, or containers with large openings must be used, to permit the entry of large delivery tubes. Sometimes, jacketed reservoir tanks can be used to raise the temperature of the product, and thereby lower its viscosity. It is normally necessary to keep suspensions, and sometimes emulsions, constantly agitated in the reservoir during filling, so that the product remains homogeneous and each subdivided unit contains the required amount of drug.

## *Filling Equipment for Solids*

Sterile solids, such as antibiotics, are more difficult to subdivide accurately and precisely into individual-dose containers than are liquids. The rate of flow of solid material tends to be slow and irregular, particularly if finely powdered. Small, granular particles flow most evenly. Containers with a relatively large opening must be used; even so, the filling rate is slow, and the risk of spillage is ever present. For these reasons, the tolerances permitted for the content of such containers must be relatively large. Suggested tolerances may be found tabulated in the USP.

Sterile solids can be subdivided into containers by individual weighing. The operator can use a scoop that holds a volume approximately equal to the weight required, but the quantity filled into the container is finally weighed on a balance. This is a slow process.

When the solid is obtainable in a relatively free-flowing form, machine methods of filling may be employed. In general, these methods involve the measurement and delivery of a volume of the solid material, which has been calibrated in terms of the weight desired. Among the major problems in the use of such machines are stratification of particles due to varying particle sizes, development of electrostatic charge within the mass of dry solid particles, formation of air pockets, and uneven flow due to clumping of the particles. These all result in uneven filling of the container. The problems usually, can be minimized if uniform particle size of the solid is achieved and a small electric current is used to neutralize the developing charge.

One type of machine for delivering measured quantities of free-flowing material employs an auger in the stem of the funnel-shaped hopper (Fig. 23.10). The size and rotation of the auger can be adjusted to deliver a regulated volume of granular material from the funnel stem into the container.

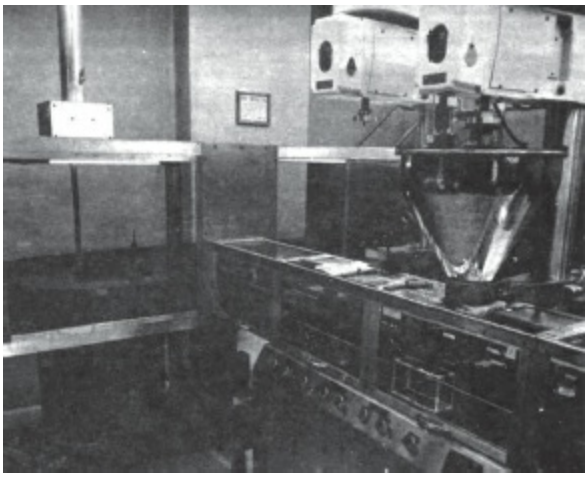
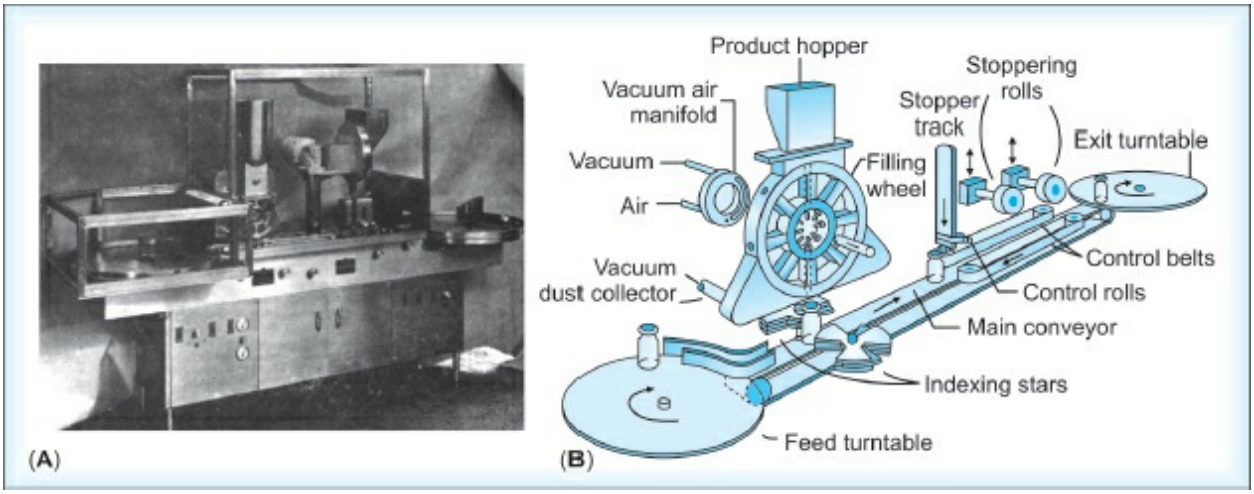


**Fig. 23.10:** High-speed automated aseptic processing line protected by a hood showing unscrambler turntable at left and powder filler at right

In another filling machine (Fig. 23.11), an adjustable cavity in the rim of the filling wheel is filled by vacuum as the wheel passes under the hopper. The contents are held by vacuum until the cavity is inverted over the container, when a jet of sterile air discharges the dry solids. This machine dispenses dry solids that flow less freely than those of other machines presently available.




**Figs 23.11A and B:** (A) Automated filling unit for dry solids. The operation is protected by hoods to maintain an aseptic environment; (B) The diagram represents the functioning parts with the unscrambler turntable at left, the filling machine, the rubber closure inserting device, and the collecting turntable for the filled and stoppered vials at right

## *Sealing*

Containers should be sealed in the aseptic area immediately adjacent to the filling machine. In addition to retaining the contents of a sterile product, sealing of containers assures the user that it has not been opened. It is obvious that a sterile container that has been opened can no longer be considered to be sterile. Therefore, tamper-proof sealing is essential.

## *Sealing Ampoules*

Ampoules may be closed by melting a portion of the glass of the neck to form either bead-seals (tip-seals) or pull-seals. Tip-seals are made by melting sufficient glass at the tip of the ampoule neck to form a bead of glass and close the opening. Pull-seals are made by heating the neck of a rotating ampoule below the tip, then pulling the tip away to form a small, twisted capillary just prior to being melted closed.

The heating with a high-temperature gas-oxygen flame must be even and carefully controlled to avoid distortion of the seal. Excessive heating of air and gases in the neck causes expansion against the soft glass with the formation of fragile bubbles at the point of seal. Open capillaries at the point of seal or cracks result in “leakers”. Pull-sealing is a slower process, but the seals are more reliable than those from tip-sealing. Powder ampoules or other types having a wide opening must be sealed by pull-sealing.

Fracture of the neck of ampoules often occurs during sealing if wetting had occurred at the time of filling. Also, wet glass at the neck increases the frequency of bubble formation and of unsightly and contaminating deposits of carbon or oxides as a result of the effect of the heat of sealing on the droplets of product.

With some sensitive products, it may be necessary to close the ampoules with pull-seals to prevent combustion products of the flame from entering the ampoule at the time of sealing, as might occur with tip-sealing. In addition, it is sometimes necessary to displace the air in the space within the ampoule above the product to prevent decomposition. This may be done by introducing a stream of inert gas, such as nitrogen or carbon dioxide, during or after filling with the product. Immediately thereafter, the ampoule is sealed before the gas can diffuse to the outside.

## *Sealing Bottles, Cartridges and Vials*

Rubber closures must fit the opening of the container snugly enough to produce a seal, but not so snugly that it is difficult to position them in or on the container. They may be inserted by hand, using sterile forceps (*see* Fig. 23.9). A faster hand method involves picking up the closure and inserting it into a vial by means of a tool connected to a vacuum line.

When closures are to be inserted by machines, the surface of the closure is usually halogenated or coated with silicone to reduce the friction. This makes it possible for a closure to slide from a rotating or vibrating drum to the bottom of a chute, where it is positioned over a container, ready for insertion by a plunger or some other pressure device (Fig. 23.11). Stoppering can be done at productionline speeds with such a machine.

Aluminum caps are used to hold rubber closures in place. Single caps may have a permanent center hole or a center that is torn away at the time of use to expose the rubber closure. Double aluminum caps usually have an inner cap with a permanent center hole, which in use is exposed when the entire outer cap is torn off. The triple aluminum caps are used for large bottles with rubber closures having permanent holes for attachment to administration sets. The inner cap with a permanent center hole remains in place during use to secure the rubber closure. The thin disc is used in conjunction with a thin rubber disc to seal the holes through the closure. The outer cap holds the disc in place and is torn away at the time of use.

When applied, the bottom edge of an aluminum cap is bent (crimped) around and under the lip of the glass container. It cannot be removed without destroying the cap, but perforations permit tearing away the portions of the cap to be discarded preparatory to use. Single aluminum caps may be applied by hand crimping devices, but double-or-triple-caps, or large production lots require the use of heavy-duty motorized crimping machines.

## *Automation of Processing*

The need for increased production rates for sterile products eventually justifies the cost of developing and operating automatic machinery. Unquestionably, machines can be designed to carry out certain operations more rapidly and with more reliable repetition than can be performed by people. A further advantage of mechanization in processing sterile products is the elimination of the human body as a source of biologic contamination; however, the contaminating effects of abraded particles, lubricants, and dirt

from the moving parts of even the cleanest aseptic machine must not be forgotten.

When machines are designed or used so that the constant attention of a human operator is required, the operation is identified as being semiautomatic. For automatic operation, machines are usually linked together by conveyor belts in an arrangement that requires little attention from an operator (Fig. 23.11). The unscrambler or feed turntable lines up a large number of vials and feeds them, one at a time, to the conveyor belt. The belt carries each vial in sequence to the filling wheel, to the stoppering machine, and then out to the collecting turntable. A crimping machine could be inserted after the stoppering machine. This processing line has been permanently linked together. For greater flexibility, each machine may remain a separate unit, linked together in use by conveyor belt units in an arrangement suitable for a particular process.

Automation of the entire process would convey an empty dose container from its supply carton through the entire process until it is filled with a product, labeled, and placed in the shipping carton. A portion of such a processing line is illustrated in Fig. 23.10. In this line, the vials are removed by operators from their protective cartons or “shrink-packs” and given an air rinse. A covered conveyor carries them at high speed through a sterilizing tunnel. They emerge, cooled, on the unscrambler turntable (left, Fig. 23.10) and are conveyed in sequence to the dry solids filler (right, Fig. 23.10), stoppering and capping machines, and then into the next room for packaging, all without contact with human hands. A compact line of similar concept for processing small ampoules and vials has been developed. Such lines are usually located in an aseptic room with the critical portions further protected with covers and bathed with HEPA-filtered air. An automated processing line, such as this, can be justified when high production rates are required.

## *Sterilization of Product*

A product must be sterilized by the most reliable method possible. The methods of sterilization, their application, and the bases for their selection have been discussed in Chapter 22.

**Freeze-drying:** Freeze drying (lyophilization) is a drying process applicable to the manufacture of certain pharmaceuticals and biologicals that are thermolabile or otherwise unstable in aqueous solution for prolonged

storage periods, but are stable in the dry state.

A product to be freeze-dried is prepared and handled as an aqueous solution or suspension in the same manner as discussed previously for an aseptic fill. The aqueous preparation is frozen rapidly and cooled to an experimentally determined temperature below its eutectic point. Most commonly, it would be frozen by a mechanical refrigeration device, often the refrigerated shelves in the freeze-drying chamber, at a temperature of −50°C or lower.

When the product is completely frozen and properly cooled, the chamber is sealed and evacuated. The ice in the frozen product gradually sublimes from the frozen surface and is collected in a refrigerated condenser chamber or on plates within the chamber containing the product. As the ice leaves the product, the drying residue maintains essentially its original volume and becomes porous, owing to the loss of ice molecules. This porous structure usually increases the subsequent rate of solution of the product as compared to the original material. The rate of drying depends largely on the thermal conductance of the frozen product, rate at which the vapour can diffuse through the progressively thicker layer of dried porous material, and the rate of transfer of the vapour through the system to the condenser surface. It has been said that the drying rate can be estimated as approximately one hour for each millimeter of depth of the product. The theory of freeze-drying is more fully discussed in Chapter 3 and in the literature.

In production, large freeze-driers are usually operated by an automatic control system. By means of a thermocouple frozen in a sample of the product, the temperature curve of the sample is duplicated in comparison with a curve experimentally found to produce a satisfactory product. During processing, the shape of the sample curve is adjusted primarily by the heat input, which provides the energy for sublimation.

The product is usually processed until there is less than 1% moisture in the dried material. After completion of the drying cycle, reabsorption of moisture must be prevented. The product must, therefore, be removed from the chamber and sealed as rapidly as possible under controlled low-humidity conditions.

Freeze driers also may be equipped for stoppering vials within the drying chamber. Special slotted rubber closures (*see* Fig. 23.2) are partially inserted into the neck of vials prior to freezing the product. The slots permit the

escape of water vapour from the vials during the drying cycle. At the end of the drying cycle a hydraulically-operated plate or an expandable rubber diaphragm presses the closures firmly into the neck of the vials and seals them under vacuum. A butyl rubber compound is usually used for these closures because of its low water vapour permeability.

Numerous biologic preparations, tissue sections, and viable microorganisms are being preserved in the freeze-dried state. Multiple vitamin combinations, antibiotics, and hormone preparations are examples of pharmaceutical products preserved by this method.

## Packaging

A thorough discussion of the packaging of sterile products is beyond the scope of this chapter. A few pertinent facts are presented, however. The package is an extremely important part of the product, for it presents the product to the user. It must be particularly dignified, neat, and attractive in appearance if it is to convey to the user the quality, purity, and reliability that conformance to the aforementioned principles and procedures is an inherent part of the product. The package also must accurately and completely provide the user with the information necessary for its use. The labeling should be legible and the identity and strength of the drug clearly distinguishable. This is particularly difficult with the small containers used for many sterile products. Furthermore, the package should protect the product against physical damage during shipping, handling, and storage, and should protect light-sensitive substances from ultraviolet radiation.

Further packaging requirements for injections are given by the USP essentially as follows:

1. The volume of an injection in single-dose containers should provide the amount specified for administration at one time and in no case is more than one liter. This requirement is intended to minimize the likelihood of someone attempting to use at a later time, a residue in a container after exposure to contamination from the environment.
2. Preparations intended for intraspinal, intracisternal, or peridural administration should be packaged only in single-dose containers because of the sensitivity of nerve tissue to irritation from added substances such as antibacterial agents.
3. Normally, no multiple-dose container shall contain a volume of injection more than is sufficient to permit the withdrawal of 30 ml, because larger volumes would provide for the withdrawal of more doses, thereby increasing the potential for contamination.
4. Injections labeled for veterinary use are exempt from the above limitation to singledose containers, and to the volume of multiple-dose containers because large animals require larger doses than man.

Details of labeling content requirements for injections can be found in the USP. When a label is applied to a container, it must be so arranged that a

sufficient area of the container remains uncovered to permit inspection of the contents. While these packaging and labeling stipulations apply specifically to official injections, the FDA looks upon these as basic guidelines for all products.

The operation of the packaging department for sterile products is essentially the same as for other pharmaceutical packaging departments. The overriding objective must be that every unit is properly labeled and packaged, with adequate controls to be assured that this is accomplished.

## Stability

General principles of evaluation for stability are discussed in Chapter 28. In this chapter, it is sufficient to mention that the previous consideration of ingredients and packaging components in intimate contact with the product, and some of the problems associated with formulation, draw attention to the necessity of evaluating the effect of all the components on the stability of the product, particularly when the product is subjected to the accelerating reactivity of thermal sterilization.

## Quality Control

The responsibilities of the quality control department have been discussed elsewhere in this text (Chapter 29). The discussion here is limited to those aspects of this important function that are peculiar to sterile products.

The three general areas of quality control are incoming stock, manufacturing (processing), and the finished product. For sterile products, incoming stock control encompasses routine tests on all ingredients as well as special evaluations such as pyrogen tests on WFI, glass tests on containers, and identity tests on rubber closures. It also may be necessary to perform microbial load (bioburden) tests to determine the number and types of microorganisms present. Process control in the manufacture of sterile products involves all of the innumerable tests, readings, and observations made throughout the manufacturing process of a product, including conductivity measurements during the distillation of WFI, confirmation of volume of fill in product containers, recording of cycle time and temperature for thermal sterilization of the product, and confirming the count and identity of labels for the product. The production control includes all of the final assays and tests to which the product is subjected. In addition to the usual chemical and biologic tests, a sterile product is subjected to a leak test (when applicable), a clarity test, a pyrogen test (when applicable), and a sterility test.

### *Leak Test*

Ampoules are intended to provide a hermetically sealed container for a single dose of a product, thereby completely barring any interchange between the contents of the sealed ampoule and its environment. Should capillary pores or tiny cracks be present, microorganisms or other dangerous contaminants may enter the ampoule, or the contents may leak to the outside and spoil the appearance of the package. Changes in temperature during storage cause expansion and contraction of the ampoule and contents, thereby accentuating interchange if an opening exists.

The leak test is intended to detect incompletely-sealed ampoules so that they may be discarded. Tip-sealed ampoules are more likely to be incompletely sealed than are those that have been pull-sealed. In addition, small cracks may occur around the seal or at the base of the ampoule as a result of improper handling.

Leaks usually are detected by producing a negative pressure within an incompletely sealed ampoule, usually in a vacuum chamber, while the ampoule is entirely submerged in a deeply colored dye solution (usually 0.5 to 1.0% (w/v) methylene blue). Subsequent atmospheric pressure then causes the dye to penetrate an opening, being visible after the ampoule has been washed externally to clear it of dye. The vacuum (27 inches Hg or more) should be sharply released after 30 min. Only a tiny drop of dye may penetrate a small opening.

A reported study has shown that detection of leakers is more effective when the ampoules are immersed in a bath of dye (FD and C Red No. 1, 0.5%; FD and C Red No. 2, 0.1%; and sodium lauryl sulfate, 0.25% in water) during the autoclaving cycle. This has the added advantage of accomplishing both leak detection and sterilization in one operation. Capillaries of about 15 μm in diameter or smaller may or may not be detected by these test methods.

Vials and bottles are not subjected to such a leak test because the rubber closure is not rigid; however, bottles are often sealed while a vacuum is being pulled so that the bottle remains evacuated during its shelf-life. The presence of a vacuum may be detected by striking the base of the bottle sharply with the heel of the hand to produce the typical “water hammer” sound. Another test is to apply a spark tester probe to the outside of the bottle, moving from the liquid layer into the air space. A blue spark discharge occurs if the airspace is evacuated.

## *Clarity Test*

Clarity is a relative term, the meaning of which is markedly affected by the subjective evaluation of the observer. Unquestionably, a clean solution having a high polish conveys to the observer that the product is of exceptional quality and purity. It is practically impossible, however, to prepare a lot of a sterile product so that every unit of that lot is perfectly free from visible particulate matter, i.e. is, from particles that are 30 to 40 μm and larger in size. Consequently, it is the responsibility of the quality control department to detect and discard individual containers of a product that the ultimate user would consider to be unclean. Further, the USP states that good pharmaceutical practice requires that all containers be visually inspected and that any with visible particles be discarded. In addition, for large-volume infusions, the USP has established a limit of 50 particles of 10 μm and larger

and 5 particles of 25 μm and larger per milliliter.

Normally, manufacturers of parenteral solutions can meet the above standard with the technology currently available. Further, product development, raw material quality control, and process control have eliminated any potential for widespread particulate development. In any lot of product, however, there may be a few containers having visible particles. These must be detected by visual inspection and discarded.

Although particulate matter is of primary concern in products given intravenously, all parenteral products should be free from insoluble particles. Several years ago, it was shown that the formation of pathologic granulomas in vital organs of the body can be traced to fibres, rubber fragments, and other solids present in i.v. solutions. More recently, a double blind study has shown that the use of a final filter for the administration of i.v. solutions reduced the incidence of thrombophlebitis, a result believed to be due to the elimination of subvisual particles from the administered solution. While there are unanswered questions related to these toxic effects, these findings have highlighted the importance of the preparation of exceptionally clean parenteral products.

Suspensions, emulsions, or dry solids, in addition to solutions, should be compounded and processed under clean conditions to minimize the presence of foreign particles.

The visual inspection of a product container is usually done by individual human inspection of each externally clean container under good light, baffled against reflection into the eyes, and viewed against a black and white background, with the contents set in motion with a swirling action, since a moving particle is much easier to see than one that is stationary. But, care must be exercised to avoid introducing air bubbles, which are difficult to distinguish from dust particles. To see heavy particles, it may be necessary to invert the container as the final step in inspection. Although a human inspector is subject to reduction in efficiency by eye strain, fatigue, distractions, and emotional disturbances, visual inspection can be done on 100% of the product units and can be done at a level of discrimination at least equal to that of the user.

Instrumental methods of evaluation for particulate matter in liquids utilizing the principles of light scattering, light absorption, and electrical resistance have been used to obtain particle counts and size distribution. All 

of them, however, require destruction of the product unit to obtain the test sample, making them useful only for quality control testing. A method utilizing video image projection coupled with electronic circuitry detects moving particles without destruction of the product unit. Therefore, it can be used for in-line detection of particles in product units, but at present, its use is limited to 1 to 5 ml containers.

## *Pyrogens and Pyrogen Test*

## *Pyrogens*

Water used in parenteral and irrigating solutions should be free of pyrogens. To achieve this, proper controls must be maintained in the preparation and storage of water.

Pyrogens are products of metabolism of microorganisms. Most bacteria and many molds and viruses have been reported as producing pyrogens. The gram-negative bacteria produce the most potent pyrogenic substances as endotoxins. Chemically, pyrogens are lipid substances associated with a carrier molecule, which is usually a polysaccharide but may be a peptide. About 1 h after injection into man, pyrogens roduce a marked rise in body temperature, chills, body aches, cutaneous vasoconstriction, and a rise in arterial blood pressure. Antipyretics eliminate the fever, but not the other systemic effects of pyrogens.

The fever response to pyrogens in rabbits is the basis for the official pyrogen test, which is described later in this chapter. For further information, the reader is referred to the extensive reviews on the nature and significance of pyrogens that have appeared in the literature.

## *Source and Elimination of Pyrogen Contamination*

Pyrogens may enter a product by any means that may introduce microorganisms or the products of their growth. The most likely sources are water, contaminated solutes, and containers. Water is free from pyrogens if it has been distilled so that the condensed molecules have gone through the vapour state protected from inadvertent contamination, and if the distillate has been collected and stored in a sterile condition. To be pyrogen-free, solutes must be prepared from vehicles free from pyrogens, and must be stored in a manner designed to prevent subsequent contamination. Opened

containers of solutes, capable of supporting the growth of microorganisms, invite such contamination. Containers may be rendered free from pyrogens by adequate cleaning and heating, usually at 210°C for 3 to 4 h. Studies also have shown that heating at 650°C for 60 s destroys pyrogens; however, autoclaving temperatures do not destroy pyrogens during a normal cycle.

Pyrogens sometimes can be removed from solutions by adsorption on the surface of select adsorbants, but the often concurrent phenomenon of adsorption of solute ions or molecules may prevent the use of such a method. Selective solvent extraction methods are useful in the production of antibiotics where heavy pyrogen contamination results from the fermentation process. New developments in ultrafiltration show promise of moving this process from limited research applications in molecular separations to practical production processes, which may include pyrogen separation and elimination. For most pharmaceutical preparations, however, it is better to prevent pyrogenic contamination than to attempt to remove pyrogens, a task that is difficult to accomplish without adversely affecting the product.

The product development department therefore must develop purity requirements for Water for Injection which are sufficiently stringent for its use as a vehicle in the product most sensitive to contaminants. Tests other than those for solids and pyrogenic content might be required, e.g. qualitative and quantitative tests for the presence of ions such as copper and iron.

## *Pyrogen Test*

The presence of pyrogenic substances in parenteral preparations is determined by a qualitative biologic test based on the fever response of rabbits. Rabbits are used as the test animal because they show a physiologic response to pyrogens similar to that of human beings. If a pyrogenic substance is injected into the vein of a rabbit, an elevation of temperature occurs within a period of 3 h. The specification limits and procedural details are given in the official test in the USP.

The housing conditions and handling are critical to obtaining consistent results with rabbits in the test. Because of this, the use of rectal thermometers has largely been replaced by rectal thermocouples, which remain in place throughout the test, eliminating the handling of rabbits for individual temperature readings. By this method, one person can handle 100 or more animals a day as compared to about 15 by the individual thermometer

method. Critical evaluations of pyrogen testing with rabbits may be found in the literature.

Many medical agents, if present, interfere with the test results because of their antipyretic or other interfering effects. Therefore, the pyrogen test is performed on all vehicles used for injections, but only on those finished products that do not interfere with the test, that have a high propensity for contamination with pyrogens, or that are given in large quantities. Considerably greater danger exists from the injection of large volume solutions containing pyrogens than from small volumes. Also, the pyrogenic effect is less with IM injection than with IV injection.

Recently, an in vitro test method for pyrogens has been developed utilizing the gelling property of the lysate of the amebocytes of *limulus polyphemus* (the horseshoe crab). In the presence of pyrogenic endotoxins from gram-negative bacteria, a firm gel is formed within 60 min when incubated at 37°C. Although only endotoxins from gramnegative bacteria react in this way, they constitute the majority and the most potent of contaminating pyrogens. The ***limulus amebocyte lysate (LAL)*** test has been found to be 5 to 10 times more sensitive than the rabbit test and by the use of serial dilutions has been shown to be semiquantitative. There do not seem to be large numbers of substances, other than proteins, that interfere with the reaction. The favourable results from many studies have brought increasing acceptance of the LAL test for in-process testing and selective product release testing. The USP now contains the specifications for a Bacterial Endotoxins Test designed to provide a means for estimating the concentration of bacterial endotoxins present in samples or on devices being tested. Debate is continuing on certain aspects of the test specifications, however, i.e. what the limits should be and how the tests are to be utilized?

## *Sterility Test*

All products labeled “sterile” must pass the sterility test, having been subjected to an effective process of sterilization. The test for sterility is intended for detecting the presence of viable form of microbes in pharmacopoeial preparations. Sterility testing of the products is carried out either by the Membrane Filtration method (Method A) or by the Direct Inoculation method (Method B). The technique of membrane filtration is used whenever the nature of the product permits; that is, for filterable

aqueous preparations, for alcoholic or oily preparations, and for preparations miscible with, or soluble in, aqueous or oily solvents, provided these solvents do not have an antimicrobial effect in the conditions of the test. The test for sterility is carried out under aseptic conditions with precautions taken to avoid contamination. The working conditions in which the tests are performed are monitored regularly by appropriate sampling of the working area and by carrying out appropriate controls.

## *Media for Sterility Test*

Sterility test media should be such that it supports growth of aerobes, anaerobes, and fungi. Fluid thioglycollate medium is primarily intended for the culture of anaerobic bacteria, however, it also supports the growth of aerobic bacteria. Soybean-casein digest medium is suitable for the culture of both fungi and aerobic bacteria. The composition of both the medium is shown in Table 23.7. Fluid thioglycollate medium is employed for clear fluid products whereas if the products are turbid and vicid then alternative thioglycollate medium is used. Alternative thioglycollate medium is fluid thioglycollate medium devoid of agar and resazurin.

**Table 23.7:** Composition of culture medium for sterility testing

| **Components** | **Culture medium** | |
|---|---|---|
| | **Fluid thioglycollate** | **Soybean–casein digest** |
| L-cystine | 0.5 g | – |
| Sodium chloride | 2.5 g | 5.0 g |
| Dextrose | 5.0/5.5 g | 2.3/2.5 g |
| Pancreatic digest of casein | 15.0 g | 17.0 g |
| Papaic digest of soya bean | – | 3.0 g |
| Dibasic potassium phosphate | – | 2.5 g |
| Granular agar (moisture < 15%) | 0.75 g | – |
| Yeast extract (water-soluble) | 5.0 g | – |
| Sodium thioglycollate | 0.5 g | – |
| or thioglycolic acid | or 0.3 ml | – |

| | | |
|---|---|---|
| Resazurin (0.10% (w/v) fresh solution) | 1.0 ml | – |
| Purified water | 1000 ml | 1000 ml |

## *Method A: Membrane Filtration*

A suitable unit consists of a closed reservoir and a receptacle between which a properly supported membrane of appropriate porosity is placed. A membrane suitable for sterility test has a nominal pore size not greater than 0.45 µm, diameter of approximately 47 mm and whose effectiveness to retain microorganisms has been established. Cellulose nitrate filters, for example, are used for aqueous, oily, and weakly alcoholic solutions; and cellulose acetate filters, are used for strongly alcoholic solutions. Specially adapted filters may be needed for certain products such as antibiotics.

Before the start of sterility test, the membrane is rinsed with diluting or rinsing fluid. For aqueous solutions or soluble solids the membrane is rinsed with sterile solution such as fluid A (pH 7.1 ± 0.2) consisting of 1 g of peptic digest of animal tissue in 1 L water. When oil, oily solutions, ointment and water-in-oil type emulsions are tested, the membrane is rinsed three times by filtering through it each time about 100 ml of a suitable sterile solution containing a suitable emulsifying agent such as polysorbate 80 at a concentration of 10 g/L (fluid D). Addition of surfactant, as in fluid D, makes it useful for formulations containing lecithin or oil. If the product has antimicrobial properties, the membrane is washed not less than three times by filtering through it each time the volume of the chosen sterile diluent. The rinsing should never exceed a washing cycle of 5 times 200 ml, even if such a cycle does not fully eliminate the antimicrobial activity. After rinsing the membrane, the contents of the container or product to be tested is transferred to the membrane. The selection of the amount and number of samples of the product are made according to Tables 23.8 and 23.9, respectively. Aqueous and oily solutions of sufficiently low viscosity may be filtered without dilution through a membrane. Viscous oils, ointments in a fatty base and emulsions of the water-in-oil type may be diluted as necessary with a suitable sterile diluent such as isopropyl myristate. Fluid products in pressurized aerosol form are freezed in an alcohol—dry ice mixture at least at −20°C and added to fluid D, prior to transfer of the contents. For pre-filled syringes

without attached sterile needles, the contents of each syringe are expelled into one or two separate membrane filter funnels or into separate pooling vessels prior to transfer. If a separate sterile needle is attached, the syringe contents are directly expelled. After filtration of the contents, the membrane is aseptically cut into two equal parts, and each one-half is transferred to both the culture medium. The Fluid thioglycollate medium is incubated at 32.5 ± 2.5°C, whereas soybean-casein digest medium is incubated at 22.5 ± 2.5°C, for not less than 14 days.

**Table 23.8:** Product quantity for sterility testing

| **Quantity per container** | **Minimum quantity to be used (unless otherwise justified and authorized)** |
|---|---|
| **Liquids (other than antibiotics)** | |
| Less than 1 ml | The whole contents of each container |
| 1–40 ml | Half the contents of each container, but not less than 1 ml |
| Greater than 40 ml, and not greater than 100 ml | 20 ml |
| Greater than 100 ml | 10% of the contents of the container, but not less than 20 ml |
| Antibiotic liquids | 1 ml |
| Other preparations soluble in water or in isopropyl myristate | The whole contents of each container to provide not less than 200 mg |
| Insoluble preparations, creams, and ointments to be suspended or emulsified | Use the contents of each container to provide not less than 200 mg |
| **Solids** | |
| Less than 50 mg | The whole contents of each container |
| 50 mg or more, but less than 300 mg | Half the contents of each container, but not less than 50 mg |

| | |
|---|---|
| 300 mg–5 g | 150 mg |
| Greater than 5 g | 500 mg |
| **Devices** | |
| Catgut and other surgical sutures for veterinary use | 3 sections of a strand (each 30 cm long) |
| Surgical dressing/cotton/gauze (in packages) | 100 mg per package |
| Sutures and other individually packaged single-use material | The whole device |
| Other medical devices | The whole device, cut into pieces or disassembled |

**Table 23.9:** Sample size for sterility testing

| **Number of items in the batch** | **Minimum number of items to be tested for each medium (unless otherwise justified and authorized)** |
|---|---|
| **Parenteral preparations** | |
| Not more than 100 containers | 10% or 4 containers, whichever is the greater |
| More than 100 but not more than 500 containers | 10 containers |
| More than 500 containers | 2% or 20 containers, whichever is less |
| For large-volume parenterals* | 2% or 10 containers, whichever is less |
| **Antibiotic solids** | |
| Pharmacy bulk packages (<5 g) | 20 containers |
| Pharmacy bulk packages (>5 g) | 6 containers |
| Bulks and blends | Bulk solid products |
| **Ophthalmic and other non-injectable preparations** | |
| Not more than 200 containers | 5% or 2 containers, whichever is the |

| | |
|---|---|
| | greater |
| More than 200 containers | 10 containers |
| If the product is presented in the form of single-dose containers, apply the scheme shown above for preparations for parenteral use | |
| **Devices** | |
| Catgut and other surgical sutures for veterinary use | 2% or 5 packages, whichever is the greater, up to a maximum total of 20 packages |
| Not more than 100 articles* | 10% or 4 articles, whichever is greater |
| More than 100, but not more than 500 articles | 10 articles |
| More than 500 articles | 2% or 20 articles, whichever is less |
| **Bulk solid products** | |
| Up to 4 containers | Each container |
| More than 4 containers, but not more than 50 containers | 20% or 4 containers, whichever is greater |
| More than 50 containers | 2% or 10 containers, whichever is greater |

* If the contents of one container are enough to inoculate the two media, this column gives the number of containers needed for both the media together

The media is examined for macroscopic evidence of microbial growth at intervals during the incubation period and at its conclusion. If no evidence of microbial growth is found, the product to be examined complies with the test for sterility. If evidence of microbial growth is found in the form of development of bacterial colonies, the product to be examined does not comply with the test for sterility, unless it can be clearly demonstrated that the test is invalid for causes unrelated to the product to be examined. The test may be considered invalid only if one or more of the following conditions are fulfilled (i) the data of the microbiological monitoring or sterility testing

facility or the the testing procedure used during the test reveals a fault and (ii) negative controls also show microbial growth. If the test is declared to be invalid, it is repeated with the same number of units as in the original test. If no evidence of microbial growth is found in the repeat test, the product examined complies with the test for sterility. If microbial growth is found in the repeat test, the product examined does not comply with the test for sterility.

## *Method B: Direct Inoculation*

Apart from testing oily solutions, creams, ointments and solid products, direct inoculation method is utilized particularly for surgical devices, sterile devices, surgical dressings and sutures, in case where membrane filtration method appears difficult. In this test, the quantity of the preparation to be examined, prescribed in Tables 23.8 and 23.9, is transferred directly into the culture medium so that the volume of the product is not more than 10% of the volume of the medium, unless otherwise prescribed. Oily liquids, creams and ointments are diluted to about 1 in 10 by fluid D. Sterile devices can be immersed intact in to the selected medium, while two or more portions from the innermost part of the surgical dressing are aseptically removed and then immersed intact in the unit. For catheters where the inside lumen and outside are required to be sterile, either cut them into pieces such that the medium is in contact with the entire lumen or fill the lumen with medium. If the product to be examined has antimicrobial activity, carry out the test after neutralizing this with a suitable neutralizing substance or by dilution in a sufficient quantity of culture medium. When it is necessary to use a large volume of the product, it may be preferable to use a concentrated culture medium prepared in such a way that it takes into account the subsequent dilution. Where appropriate, the concentrated medium may be added directly to the product in its container.

The inoculated media is then incubated for not less than 14 days. Observe the cultures several times during the incubation period. When thioglycollate medium is used for the detection of anaerobic microorganisms, keep shaking or mixing to a minimum in order to maintain anaerobic conditions. If the material being tested renders the medium turbid so that the presence or absence of microbial growth cannot be readily determined by visual examination, 14 days after the beginning of incubation transfer portions (each

not less than 1 ml) of the medium to fresh vessels of the same medium, and then incubate the original and transfer vessels for not less than 4 days. The results are interpreted as described previously under membrane filtration method.

# 24:
# Novel Drug Delivery Systems

Recent advances in the understanding of pharmacokinetic and pharmacodynamic behaviours of drugs have offered a more rational approach to the development of optimal drug delivery systems. In addition, it has now become apparently appreciable that future successes in drug delivery research will largely be the result of multidisciplinary efforts. Any therapeutic agent that can be more efficacious and safe using an improved drug delivery system represents both lucrative marketing opportunities for pharmaceutical companies and advances in the treatment of diseases of mankind. An ideally designed drug delivery system delivers a specified amount of drug to the target site at an appropriate time and rate as dictated or desired by the etiological and physiological needs of the body. Conventional pharmaceutical dosage forms are incapable of controlling the rate of drug delivery to the target site. As a result, the massive distribution of drugs in non-target tissues and body fluids necessitate therapeutic doses that could far exceed the amount required in target cells. The higher doses often lead to serious adverse effects during treatment. Thus, novel drug delivery systems (NDDS) are the carriers which maintain the drug concentration in therapeutic range for longer period of time and also, in addition, may deliver the content to the site of action if so desired as per requirements. There are several advantages of novel drug delivery systems over conventional drug delivery.

1. Optimum therapeutic-drug concentration in the blood or in a tissue may be maintained over a prolonged period of time.
2. Pre-determined release rates for an extended period of time may be achieved.
3. Duration for short half-life drugs may be increased.
4. By targeting the site of action, side effects may be eliminated.
5. Frequent dosing and wastage of the drug may be reduced or excluded.
6. Better patient compliance may be ensured.

In this chapter, we discuss various novel drug delivery systems and their potential applications in sustained, controlled and targeted delivery of therapeutic molecules.

## NOVEL DRUG DELIVERY SYSTEMS

Various drug delivery systems have been developed and some of them are under development with an aim to minimize drug degradation or loss, to prevent harmful side effects, and to improve drug bioavailability and also to favour and facilitate the accumulation of the drug in the required bio-zone (site). There are a number of novel carriers which have been established and documented to be useful for controlled and targeted drug delivery. It is important to critically evaluate different terms used under different broad categories of novel drug delivery systems:

- Sustained-or controlled-drug delivery systems provide drug action at a pre-determined rate by providing a prolonged or constant (zero-order) release, respectively, at therapeutically effective levels in the circulation.
- Localized drug delivery devices provide drug action through spatial or temporal control of drug release (usually rate-limiting) in the vicinity of the target.
- Rate-preprogrammed drug delivery systems provide drug action by manipulating the release of drug molecules by system design, which controls the molecular diffusion of drug molecules.
- Targeted drug delivery provides drug action by using carriers either for passive or active targeting or one based on self-programmed approach, usually anchored with suitable sensory devices, which recognize their receptor at the target.

Carrier systems for the targeted and controlled drug delivery purpose may be classified on the basis of their nature, mechanism of drug release and the nature of drug incorporation (Table 24.1 and Fig. 24.1). Diffusion occurs when bioactive agent is hydrophilic and passes through the polymer(s), the latter constitutes the key building block of controlled-release concept. Many environmentally-responsive systems are also designed that retain their content until appropriately placed in biological by an environment and are activated by an external or internal stimulus for the release of drug. Table 24.1 shows the mechanism of drug release from various drug-delivery systems. In the following paragraphs, some of these release systems are discussed in brief.

**Table 24.1:** Classification of sustained or controlled release systems based on their rate-control mechanism

| Type of system | Rate-control mechanism |
|---|---|
| **Diffusion-controlled** | |
| Reservoir systems (Ocusert®) | Diffusion through membrane |
| Monolithic systems (transdermal drug delivery systems-Nitro-dur®) | Diffusion through membrane |
| **Water penetration controlled** | |
| Osmotic systems (Oros®, Alzet® osmotic pump) | Osmotic transport of water through semipermeable membrane |
| Swelling systems (hydrogel) | Water penetration into glassy polymer |
| **Chemically-controlled** | |
| Monolithic systems | Either pure polymer erosion (surface erosion) or a combination of erosion and diffusion (bulk erosion) |
| Pendent systems | Combination of hydrolysis of pendent group and diffusion from bulk polymer |
| Ion-exchange resins | Exchange of acidic or basic drug with the ions present on resins |

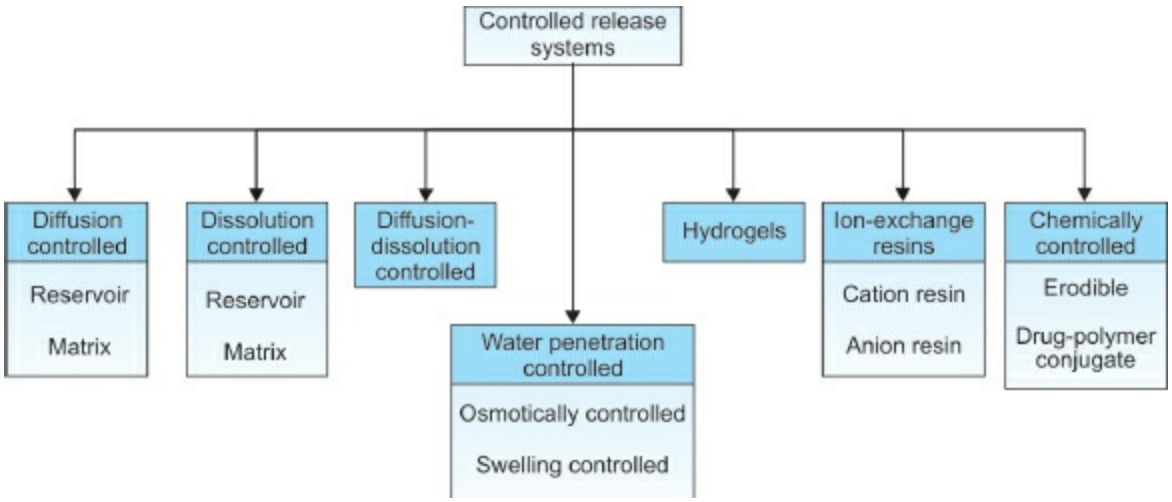


**Fig. 24.1:** Schematic depiction of various classes of controlled-release systems

## RESERVOIR-TYPE DRUG DELIVERY SYSTEM

In the reservoir-type drug delivery systems, drug is encapsulated in the drug reservoir compartment whose drug-releasing surface is covered by a rate-controlling an embryonic polymeric membrane. The drug in the reservoir compartment can be drug solid particles or a molecular dispersion of drug in a liquid- or solid-type dispersing medium. The polymeric membrane can be fabricated from a homogeneous or a heterogeneous non-porous polymeric material or semi-permeable membrane. The release of drug from this type of delivery systems occurs at a nearly constant rate (Q/t), as described in Eq. (1).

$$\frac{Q}{t} = \frac{K_{m/r} K_{a/m} D_d D_m}{K_{m/r} D_m + K_{a/m} D_d h_m} C_R \qquad \ldots (1)$$

where, $K_{m/r}$ and $K_{a/m}$ arethepartitioncoefficients for the interfacial partitioning of drug molecules from the reservoir to the membrane and from the membrane to the aqueous diffusion layer, respectively. $D_m$ and $D_d$ are the diffusion coefficients in the rate-controlling membrane with a thickness of $h_m$, and in the aqueous diffusion layer with a thickness of $h_d$, respectively. $C_R$ = drug concentration in the reservoir compartment.

### *Ocusert®*

A truly continuous, controlled-release and zero-order kinetic fashion was achieved using Ocusert®. First marketed by Alza Corporation, California, the pilocarpine Ocuserts® improved the non-compliance problems, low intra-ocular drug bioavailability and potential systemic side effects of pilocarpine. The system consists of a pilocarpine-alginate core (drug) sandwiched between two transparent, rate-controlling ethylene-vinyl acetate copolymer-based thin membranes. When this is placed under the upper or lower eyelid, the pilocarpine molecules after getting dissolved in the lachrymal fluid are released through the rate-controlling membranes at a pre programmed rate.

A mixture of pilocarpine and alginic acid in the drug reservoir releases the drug for almost one week. A thin membrane of ethylene-vinyl acetate (EVA) copolymer encloses the reservoir above and below. A retaining ring of the same material impregnated with titanium dioxide encloses the drug reservoir circumferentially (Fig. 24.2). The typical dimensions of the elliptical device are: major axis, 13.4 mm, minor axis, 5.7 mm; thickness, 0.3

mm. Two types of Ocusert® are available: Ocusert® Pilo-20 and Ocusert® Pilo-40. The Ocusert® Pilo-20 can release pilocarpine at a rate of 20 g/h for 7 days (total amount of drug released, 3.4 mg) and Ocusert® Pilo-40 at a rate of 40 g/h for 7 days (total amount of drug released, 6.7 mg). Ocusert® Pilo-20 contains 5.0 mg drug and Ocusert® Pilo-40, 11.0 mg of drug to maintain a constant release rate from the drug reservoir. Ocusert® Pilo-40 contains about 90 g of di-(2-ethylhexyl) phthalate as flux-enhancer to maintain the higher release rate (40 g/h).

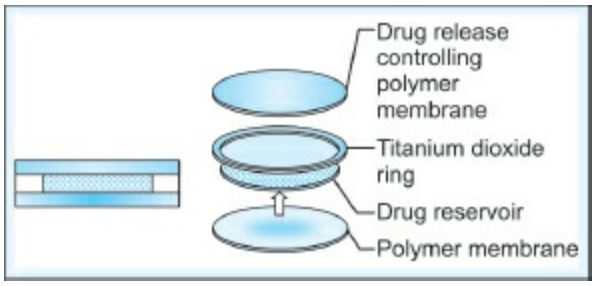



**Fig. 24.2:** Schematic representation of Ocusert®

## Matrix-type Drug Delivery Systems

In the matrix-type drug delivery systems, the drug is homogeneously dispersed in the lipophilic or hydrophilic polymer matrix. The drug dispersion in the polymer matrix may be produced by any of the following processes:

1. Blending finely-grounded drug particles with a viscous liquid (or a semisolid) polymer, followed by cross-linking of polymer.
2. Mixing drug solids with a melted polymer at an elevated temperature. The resultant drug-polymer dispersion is then moulded or extruded to form drug delivery devices of various shapes and sizes designed for a specific application.
3. It can also be fabricated by dissolving the drug and the polymer in a common solvent, followed by solvent evaporation, at an elevated temperature and/or under a vacuum, in a mould.

The release profile of drug from these type of drug delivery systems is not constant, because the rate of drug release is time dependent as defined by Eq. (2):

$$Q = \sqrt{(2AC_R D_p)} \times \sqrt{t} \qquad \ldots (2)$$

where,

$A$ = The initial loading dose of drug dispersed in the polymer matrix;

$C_R$ = Drug solubility in the polymer, which is also the drug reservoir concentration in the polymer matrix;

$D_p$ = Diffusivity of the drug molecules in the polymer matrix.

## *Nitro-Dur®: A Transdermal Drug Delivery System*

This system is designed for application onto intact skin to provide a continuous transdermal infusion of nitroglycerin, at a daily dose of 0.5 mg/cm$^2$, for the prevention of angina pectoris. This controlled-release transdermal therapeutic system is fabricated by first heating an aqueous solution of water-soluble polymer, glycerol, and polyvinyl alcohol, and then lowering the temperature of the mixture to form a polymer gel. Nitroglycerin/lactose triturate is dispersed in the gel, and the mixture is then

solidified at room temperature to form a medicated polymer disc by a moulding and slicing technique. After assembly onto a drug-impermeable metallic plastic laminate, a patch-type transdermal therapeutic system is developed with an adhesive rim surrounding the medicated disc (Fig. 24.3). The drug reservoir can also be formulated by directly dispersing the drug in an adhesive polymer, such as poly (isobutylene) or poly(acrylate) adhesive, and then spreading the medicated adhesive by solvent casting or hot melt, onto a flat sheet of drug-impermeable backing support to form a single- or multiple-layer drug reservoir.

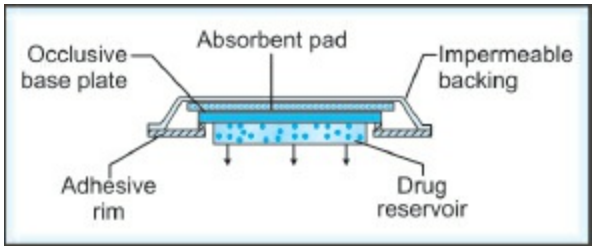



**Fig. 24.3:** Schematic view of Nitro-dur® system

## Osmotic Pumps

Osmotic pumps are controlled drug delivery devices based on the principle of osmosis. Wide spectrums of osmotic devices are available, out of which osmotic pumps are unique, dynamic and widely employed in clinical practice. Osmosis is a universal biophenomenon, which is exploited for the development of delivery systems with every desirable property of an ideal controlled drug delivery system. Osmotic pumps offer many advantages like easy formulation and simple operation, improved patient compliance with reduced dosing frequency, more consistancy and prolonged therapeutic effect with uniform blood concentration. Moreover, they are inexpensive and their industrial adaptability vis-a-vis production scale-up is easy.

An elementary osmotic pump was developed by Alza under the name OROS®, for controlled release oral drug delivery formulations. It has been discussed earlier in Chapter 16 (Sustained Release Oral Dosage Forms).

Second type of osmotic device is a small osmotic pump sold under the trade name Alzet® (Alza Corporation) that is implanted in the tissue of animals, where it delivers a chosen therapeutic agent at predetermined rate. The active agent is placed in an impermeable flexible wall reservoir that is surrounded and sealed within a rigid cellulose acetate membrane. When the device comes in contact with an aqueous environment, water is osmotically driven across the cellulose acetate membrane and the resultant pressure on the reservoir wall forces the agent out of the orifice (Fig. 24.4).

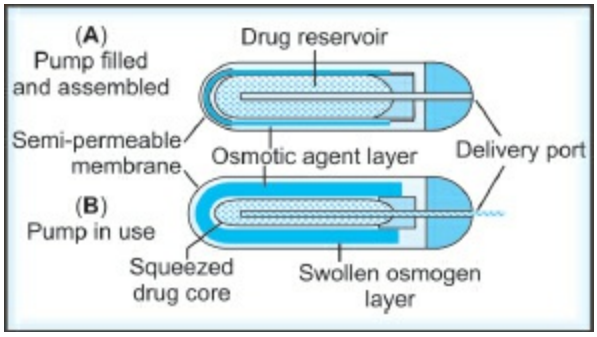
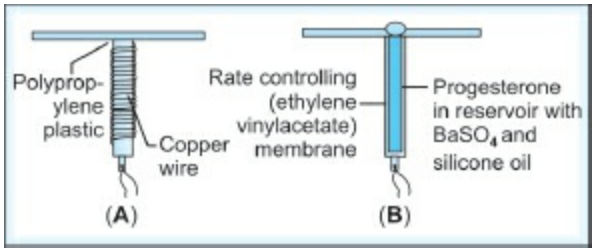



**Fig. 24.4:** Alza mini-osmotic pump (Alzet®)

## Intrauterine Devices (IUDs)

Intrauterine devices belong to novel therapeutics, commonly used for the delivery of contraceptive steroidal hormones. Medicated IUDs are those that serve as carriers or vehicles for pharmacologically-effective antifertility agents, such as the copper-bearing IUD and the progesterone-releasing IUD. The advantages of these delivery systems include prolonged drug release, minimum side effects and increased bioavailability. Copper and other metals such as zinc, cadmium, lead, etc., have been reported to enhance the contraceptive effectiveness of IUDs. Clinical studies on a T-shaped polyethylene plastic device with 30 $mm^2$ of copper wire (CU-T30) (Fig. 24.5A) showed that the pregnancy rate was declined upto 5% from the 18% level as achieved with non-medicated T-shaped device. It has been reported that the copperbearing IUD is more efficacious when the copper wire is located on the transverse arm of the device, which remains in contact with the fundus after insertion. This fact leads to the development of new generation of copperbearing T-shaped IUDs, CU-T 380A having two sleeves of copper located on the transverse arms of the letter "T". The additional surface area of 30 $mm^2$ and relatively small surface area of copper (2 × 30 $mm^2$) comes in close contact with the upper portion of the endometrial cavity resulting into significantly enhanced anti-fertility efficacy of copperbearing IUD.

In progesterone-releasing device (Progestasert® IUD), the drug reservoir exists as a dispersion of progesterone crystals in silicone fluid encapsulated in the vertical limb of a T-shaped device walled by a non-porous membrane of ethylene-vinyl acetate copolymer (Fig. 24.5B). It is engineered to release continuously a daily dose of 65 mg progesterone inside the uterine cavity to achieve contraception for one year.




**Figs 24.5A and B:** Intrauterine devices (IUDs): (A) Copper-bearing IUD; (B) Progestasert®, progesterone releasing IUD

## Implants

Implants are the drug delivery systems which act as depot or reservoir of loaded drug and deliver it over a prolonged period of time in a controlled manner upon insertion into the body. Briefly, the intramuscular, subdermal, intracranial or other organospecific depots are largely based on implants which either limit high drug concentrations to the immediate area surrounding the pathology or to provide sustained drug release for systemic therapy. Implantable drug delivery devices are devoid of limitations associated with oral, IV, and topical drug administration. The subcutaneously implantable drug delivery devices offer one unique advantage of retrieving and recovery opportunity. This feature enables a readily reversible termination of drug delivery whenever so required.

Implants can be classified into transdermal implantable systems, subdermal implantable systems, and implantable polymeric matrices. Clinically, implant systems are recommended in situations where chronic therapy is indicated, such as hormone replacement therapy and chemical castration in the treatment of prostate cancer. Parenteral implants can be highly viscous liquids or semisolid formulations, both of which may be injected with the help of a needle. Alternatively, implants may be in the form of tiny rods impregnated with drug substances or a liquid, which gels following administration. In in situ forming gels, either gel on diffusion of polymer solubilizing solvent from the injection site, leaving the polymer in contact with an aqueous environment in vivo or may gel on cooling after they are injected in the body at an elevated temperature. In situ gels may be used to prepare sustained-or controlled-release formulations of the macromolecules such as proteins and oligonucleotides.

Implants intended for parenteral administration are prepared from a variety of polymeric materials including polysaccharides, polylactide-co-glycolide, and the non-biodegradable methacrylates. Biodegradable materials, such as polylactide-co-glycolide, are preferred as they exclude the need for surgical removal of the implant after treatment ended. However, non-biodegradable materials do provide therapeutic levels of the drug for up to one year in vivo. Solid implants typically exhibit biphasic release kinetics, with an initial burst of drug followed by a slower release. The initial burst is usually due to the release of drug deposited on the surface of the implant,

although zero-order release kinetics may be achieved by, for example, coating the implant with a drug-impermeable polymeric material. Varying polymer composition may vary or control drug release. Drug release could be biophysically modulated using stimuli, to which a system responds, such as electrical stimuli in polyelectrolyte systems. Solid implants avoid the peak levels associated with the administration of the drug in conventional solution, thus limiting the toxic effects associated with the free drug. Implants are used for the delivery of anticancer agents, as they are able to confine potentially-toxic anticancer drugs to tumour sites and also allow sustained drug release. In order to achieve high drug doses in traditionally inaccessible areas, such as the central nervous system, bone tissue and beyond the blood-retinal barrier, implants could effectively be used clinically. For example, ethylene vinyl acetate copolymer dexamethasone intracranial implants have been reported to produce high drug levels in the brain, however, at the same time plasma levels remain normal. Similarly, a polylactide-co-glycolide gentamycin bone implant demonstrated superior therapeutic regimen compared to an intramuscular injection of the drug in eradicating bone infections in a canine model. It could be concluded that implant systems, although have the obvious drawback of requiring administration through minisurgical procedures, offer a means of achieving high drug concentrations in areas that are usually inaccessible to peripherally or vascularly-administered drug. In addition, high drug levels are maintained in a sustained manner in these areas.

## Hydrogels

Hydrogels normally consist of hydrophilic polymers that are covalently cross-linked to prevent their dissolution. Depending on the polymer used, these systems can imbibe water content ranging from 30 to 90%. Because of their highly swollen nature, hydrogel membranes are usually quite permeable to hydrophilic or high-molecular weight agents. This property, together with good biocompatibility, has popularized the use of hydrogels as rate-controlling membranes in devices for delivering proteins, such as insulin, aprotinin, tumour antigenesis factor, and leuteinizing hormone. Low-molecular weight, water-soluble drugs often permeate across the hydrogel matrices at a high rate which limits their usefulness, yet hydrogel membranes have found application as rate-controlling barriers for water-insoluble drugs, such as steroids. Almost all hydrogel controlled-release devices are based on monolithic systems, releasing the active agent following $t_{1/2}$, or first-order kinetics. Many hydrogels are prepared using free-radical polymerization, and ultraviolet light or gamma radiation can also initiate the reaction followed by propagation by free-radical reaction. The hydrophilicity of the gel is usually controlled by copolymerizing hydrophilic and hydrophobic monomers. The permeability of a hydrogel is determined by the extent of crosslinking, the degree of hydration of the gel, the nature of the permeant, and the device design. Hydrogels may also be prepared in the absence of water, and subsequently equilibrated with water or a concentrated aqueous solution of the active agent.

It has been proposed that water exists in two forms in gels: bound water closely associated with the polymer matrix (water of solvation) and bulk water lying between the polymer chains. However, due to the nature of the system, hydrogels still have a relatively high permeability for all water-soluble agents, upto a fairly large molecular size. The simplest form of controlled-release hydrogel is obtained by equilibrating the polymer, in either the dry or hydrated state, with a concentrated solution of the active material. Hydrogels have been used to prepare reservoir devices, usually by coating a high-permeability monolithic core containing dispersed or dissolved active material with a less-permeable coating.

Hydrogels have been used in the controlled-delivery of a variety of bioactive agents, such as contraceptives, ophthalmic, antibacterial agents,

drug antagonists, antiarrhythmics, anticancer drugs, anticoagulants, enzymes, and antibodies. They can be applied as inserts or implants, or can be administered subcutaneously, intramuscularly, or per-orally. In the development of controlled-release systems for ophthalmological applications, hydrogels have been used as ocular insert devices, delivering the drugs directly to the eye. The introduction of ionogenic groups may affect the rate of release of drugs from hydrogels. Increasing the content of acidic or basic groups in the hydrogel matrix may transform them into ion-exchange resins. Principally, ion-exchange resins, having the ability to swell in aqueous media, can be categorized as special type of hydrogel materials. Cationic and anionic ion-exchangers have been used in the past to prolong the effect of drugs. Their use was based on the principle that acidic or basic pharmaceuticals, combined with appropriate resins, yield insoluble polysalt resinates. Reticulated hydrogel sponges made up of cross-linked diesters, are particularly useful as general fluid sorbants, and those of amylose glutarate are good bioabsorbable hemostatic agents.

## In situ Gelling Systems

The field of in situ gelling system has grown exponentially in recent years. Liquid formulations generating a semisolid depot after administration into the body are attractive delivery systems. Being less invasive and painful compared to implants they can be administrated through various routes. Localized or systemic drug delivery can be achieved for prolonged periods of time, typically ranging from one to several months. These systems can be used for the sustained delivery of bioactives through various routes, such as parenteral, ocular and nasal, etc. Generally, parenteral depot systems could minimize side effects by achieving constant, 'infusion-like' plasma-level time profiles, especially important in the case of potent drugs with narrow therapeutic indices. From the manufacturing point-of-view, in situ depot forming systems offer the advantage of being relatively simple to manufacture from polymers. The gelling of polymers may occur by any process including ionic gelation due to bio-environment and pH and temperature conditions of the cavity.

Parenterally thermosensitive hydrogel, ReGel® (triblock copolymer PLGA-PEG-PLGA), is being used as a drug delivery carrier for the continuous release of human insulin. It has been observed that from the ReGel® formulations, steady amount of insulin secretion is achieved for up to 15 days. A biodegradable poly (ethylene oxide) and poly(L-lactic acid) gel is also formed, which exists as a sol at an elevated temperature (around 45°C), but forms a gel after subcutaneous injection and subsequent rapid cooling to body temperature. Poloxamers as thermogelling polymers can be applicable for the development of effective ophthalmic drug delivery systems. However, combination with Pluronic® analogs or addition of further polymer, e.g. PEG, PAA, methylcellulose (MC), HPMC, CMC is often required to reduce the concentration of polymer and/or to achieve a phase-transition temperature higher than the room temperature (25°C) and gelling at precorneal temperature (35°C). Aqueous solutions of PAA that transform into gels upon increase in pH may also be used as in situ gelling, ophthalmic drug-delivery systems. Other natural polymers, such as xyloglucan, have also been used for ocular delivery of therapeutics. Slightly viscous gellan gum solutions in low concentrations (<1%) show a marked increase in apparent viscosity, when introduced in the presence of physiological cations, without requiring

additional ions, more than 10–25% of those available in tear fluid. The pre-corneal contact times for drugs can thus be extended up to 20 h.

## Stimuli-sensitive or ‘Smart’ Polymers as Drug Delivery Systems

The term ‘stimuli-sensitive polymers’ refers to the polymeric systems, which demonstrate noticeable and sharp property changes in response to a small change in physical or chemical stimuli, such as pH, temperature, ionic concentration, solvents, mechanical stress and electric or magnetic field. These microscopic changes are prominent apparently at the macroscopic level and reversible on withdrawal of the stimulus. Stimuli-sensitive polymers are either synthetic or natural polymer that exhibit reversible conformational changes, which are reflected in their physicochemical properties such as solubility in response to environment. For this additional property, they are called environment-sensitive polymers, or more commonly “ intelligent” or “ smart” polymers. The environmentally-sensitive polymers used for designing and preparing drug delivery devices are listed in Table 24.2. As shown in Fig. 24.6, a smart polymer-based hydrogel can undergo swelling due to a change in the pH of the environment, resulting in easier diffusion of drug molecules through the expanded polymer network. Alternatively, the temperature of the environment can be increased to consolidate the temperature-sensitive hydrogels for faster release of the drugs through a squeezing action.

**Table 24.2:** Some environmentally sensitive polymers for drug delivery

| Stimulus | Hydrogel | Mechanism |
|---|---|---|
| pH | Acidic or basic hydrogel | Change in pH leads to swelling and subsequent release of drug |
| Thermal | Thermoresponsive hydrogel | Change in temperature alters polymer-polymer and water-polymer interactions leading to change in swelling profile and release of drug |
| Electrical | Polyelectrolyte hydrogel | Application of electric field charges membrane and results into electrophoresis |

| | | |
|---|---|---|
| | | of charged drug, alteration in swelling of gel and release of drug |
| Enzyme-substrate | Hydrogel containing | Enzyme acts on the substrate, the product formed immobilized enzymes changes swelling characteristics of gel and releases the drug |
| Magnetic field | Magnetic particles dispersed in alginate microspheres | Application of magnetic field alters the pores in gel leading to swelling and release of drug |
| Ultrasound irradiation | Ethylene-vinyl alcohol | Ultrasound irradiation increases temperature and leads to the release of drug |
| Ionic strength | Ionic hydrogel | Change in ionic strength alters the ionic concentration inside gel leading to swelling and release of drug |
| Chemical moieties | Hydrogel bearing electron-accepting groups | Electron donating compounds form charge transfer complex leading to alteration in swelling of gel and release of drug |

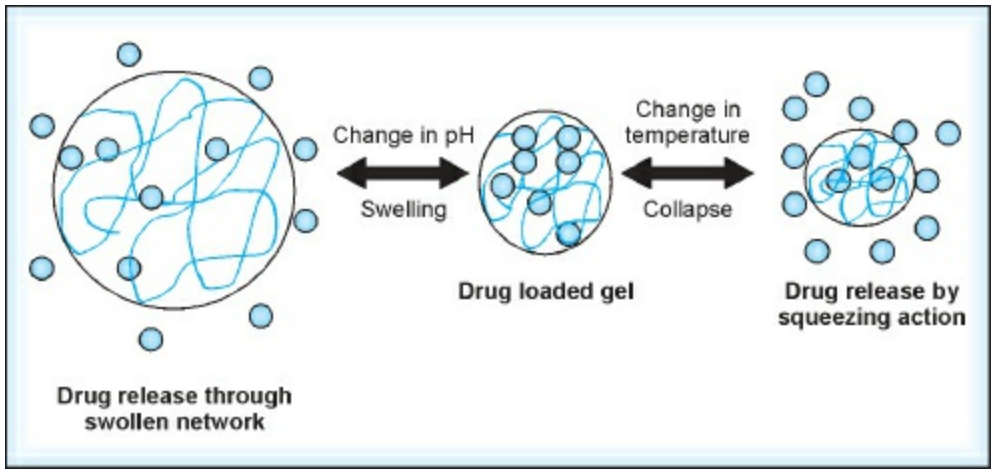



**Fig. 24.6:** Schematic representation showing the mechanism of drug release from stimuli-responsive polymeric carriers

## *pH-sensitive Polymers*

Polymers containing ionizable functional groups that respond to a change in pH are termed as pH-sensitive polymers. All the pH-sensitive polymers contain pendant acidic or basic groups that either accept or release protons in response to changes in environmental pH. When the external pH increases, swelling of hydrogel increases in the case of weakly-acidic (anionic) groups, but decreases in the case of weakly-basic (cationic) groups-containing polymers. Charge can be generated along the polymer backbone, which results in an increase in the hydrodynamic volume of the polymer. Several polymers including polyacrylic acid, polymethacrylic acid (PMAA), poly(ethylene imine), poly(L-lysine), and poly(N,N-dimethyl aminoethyl methacrylate) are typical examples of pH-sensitive polymers. pH sensitive polymers find applications in oral drug delivery. Rohm and Haas commercialized a series of polymers under the trade name Eudragit®. More recently, another series of polymers for the same applications under the trade name Eastacryl® (Eastman Kodak) and Kollokot® (BASF) have been commercialized.

## *Temperature-sensitive Polymers*

Thermosensitive or thermoprecipitating polymers are the polymers, which undergo phase transitions as a function of temperature. Temperature-sensitive hydrogels are the most extensively studied class of environment-sensitive polymer systems in drug-delivery research. These hydrogels are able to swell

or deswell as a result of variations in the temperature of the surrounding fluid. For convenience, temperature-sensitive hydrogels are classified into negatively thermosensitive, positively thermosensitive, and thermally reversible gels.

Negative temperature-sensitive hydrogels have a lower critical solution temperature (LCST) and contract upon heating above the LCST. Copolymers of (N-isopropylacrylamide) (NIPAAM) are usually used for negative temperature release. A positive temperature-sensitive hydrogel has an upper critical solution temperature (UCST), such hydrogel contracts upon cooling below the UCST. Polymer networks of poly(acrylic acid) (PAA) and polyacrylamide (PAAm) or poly (acrylamide-co-butyl methacrylate) have positive temperature dependence for swelling.

To date, many thermosensitive polymers have been reported for diverse applications. Poly(N-isopropylacrylamide) has attracted applications in pharmaceutical and biomedical technology as its LCST is close to body temperature. The most commonly used thermoreversible gels are those prepared from poly (ethylene oxide)-b-poly(propylene oxide)-b-poly(ethylene oxide) (Pluronics®, Tetronics®). Recently, new series of biodegradable triblock copolymers has been chemically designed. The polymers consisting of poly (ethylene glycol)-poly-(DL-lactic acid-co-glycolide)-poly(ethylene glycol) (PEG-PLGA-PEG) or PLGA-PEG-PLGA were investigated for sustained injectable drug delivery systems.

## *Electro-sensitive Polymers*

Like pH and temperature, electric current can also be used as an environmental signal to induce hydrogels for responses. Hydrogels sensitive to electric current are usually made up of polyelectrolytes. They are pH-sensitive hydrogels. Hydrogels of such polymers may undergo shrinking or swelling in the presence of an applied electric field. Sometimes, the hydrogels show swelling on one side and deswelling on the other side, resulting in bending of the hydrogels. The hydrogel shape change (including swelling, shrinking and bending) depends on a number of conditions. If the surface of hydrogel is in contact with the electrode, the result of applying electric field to the hydrogel may be different from systems where the hydrogel is placed in water (or acetone-water mixture) without touching the electrode.

Electro-sensitive hydrogels have been applied in controlled drug delivery. Hydrogels made of poly(2-acrylamido-2-methyl-propane sulfonic acid-co-n-butylmethacrylate) are able to release edrophonium chloride and hydrocortisone in a pulsatile manner using electric current. Chemomechanical shrinking and swelling of PMA hydrogels under an electric field is used for the pulsatile delivery of pilocarpine and raffinose. Microparticles of PAA hydrogel which showed rapid and sharp shrinkage with the application of electric current, recover their original size when the electric field is turned off. The electric field-induced changes in the size of the microparticles result in 'on-off' release profiles. The electric field-induced volume changes of poly(dimethylaminopropyl acryl-amide) hydrogels were accounted for pulsatile release of insulin.

## *Glucose-responsive Insulin Delivery*

Glucose-responsive delivery systems utilize the enzyme glucose oxidase which converts glucose into gluconic acid. Several approaches have been designed for chronic treatment of diabetes. Systems based on pH-sensitive polymers consist of immobilized glucose oxidase in a pH-responsive hydrogel encapsulating a saturated insulin solution. With the diffusion of glucose into the hydrogel, glucose oxidase converts it into gluconic acid that decreases the pH of the microenvironment of the hydrogel and subsequently causes gel swelling. The insulin permeates the swollen hydrogel more rapidly; hence relatively faster delivery of insulin in presence of glucose occurs. As the glucose concentration decreases in response to the released insulin, the hydrogel due to pH rise contracts and collapses to further retard or impede the rate of insulin delivery. Polyortho esters contain pH-sensitive linkages in the polymer backbone, which erodes as pH decreases. Insulin release from this polymer system can be modulated by glucose oxidase-glucose reaction. The acid sensitivity of the poly(orthoesters) can be further increased by incorporating tertiary amine functions in the polymer backbone (by using n-alkyldiethanolamines and triethanolamines). This pH sensitivity is due to the swelling of polymers induced by protonation of the tertiary amines in the polymer backbone.

Ito et al. developed a system using a porous cellulose membrane with surface-grafted poly (acrylic acid) as a sensitive membrane. On immobilization of glucose oxidase onto the poly(acrylic acid)-grafted

cellulose membrane, it becomes responsive to glucose concentration (Fig. 24.7). Basically in the absence of glucose, the chains of poly(acrylic acid) grafts are rod like, that reduce the porosity of the membrane and suppress insulin permeation, however in the presence of glucose, gluconic acid produced by glucose oxidase (GOD) which protonates poly(acrylic acid), making the graft chains coil-like thus opening the pores wide to enhance insulin permeation and release.

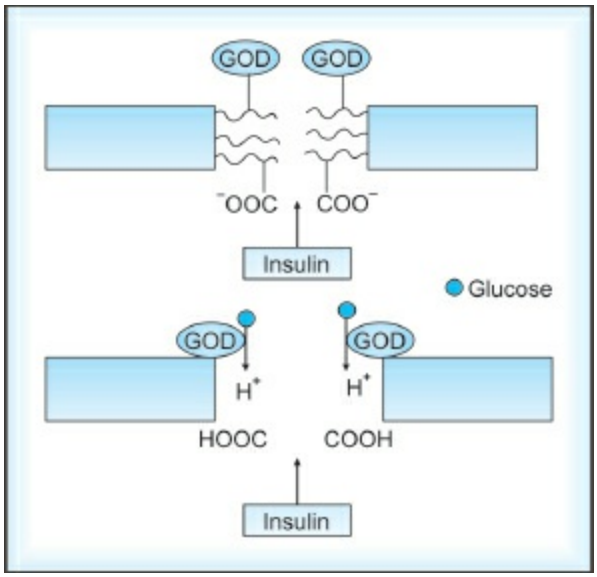



**Fig. 24.7:** Glucose-responsive insulin delivery system

## *Photoresponsive Systems*

Photoresponsive gels change their physical or chemical characteristics reversively upon exposure to photoradiation. A photoresponsive polymer consists of a photoreceptor, usually a photochromic chromophore, and a functional part. The photochromic molecules capture the optical signal and subsequent isomerization of the chromophore in the photoreceptor converts it to a chemical signal. Copolymer gels of N-isopropyl acrylamide and a photosensitive molecules, bis(4-dimethyl amino)phenyl) (4-vinylphenyl) methyl leuco-cyanide, exhibited a discontinuous volume-phase transition upon ultraviolet irradiation due to osmotic pressure of cyanide ions generated by irradiation. Suzuki and Tanaka suggested that phase transition in polymer gels is induced by visible light and the mechanism proposed was direct heading to the network polymer response to the light.

## Polymer-drug Conjugate

Polymer-drug conjugates are synthesized by either linking the drug using pendent chains along the polymer backbone or by conjugation of drug to one or both end groups of the polymer chain. The latter method has been preferred for binding mono-methoxy poly (ethylene glycol) to proteins, where several polymer chains are conjugated per protein molecule in order to coat the protein surface. Two broad classes of pendent chain linkers are (1) Peptidyl linkers that are stable in the blood stream but degradable by lysosomal enzymes and thus release the drug intracellularly and (2) acid labile, pH dependent linkers designed to remain stable in plasma (pH 7.4) but to release drug intracellularly following hydrolysis at low pH of endosomes and lysosome (pH 5.5 to 6.5).

Linkers that are lysosomally-degradable peptides or pH labile moieties have been used to prepare soluble polymer-drug conjugates, which increase the circulation time compared to free drug. Slow release of them may result in sustained drug action. The drug can be conjugated directly to an existing pendent chain such as conjugation of paclitaxel to the side chain of poly(L-glutamic acid).

It has been established that cellular uptake of low-molecular weight molecules usually occurs by rapid transmembrane passage, the uptake of soluble macromolecules occurs exclusively by endocytosis. So, when drug-polymer conjugate (macromolecule) reaches the lysosome, pendent chain of the molecules breaks in the presence of hydrolytic enzymes. Free drug molecules are released from the lysosome to the cytosol for their pharmacological action (Fig. 24.8).

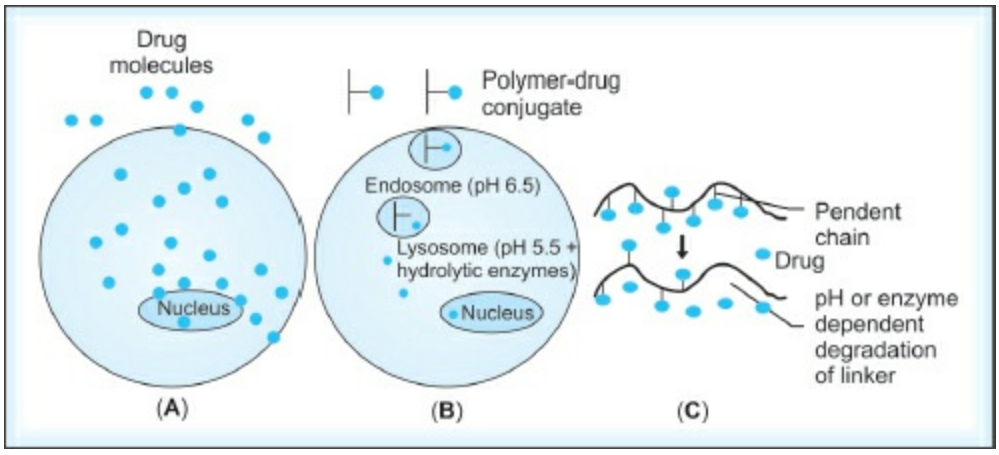


**Fig. 24.8:** Mechanism of cellular uptake: (A) Low molecular weight drugs; (B) Polymer-drug conjugates; (C) Drug release from pendent chain system due to scission of bonds connecting the drug to the polymer backbone

## NOVEL CARRIERS FOR CONTROLLED AND TARGETED DRUG DELIVERY

As the knowledge of the molecular biology and pathophysiology of diseases has expanded, more therapeutically precised and purpose-specific drugs are being developed. These newly-developed drugs have high potency (low therapeutic window) and require their localization to the particular site of their action. Most drugs are administered by conventional immediate-release dosage forms. They distribute freely throughout the body and accumulate in the non-specific organs in an undesirable manner and thus produce adverse side effects. To reduce these side effects and increase their therapeutic benefits, they should be delivered to their respective site of action, and hence a suitable carrier system becomes a mandatory requirement. Various novel carriers have been developed for the purpose. Among these colloidal carriers such as liposomes, nanoparticles and supramolecular systems, i.e. micelles have gained more attention in the field of controlled and targeted drug-delivery. Recently, new carriers such as inorganic particles, liquid crystals, aquasomes, carbon nanotubes, dendrimers, etc. are also investigated for the specialized purpose. In the following section, these carriers including other most popular drug carriers for the same purpose are discussed in brief.

## Colloidal Carriers

### *Liposomes*

Liposomes were discovered in the early 1960s by Bingham and co-workers and subsequently became the most extensively-explored drug-delivery system. Initially, though they were used to study in vitro simulated-biomembrane behaviour, subsequently, they emerged as strong therapeutic tools most notably in drug delivery and drug targeting. Structurally, liposomes are phospholipid-based colloidal vesicular structures in which hydrophilic core is entirely enclosed by membranous lipid bilayers. They may be classified on the basis of the method of preparation, structural parameters or special functions assigned to them (Table 24.3).

**Table 24.3:** Liposome classification based on pharmaceutical and therapeutic aspects

| Type | Specifications |
|---|---|
| **Based on structural parameters** | |
| MLV | Multilamellar large vesicles, >0.5 µm |
| OLV | Oligolamellar vesicles, 0.1–1 µm |
| UV | Unilamellar vesicles, (all size range) |
| SUV | Small unilamellar vesicles, 20–100 nm |
| MUV | Medium-sized unilamellar vesicles |
| LUV | Large unilamellar vesicles, > 100 nm |
| GUV | Giant unilamellar vesicles, > 1 µm |
| MV | Multivesicular vesicles, >1 µm |
| **Based on method of liposome preparation** | |
| REV | Single or oligolamellar vesicles made by reverse-phase evaporation method |

| | |
|---|---|
| MLV-REV | Multilamellar vesicles made by reverse-phase evaporation method |
| SPLV | Stable plurilamellar vesicles |
| FATMLV | Frozen and thawed MLV |
| VET | Vesicles prepared by extrusion technique |
| DRM | Dehydration-rehydration method |
| **Based upon composition andapplications** | |
| Conventional liposomes (CL) | Neutral or negatively charged phospholipids and Cholesterol |
| Fusogenic liposomes | Reconstituted Sendai virus envelopes (RSVE) |
| pH-sensitive liposomes | Phospholipids such as PE or DOPE with either CHEMS or OA |
| Cationic liposomes | Cationic lipids with DOPE |
| Long-circulatory (stealth) liposomes (LCL) | Neutral high Tc°, Cholesterol and 5–10% of PEG-DSPE or GM1 |
| Immuno-liposomes | CL or LCL with attached monoclonal antibodies or recognition sequences |

Liposomes are manufactured using various procedures in which the water-soluble (hydrophilic) materials are entrapped by using aqueous solution of these materials as hydrating fluid or by the addition of drug/drug solution at some stage during the manufacturing. The lipid-soluble (lipophilic) materials are solubilized in the organic solution of the constitutive lipid(s) and then evaporated to a dry drug containing lipid film followed by its hydration (Fig. 24.9).

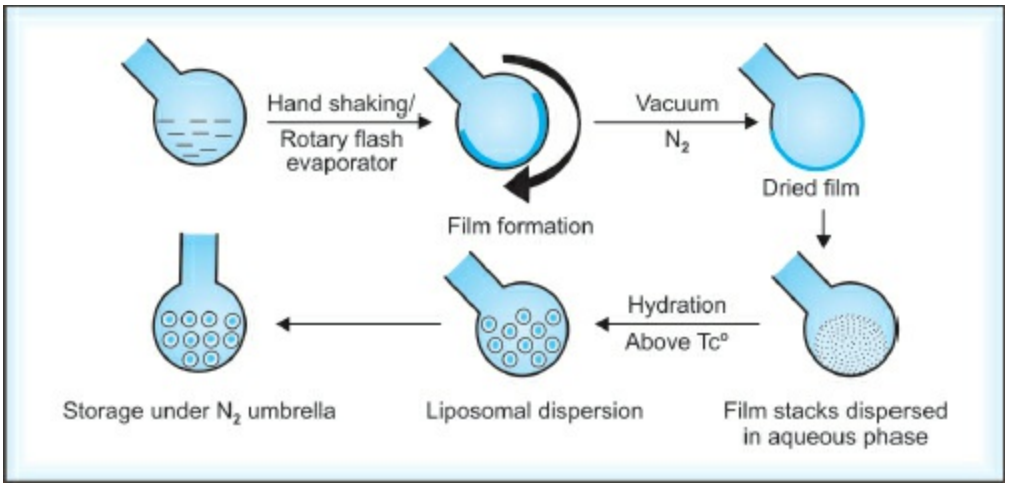



**Fig. 24.9:** Schematic representation depicting various steps involved in the preparation of liposomes by film-hydration method

Liposomes, due to their biphasic environments, can act as carriers for both lipophilic and hydrophilic drugs. Highly hydrophilic drugs (log P<−0.3) are located exclusively in the aqueous domains, whereas highly-lipophilic drugs (log P>5) are entrapped within the lipid bilayers of the liposomes. It is interesting to note that the loss of drug (drug leakage) on long-term storage is maximal with the former, and minimal or none with the latter. Drugs with intermediary partition coefficients, i.e. 1.7 (log P<4) impose problem for loading as they equilibrate between the lipid and aqueous domains and are prone to appreciable degree of leakage on storage. However, drugs with poor biphasic insolubility, mostly anticancer drugs like 6-mercap-topurine, azathioprene and allopurinol, are experienced to be problematic due to their immiscibility with both aqueous and lipidic domains. In liposomal formulations, the drug: phospholipid ratio is kept maximum that favours homing/loading of the drug. However, this is always not feasible as the extent of incorporation of a drug in the lipid vesicle strongly depends upon its solubility in the aqueous milieu and/or the lipid bilayer. Based on whether the drug is hydrophilic, lipophilic, amphiphilic or biphasic, the degree of entrapment may range from nil to 100%.

Upto now, many studies have been performed to explore their use in targeted and controlled delivery of various categories of drugs including antimicrobial, antifungal, antiviral and anticancer agents. AmBisome® (liposomal formulation of amphotericin B) is the hallmark in the field of liposomal research as the product is now available in the market for clinical

uses. Amphotericin B-based several other products are available also in UK and USA. Liposomes have also been successfully used for several other practices in drug delivery such as solubilization of water-insoluble drugs, protection of sensitive drug molecules, alteration of pharmacokinetics, and biodistribution, and enhancing intracellular uptake. To increase the targeting potential of liposomes and decrease the RES uptake, 'stealth liposomes®' (PEGylated) are the recent innovation in the field of drug delivery. Doxil™ is the marketed product of doxorubicin based on liposomes. It is being used for the targeted delivery of anticancer agent, doxorubicin, in the treatment of AIDS related Kaposi's sarcoma.

The delivery aspect of liposomes could also be exploited for drugs that must penetrate the plasma membrane in order to be therapeutically beneficial. The use of liposomal carrier system assists to surpass the membrane barriers (as they structurally mimic natural membrane) and promotes the nonspecific entry of drugs into the cellular interiors. Another field that utilizes the use of liposomal delivery system is the formulation of better tolerated pre-clinical and clinical formulations, where these carriers serve as a formulation adjuvant/excipient(s). Hydrophobic drugs such as cyclosporin A and paclitaxel (Taxol) are usually formulated in surfactants and organic co-solvents for systemic administration in humans; however, their liposomal encapsulation provides maximum therapeutic benefits. Liposomes are relatively non-toxic, non-immunogenic, biocompatible, and biodegradable; and can deliver the drug systemically with an increased therapeutic index and minimized toxicity index.

## *Niosomes*

Non-ionic surfactant vesicles (niosomes or NSVs) are now widely studied as an alternative to liposomes. Non-ionic surfactant vesicle results from the self-assembly of hydrated surfactant monomers. Non-ionic surfactants of a wide variety of structural types have been found to be useful alternatives to phospholipids. Though the terminology suggests that distinctions exist between niosomes and liposomes of which the former is having chemical differences in the monomer units, niosomes possess physical properties, which are similar to liposomes, which are formed from phospholipids. As the name indicated, generally, non-ionic surfactant vesicles are prepared by the incorporation of components containing non-ionic surfactants. However, they

may also be prepared with various ionic amphiphiles such as dicetylphosphate, stearylamine, etc. in order to achieve a stable vesicular suspension. It is important to identify and know the basic structural units of NSVs. While an amphiphilic nature is an inexplicable pre-requisite for molecules to form vesicular assemblies, variations abound in the nature of facilitating hydrophilic head groups. The vesicle-forming non-ionic compounds are mainly alkyl ether lipids. These can be broadly divided into two classes based on the nature of their hydrophilic head group, i.e. alkyl ethers in which the hydrophilic head group consists of repeat glycerol subunits, related isomers or larger sugar molecules, and those in which the hydrophilic head group consists of repeat ethylene oxide subunits. In addition, alkyl esters, amides and fatty acids, and amino acid compounds also form vesicles.

The ultimate identity of any niosomal system and hence its properties are determined by the factors listed in Fig. 24.10. It is thus obvious that all these variables must be carefully controlled in the design of a niosomal drug-delivery system.

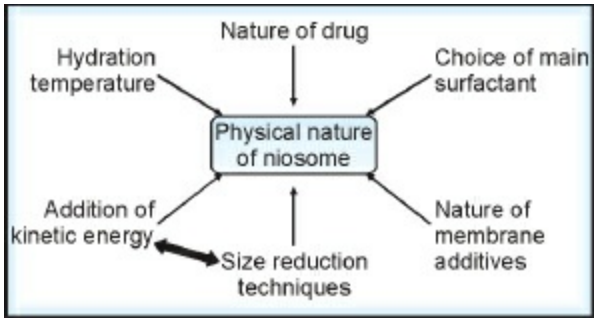



**Fig. 24.10:** Factors influencing niosome physical stability

Although pharmaceutical niosome formulations have yet to be commercially exploited, a number of studies have demonstrated the potential of niosomes in drug delivery. Niosomes have been proven to be useful in the delivery of anti-infective agents, anticancer agents, anti-inflammatory agents, and, fairly recently, as vaccine adjuvants. These systems have been proven to target certain areas of the mammalian anatomy and may be exploited as diagnostic imaging agents. Examination of the literature reveals that on IV administration of niosomes, the highest drug levels are found in the liver. However, there were exceptions. When DOX 850 nm C16G3 niosomes were administered, DOX liver levels were not significantly different from the

administration of DOX solution. The serum levels are although low (~0.5% of the administered dose 10 min after dosing) in case of solution administration, they are higher for the noisome formulation. The cause of this non-liver uptake is not apparent although smaller DOX niosomes are found to accumulate in the liver following IV administration.

## *Ethosomes*

Recently, Dayan and Touitou, described a novel vesicular carrier, the ethosomes. Ethosomes are lipid-vesicles with high content of ethanol. They can penetrate the skin and enhance compound delivery both to deep skin strata and systemically, because ethanol fluidizes both lipid bilayers of the stratum corneum and intercellular lipid. Ethosomal systems are easy to prepare, non-irritant and composed mainly of phospholipids and ethanol, compounds commonly found in pharmaceutical preparations. In contrast to liposomes, ethosomes have been shown to exhibit high encapsulation efficiency for a wide range of molecules, including lipophilic drugs, and are effective at delivering molecules to and through the skin.

In comparison to liposomes prepared in the absence of ethanol, the phospholipids in ethosomes are loosely-packed and the membrane tends to be permeable for cations. Permeation enhancement is greater in case of ethosomes than ethanol alone and this indicates that there may be some kind of synergistic mechanism between ethanol, vesicles and skin lipids.

The exact mechanism of drug delivery by ethosomes remains a matter of speculation, most likely, a combination of processes which contribute to the enhanced permeation effect. When ethosomal carriers are applied to the skin, a number of concomitant processes may takes place, involving the stratum corneum and pilosebaceous pathways. Ethanol in the ethosome seems to disrupt the organization of the stratum corneum lipid-bilayer and hence enhances its lipid fluidity. The flexible ethosome vesicle can then penetrate the disturbed stratum corneum. The release of entrapped molecule in the deeper layers of the skin, and subsequent transdermal absorption may be the result of fusion of ethosome with skin lipids and drug release at various points along the penetration pathways. Ethosomal penetration may also involve pilosebaceous pathways as they may be trapped in follicles.

Ethosomal system is much more efficient at delivering a fluorescent probe to the skin in terms of quantity and depth compared with liposomes or

hydroalcoholic solution. The ability of ethosomal systems to deliver molecules to and through the skin was investigated using Franz and side-by-side diffusion cells. It was found that the fluorescent probe D-289 penetrates the skin to a much greater depth from ethosomes than from classic liposomes, the latter, as expected, remained in the upper layers of the skin.

## *Transfersomes*

Transfersomes are complex, most often, vesicular aggregates optimized to attain extremely flexible and self-regulating membrane which makes the vesicle ultra deformable. Transfersome vesicles, therefore, can cross microporous barrier efficiently, even when available passages are relatively smaller than the average aggregate size. A transfersome crossing the skin thus mimics the behaviour of a parasite during its invasion of the host body. Transfersomes consist of natural amphipathic compound suspended in a water-based solution, sometimes containing biocompatible surfactant. Similar to liposomes, transfersomes have a lipid bilayer that surrounds an aqueous core, however in contrast to liposome, trans-fersomes contain at least one component that softens the membrane and makes it more flexible allowing an easy and rapid change of shape.

The efficacy of transpore movement of transfersomes is fairly high. A suspension of transfersomes with an average diameter of 500 nm can be transported through the pore 5 times smaller in diameter nearly as rapidly and efficiently as pure water. Such a high penetration capability is seen, when the stress suffered by transfersomes (e.g. the flowdriving pressure) is sufficiently high. This is a key characteristic of transfersomes, with roots in the self-optimizing capability of transfersomes body or membrane. The passage of transfersomes through pores that are “too small” is nearly perfect even when their size exceeds the pore diameter by a factor of approximately 4 or even 10 (Fig. 24.11).

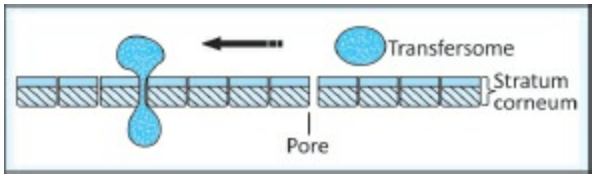



**Fig. 24.11:** Penetration of transfersome through the pore in stratum corneum

One naturally-occurring transdermal gradient is osmotic gradient. Such a

gradient is created by the difference in the total water concentration between the skin surface and the skin interior. When a lipid suspension is placed on the skin surface and partly dehydrated by water evaporation, the lipid vesicles respond to the gradient and try to escape complete drying by moving along the gradient. They can only achieve this provided they are sufficiently deformable to pass through the narrow pores in the skin. Less deformable vesicles like standard liposomes, are confined to the skin surface where they dehydrate completely and fuse.

Lipid hydrophilicity leads to xenophobia, the tendency to avoid dry surrounding and causes carriers sitting near or at the skin surface to resist dehydration in order to remain maximally swollen. Transfersomes near the skin surface, try to follow the local hydration gradient, and thereby get into deeper and better-hydrated skin strata. This causes transfersome-carriers to retract from the relatively dry skin surface and to get into more humid region in the deeper skin layers.

An important difference between trans-fersomes and liposomes is the much higher hydrophilicity of the former. This forces transfersome membrane to swell more than conventional lipid vesicle bilayer. Higher membrane hydrophilicity and flexibility both help transfersome to avoid aggregation, and thus fusion, which is observed with liposome exposed to an osmotic stress. Transfersome satisfy basic criteria for efficient skin permeation, and thus can be used as a potential carrier system for protein, interleukin, analgesic and anaesthetic agents, non-steroidal antiinflammatory drugs, corticosteroids, etc.

## *Virosomes*

Virosomes are one of the most widely investigated biological carriers. They are promising nanobiocarriers for cytosolic access vis-a-vis intracellular drug targeting. Virosomes hold potential applications in the field of vaccine development, gene and drug delivery. As a novel version of liposomes, virosomes were developed to obtain improved efficiency of interaction with the cellular target for the introduction of the molecules directly into cells. By and large, ‘Virosomes’ are liposomes spiked with viral proteins extracted from the envelope of viral virion, which imparts the liposomes viral functions, such as cell-surface adherence and fusogenicity to cellular and organelle membranes. However, they are free from virulence and infectivity.

Virosomes can be considered as hybrids of viral and liposomal carrier systems, combining the characteristics of cellular interactions of viral vectors and biosafety advantages of liposomes. Ultrastructurally, virosomes are spherical-unilamellar vesicles with a mean diameter of approximately 150 nm, possessing short-surface projections of 10–15 nm. Virosomes have attracted attention as delivery vesicles especially for cytosolic drug delivery as they are capable of transporting therapeutic molecules into the cytoplasm, owing to an as efficient encapsulation efficiency with fusogenicity which provide for a virtual sink to the cellular cytoplasm. Possibly they also avert the effect of p-glycoprotein efflux pumps by intracellular delivery drug, in sustained and controlled manner. In one of the studies, virosomes have been exploited as a drug-delivery system for cytotoxic drug, doxorubicin and found to be capable of binding and penetrating deep into tumour cell mass, ensuring effective delivery of cytotoxic drugs. Further, virosomes have been conjugated with Fab' fragments of an anti-rat Neu (anti-rNeu) monoclonal antibody to achieve efficient cell-specific targeting to the rNeu-over-expressing breast tumours.

Apart from the attainment of therapeuti-cally-effective concentration in the cytosol, virosomes can be designed to deliver the therapeutic agent to a particular cell population. This is because the mechanisms of virus entry into the cell is largely mediated through cell-surface receptors thus leading to the localization in specific cells. A wide variety of acute viral infections, such as vesicular stomatitis virus, herpes simplex virus, HSV1, encephalitis, HSV2, cytomegalovirus (hepatitis, pneumonia) and rift valley fever virus (fever) directly infect macrophages through various receptors. Epstein Barr virus (EBV) infects mainly B-lymphocytes through CR2 receptor, hepatitis B virus attacks hepatocytes through a viral-surface antigen (HBsAg) on CEA-like receptor and HIV targets T-lymphocytes and macrophages through CD4. This receptor-dependent interaction makes specialized cells or tissues a depot for these viral infections. Therefore, site-specific delivery of therapeutic agent with simultaneous intracellular therapeutic concentration is achieved by using properly designed virosome using specific membrane proteins of viral origin or whole viral envelopes.

## *Microparticles*

The "microcapsules" are defined as a spherical particles with size varying

from 50 nm to 2 mm, containing a core substance. Microspheres are, in real sense, spherical empty particles. However, the terms microcapsules and microspheres are often used interchangeably. In addition, some related terms are used as well. For example, “microbeads” and “beads” are used alternatively. Spheres and spherical particles are also used for particles of large size and rigid morphology. The dried microspheres form a free-flowing powders. They consist of proteins or synthetic polymers, which biodegrade and ideally have a size range less than 200 μm. The solid biodegradable microspheres bearing a drug dispersed or dissolved throughout particle matrix have a potential in controlled-release of drugs. These carriers received much attention not only as prolonged-release formulations but also for their carrier potential in drug targeting particularly anti-cancer drugs to the tumour.

Pre-requisites for ideal microparticulate carriers are as follows:

- Longer duration of action
- Control of drug release
- Increase of therapeutic efficiency
- Protection of drug
- Biocompatibility
- Sterllizability
- Relative stability
- Water-solubility or dispersability
- Bioresorbability
- Targetability
- Polyvalency

Microspheres can be prepared by using any of the appropriately selected methods including in situ polymerization, solvent evaporation, coacervation phase separation, spray drying and spray congealing, etc., but the choice of the technique depends on the nature of the polymer used, the drug, the intended use and the duration of therapy. The choice of method is based on the following determinants:

1. The particle size requirement.
2. The drug or the protein should not be adversely affected by the process.

3. Reproducibility of the release profile and the method.
4. No stability problem.
5. There should be no toxic product associated with the final product.

A number of different substances both biodegradable as well as non-biodegradable have been investigated for the preparation of microspheres. These materials include the polymers of natural and synthetic origin and also modified natural substances. Synthetic polymers employed as carrier materials are methyl methacrylate, acrolein, lactide, glycolide and their copolymers, ethylene vinyl acetate copolymer, polyanhydrides, etc. The natural polymers used for the purpose include albumin, gelatin, starch, collagen and carrageenan, etc.

## *Nanoparticles/Nanospheres*

Nanoparticles are nanometric colloidal particulates consisting of synthetic or semisynthetic polymers. The continual quest and maneuvering toward physical stability improvement of liposomes resulted into the development of solid core nanoparticles in nineteen eighties as an alternative drug carrier. The term particulate is suggestively general and doesn't account for morphological and structural organization of the system. Thus, they could be nanospheres, nanocapsules, nanocrystals or nanoparticulates. Nanospheres may be defined as solid core spherical particulates, which are nanometric in size. They contain drug embedded into the matrix or adsorbed onto the surface; nanocapsules however are vesicular systems in which drug is essentially encapsulated within the central void surrounded by an embryonic continuous polymeric sheath.

## *Polymeric Nanoparticles*

Over the last decades polymeric biodegradable nanoparticles have been extensively investigated as nanocarriers for the controlled- and targeted-delivery of therapeutics. Various synthetic polymers approved by US-FDA including poly-lactic acid (PLA), poly (lactide-co-glycolide) (PLGA), and polycyanoacrylate (PCA) have been studied as carrier materials. These polymers are used for the delivery of therapeutics through parenteral and implantable routes of delivery. When nanoparticles are administered intravenously, they may be easily recognized and opsonized by immune cellular component(s). To reduce opsonization and clearance, PEGylated

nanoparticles have been investigated in which PEG is used to modify the surface of nanoparticles. In another approach, different polymers may be combined to form co-polymers harnessing the property of both PEG crown and PLGA core. These polymers show anti-opsonizing character of PEG. Paclitaxel loaded in PLGA-PEG nanoparticles exhibited similar apoptotic cell death, as Taxol®, in HeLa cancer cell lines. Therefore, it can be concluded that polymeric nanocarrier-based formulations would be better and effective treatment of future. Another nanoparticles based FDA-approved formulation is Abraxane®. This is based on albumin-nanoparticles of paclitaxel which evades the hypersensitivity reaction associated with Cremophor EL, a solvent used in conventional paclitaxel formulation. Many other polymers of natural origin are also being investigated for drug delivery in which polysaccharides are most extensively studied. The investigated polysaccharides include chitosan, cyclodextrins and dextran.

## *Solid Lipid Nanoparticles*

To solve polymeric nanoparticles-associated problems, like cytotoxicity and industrial scale-up problem, lipid-based nanoparticles termed as solid lipid nanoparticles (SLN), have been developed and documented in the beginning of the 1990s. They are submicron size (50–1000 nm) colloidal particulate carriers based on physiological lipids (glycerides of fatty acids from animals). They have advantages over polymeric nanoparticles because no organic solvent is used in their preparation. The main features of SLNs include excellent physical stability, protection of the incorporated labile drugs from biodegradation, controlled (fast and sustained) and site-specific content delivery. SLNs can be used as carrier system for both hydrophilic and hydrophobic drugs. A number of studies have been performed and almost an equal number still continue to exploit SLNs as a drug carrier. SLNs provide platform for development of dosage forms for any predefined routes of delivery like parenteral, oral, topical, rectal and pulmonary, etc. The potential advantages of SLNs for parenteral route over other drug carriers especially in regard to lipophilic drugs are the use of natural lipids, which are non-toxic. Lipid-based formulations are effective tools to keep a lipophilic drug in solution and to make it available for oral absorption upon lipid digestion or extraction of drug into the GI fluids. SLNs protect labile drug in the harsh GI environment and provide sustained release of active ingredients. SLNs could successfully transverse across the BBB, and therefore can improve the

therapeutic concentration of drug(s) in brain. SLNs have also been demonstrated for topical delivery of many drugs including oxybenzone.

## *Inorganic Nanoparticles*

Recently, ceramic-based nanoparticles have been proposed as carrier system for drug and DNA delivery. These are typically inorganic systems based on silica, alumina and titania. Ceramic nanoparticles as their role in drug delivery have several distinct advantages over polymeric particles. They are biocompatible, inert, stable and porous particles. The porosity of these particles is unaffected by pH and they are not susceptible to microbial attack. Roy et al. used these particles in cancer therapy. In this study, ceramic particles have been loaded with photosensitizing drugs and used in photo-dynamic cancer treatment. The ceramic being dense in nature, concentrated at the surface of cells growing in culture and thus allowed an efficient uptake of loaded DNA through endosomal-lysosomal route. This property of ceramic particles has been exploited to increase the transfection efficiency of transfecting agents. DNA-dendrimer conjugate has been appended onto the surface of silica nanoparticles in order to increase the transfection ability of dendrimers. Recently, Aquasomes, carbohydrate-stabilized ceramic nanoparticles have shown promising carrier potentials for the delivery of a broad range of conformational-sensitive bio molecules. These are biocompatible and biodegradable nanocarriers, which exhibit a broad range of surface properties with a verisimilitude between them and live virus. These are utilized as an antigen/protein delivery system due to the presence of hydrophilic poly-hydroxyl oligomeric film of carbohydrate like cellobiose, maltose, sucrose, and trehalose on the outer surface of hydroxyapatite core. They preserve the protein from conformational changes.

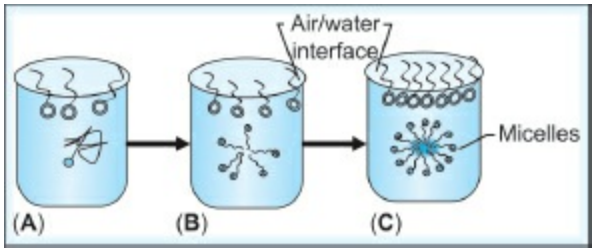
## *Supramolecular Systems*

A large number of molecules organize spontaneously, exhibiting well organized behaviour in the microscopic levels. The term supramolecular system was used first time by KL Wolf in 1937. In context of drug-delivery system, supramolecular biovector systems (SMBVs) can be defined as aggregation and assembling of molecules in such a way that they can acquire thermodynamically-stable status particularly when they suffer environmental changes. These represent a group of classically-customized drug-delivery system which is colloidal in nature and able to mimic and negotiate the

behaviour of endogeneous carriers, i.e. apoproteins and chylomicrons. They can encapsulate both hydrophilic and lipophilic drugs because they can be developed with hydrophilic/hydrophobic core-crown characteristics. Furthermore, the assemblage is principally an interfacial and surface-related phenomenon and it rarely requires sophisticated chemical methodologies. Thus, assembly of moles is as per kinetic requirement of the interface(s) in accordance with the nature of assembling molecules. The supramolecular systems are typically host-guest type and self-assembling type. Host-guest supramolecular systems are often combined at 1:1 molar ratio of host and guest. It is a spatial accommodation of guest within the host component. Further, SMBVs may be classified into three categories on the basis of their association number (1) low aggregation number (micelles, reverse micelles and polymeric micelles), (2) intermediate aggregation number (liquid crystal system) and (3) high aggregation number (lipoproteins).

## *Micelles/Reverse Micelles*

Micelles are essentially supramolecular core-shell structures resulting on account of self-assembling of amphiphiles when they are confronted with an aqueous environment at the concentration of amphiphiles above the critical micelle concentration (CMC) (Fig. 24.12). The main driving force for micelle formation in aqueous solution is the effective interaction between the hydrophobic parts of the surfactant molecules, whereas interaction opposing micellization may include electrostatic repulsive interaction between charged head groups of ionic surfactants, repulsive osmotic interactions between chain-like polar head groups, such as oligoethylene oxide chains, or steric interaction between bulky head groups. In aqueous solution, surfactants aggregate in different forms, such as spherical micelles, worm-like micelles, bilayer fragments, vesicles, or inverted structures. The way surfactant molecules aggregate is mainly determined by the attraction between the hydrophilic tails and electrostatic repulsions of the hydrophilic head groups, which are present in the surfactant molecules.



**Figs 24.12A to C:** Formation of micelles: (A) Amphiphilic molecules at low concentration; (B) Amphiphiles at higher concentration—critical miceller concentration (CMC); (C) Aggregated amphiphiles at above CMC

Recently, micelles have attracted the attention of researchers as drug delivery systems, because of their nanometric size, ability to solubilize hydrophobic drugs and site-specific delivery by passive and active targeting. Amphiphilic block copolymers with biodegradable and/or biocompatible chains have been used to incorporate hydrophobic drugs and target them to their site of action upon parenteral administration. Although much research has been carried out on micelles for parenteral delivery, their potential for oral drug administration remains largely untapped. Block copolymers that have both hydrophilic and hydrophobic polymer chains in a molecule form micelles, in water, with hydrophilic chains exposed to the bulk while hydrophobic chains agglomerate to avoid exposure to aqueous bulk. The micelles are excellent drug carriers, because they can hold drugs firmly in their inner cores (apolar) and control the drug release by changing the molecular structure of the inner cores. The micelles also have an enhanced permeability and retention effect, useful for tumour targeting. SP1049C, a miceller formulation containing doxorubicin in a mixture of L61 (PEO80-PPO27-PEO80), and F127 Pluronic (PEO101-PPO56-PEO101) is now approved by USFDA for the treatment of metastatic adenocarcinoma of the upper GI tract.

Micelles are also used as non-viral carrier system for gene delivery. Polyethylene glycol-poly-aspartic acid as anion-block copolymer mixed with polyethylene glycol-poly-L-lysine (PEG-PLL) as cation-block copolymer generates PIC with a diameter of dozens of nano-metres in a sharp diameter distribution. Micelles, 80 nm in diameter were formed with plasmid DNA and PEG-PLL. Higher expression of luciferase in cultured cells was reported in the case of micelles as compared to conventional solution of the drug.

## *Lipoproteins*

Plasma lipoproteins are spherical macromolecular complexes composed of specific protein and lipid components. They act as biological carriers and transporters for the water-insoluble lipids in blood circulation. Structurally, they contain an outer shell that is made up of phospholipids, apolipoproteins, and unesterified cholesterol and an interior core compartment,

accommodating water-insoluble lipids (triacylglycerols and choles-teryl esters) (Fig. 24.13). This arrangement provides considerable stability to the overall structure of lipoproteins and makes them particularly suitable for the transport of hydrophobic drugs.

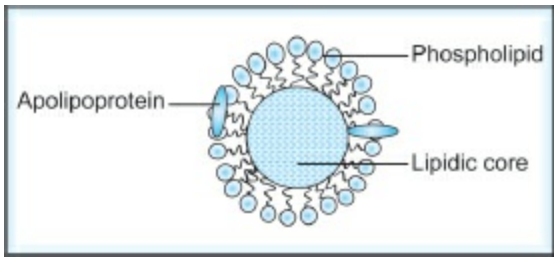



**Fig. 24.13:** Schematic representation of the structure of lipoprotein

On the basis of their densities, lipoproteins can be classified into five major classes including chylomicrons, very low-density lipoproteins (VLDL), intermediate-density lipoproteins (IDL), low-density lipoproteins (LDL), and high-density lipoproteins (HDL). Because of their superior stability and smaller size, the classes of low-density lipoproteins (LDL) and high-density lipoproteins (HDL) qualify as drug carrying nanoparticles.

There are some desirable features of lipoproteins which lend them to be suitable carriers:

- Being natural biocompatible components these provide enhanced safety and efficacy.
- They have relatively long half-life in circulation.
- Because of nanometric size range they diffuse from vascular to extravascular compartment.
- Lipidic core of lipoproteins provides a suitable compartment for hydrophobic drugs.
- These can be used for selective delivery of therapeutic molecules because apolipoproteins provide a signal for the receptor-mediated uptake of the whole particle by endocytosis.
- These can be used for cell-selective cancer therapy because of the over-expression of lipoprotein receptor by cancer cells. Thus, it is possible to attack malignant cells selectively without substantially impacting normal cells.

Among various lipoproteins, LDL has been extensively studied as a drug carrier. It also has a longer serum half-life (2–4 days) as compared to other lipoproteins. In essence, lipoprotein-based nanoparticles may be considered as being potential carriers that can execute superior drug delivery especially of anti-cancer drugs as because of their natural biodegradable components they are protected against rapid removal from the circulation. Therefore, they target the corresponding receptors selectively. These receptors are over expressed in cancer cells as compared to normal cells.

## *Liquid Crystals*

Liquid crystals combine the properties of both liquid (the ability to flow) and solid states (an ordered, crystalline structure). They can be prepared in different geometries, with alternative polar and non-polar layers (i.e. a lamellar phase) where aqueous drug solutions can be included. A pre-requirement for the formation of liquid-crystalline phases is an anisometric molecular shape that is generally associated with a marked anisotropy of the polarizability. Molecules that can form mesophases are called mesogens. Mesophase is formed either by increasing the temperature or by adding a solvent to the crystalline state of mesogens. On the basis of the procedure adopted for the formation of mesophases, differentiation can be made between thermotropic and lyotropic liquid crystals, respectively. Lyotropic liquid crystals are formed either by the mesogens that are not the molecules themselves but their hydrates or solvates, or associates of hydrated or solvated molecules anticipate in crystalline orientation.

Cubosomes are bicontinuous cubic-phase liquid crystals having many properties that make them appealing and interesting as a universal vehicle for drug delivery. Bicontinuous cubic phases consist of two separate, continuous but non-intersecting hydrophilic regions divided by a lipid bilayer that is distorted into a periodic minimal surface with zero-average curvature. These structures maintain the efficacy as well as the stability of active molecules like vitamins and proteins. Cubosomes possess the potential for controlled-delivery of drugs because the diffusion of molecules is governed by the tortuous diffusion through the “regular” channel structure (5–10 nm) of the cubic phase. They have the ability to solubilize hydrophobic, hydrophilic, and amphiphilic molecules. The most widely used lipid in the preparation of bicontinuous cubic phases is the monoglyceride-monoolein. Cubic phase is

strongly bioadhesive and is thought to be a skin penetration enhancer with excellent compatibility with topical and mucosal deposition and delivery of active ingredients. Commercial applications of cubosomes have been developed for periodontal disease that is based on triglyceride-monoolein mixtures combined with the drug metronidazole.

## Cellular Carrier Systems

### *Resealed Erythrocytes*

Red blood cells (RBCs), by their sheer numbers, 5 million/microliter in the circulation totaling 30 trillion in humans, and their long lifetimes of 120 days, are uniquely positioned to modulate the properties of blood-borne and vascular components. The breadth and diversity of vascular traffic afford a multitude of opportunities, with a vast clinical potential, for intervening in pathologies involving cellular or humoral components. The idea of using red cells as storage containers (erythrocyte encapsulation), whereby biologically-active entities can be encapsulated within the cell and later on released, continues to be pursued as a means of drug delivery and targeting. The membrane of erythrocytes can be transiently lysed or broken to yield higher permeability by using osmotic variation, or by applying electric current for nanosecond(s) in the form of pulse. During this hyperpermeation phase of membrane, the exogenous substances diffuse in and get encapsulated within the RBCs. The RBCs are prepared to be hypermermeable than loaded by bringing them in equilibrium with endogenous components and subsequently reannealing the RBCs by restoration of tonicity and incubating them at 37°C for resealing. The drug-carrying capability of resealed RBCs was realized, which subsequently became a central paradigm that resealed erythrocytes can be utilized as promising drug carriers. The desirable properties, which substantiate the suitability of RBCs as drug carriers are:

- Biodegradability and biocompatibility
- Can circulate throughout the circulatory system.
- Large quantities of materials can be encapsulated within small volume of cells.
- They can be utilized for organ targeting within the reticuloendothelial system (RES).
- A wide variety of bioactive agents can be encapsulated within them.
- Since they are biocompatible, provided that compatible cells are used in patients, there is no possibility of triggered immunological response.
- Since they are cells in common clinical usage in transfusion much is known about the technique for their collection and storage.

- Easy to prepare.

# Emulsions

## *Multiple Emulsions*

Multiple emulsions are complex multiphasic systems which are relatively a new development in the field of emulsion technology. In multiple emulsions, the dispersed phase contains smaller droplets that have the same composition as the external phase. Since this involves double emulsification, these systems are also referred to as double emulsion. The multiple emulsions are also called “Liquid Membrane Systems”, as the liquid film that separates the liquid phases acts as a thin semipermeable film through which solutes must diffuse in order to traverse from one phase to another. This liquid membrane may be hydrophilic or hydrophobic. The multiple emulsions can be either oil-in-water-in-oil (o/w/o) or water-in-oil-in-water (w/o/w) type emulsion systems.

Multiple emulsion systems have an immense potential as a drug carrier since their vesicular structure with the innermost phase is similar to that of liposomal vesicles and the selective permeability characteristic of liquid membranes. These properties have amplified the practical utility of multiple emulsions, especially as drug delivery systems. An o/w/o multiple emulsion may appear to be a more beneficial dosage form since the extrapartitioning step with drug initially in the internal oil phase is expected to be the rate-limiting step that may define the drug release characteristics.

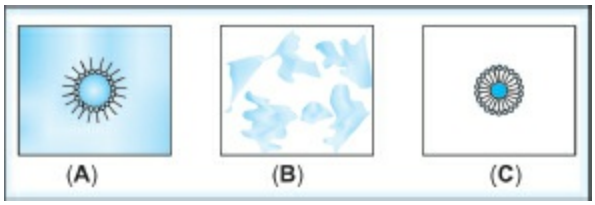
Multiple emulsions have been exploited with various diverse applications including pharmaceutics, cosmetics, food and separation techniques. Their potential pharmaceutical applications include use as RBC substitutes; treatment of drug overdosing; immobilization of enzymes; masking the taste of drugs; enhancement of GI absorption; and as carriers for sustained release, lymphatic uptake and transdermal drug delivery. In addition to these applications, w/o/w emulsions have been used as intermediates for the preparation of microspheres.

## *Microemulsion/Nanoemulsion*

Microemulsion can be defined, in general, as thermodynamically-stable, isotropically-clear dispersion of two immiscible liquids, consisting of microdomains of one or both liquids stabilized by interfacial films of surface-

active molecules. Schulman and coworkers introduced the term microemulsion to describe the clear, fluid systems obtained by titration to the point of clarity of an ordinary milky emulsion (macroemulsion) by the addition of a medium-chain alcohol such as pentanol or hexanol. Thus over the year the terms thermodynamically-stable emulsions, transparent emulsions, micellar emulsions, swollen micelles and reverse micelles have been used in the literature to describe precisely the same systems that were called microemulsion by Schulman.

Microemulsions can be classified into three types (Fig. 24.14): water-in-oil (water/oil), bicontinuous, and oil-in-water (oil/water).




**Figs 24.14A to C:** Schematic representation of different types of microemulsion system: (A) W/o microemulsion; (B) Bicontinuous microemulsion; (C) O/w microemulsion

Two types of phases are thus shown to be associated with microemulsions; droplet phases and bicontinuous phases. At higher water concentration, microemulsion consists of small oil droplets dispersed in water, i.e. o/w microemulsion, where as at lower water concentration the status is reversed and the system consists of water droplets dispersed in oil, i.e. w/o microemulsion. In each phase, the oil and water are separated by surfactant-rich film.

The oil droplets in o/w microemulsions are surrounded by the electrical double layers, which can extend into the external phase up to a considerable distance (up to 100 nm), depending on the electrolyte concentration. Thus, the hard sphere volume of the droplet is considerably greater than of the oil-core volume, which creates a strong osmotic (repulsive) force at relatively low dispersed phase concentrations.

In contrast, w/o systems are stabilized primarily by steric interactions between the absorbed films in such a way that the hard sphere volume of the droplets tends only slightly greater than that of the water pools. The droplet interaction can take place at relatively short distances of separation, where the

tail of the hydrocarbon chains can interpenetrate each other. This allows a great increase in droplet number before a strong osmotic force is felt. In many cases, it is possible to effect a gradual transition from o/w to w/o microemulsion simply by changing the volume fraction of oil and water. The intermediate region, which contains approximately equal volumes of oil and water is composed of lamellar or bicontinuous structures. In this region, both the oil and water domains extend over macroscopic distances and the surfactant forms an interface of rapidly fluctuating curvature; however the eventual net curvature comes to be near zero.

Major application areas of microemulsions are in the field of total parenteral nutrition and as carriers for various drugs, anaesthetics, and image contrast-enhancement agents.

1. Improved delivery of lipophilic drugs
2. Positively-charged microemulsions in cancer therapy
3. Photodynamic therapy of cancer
4. Neutron capture therapy of cancer
5. Perfluorochemical microemulsions in cancer therapy.

# MISCELLANEOUS

## Dendrimers

The term “Dendrimer” (Greek: dendron means tree and meros is part) defined as class of chemical molecules which are highly branched with three-dimensional macromolecular architecture that is highly controlled and defined, with all bonds emanating from a central core. Dendrimers appear to be inert, non-immunogenic and non-cytotoxic. Generally, during dendrimer formation, molecules emanate from a core and resemble a tree that more and more ramifies with each subsequent branching unit referred to as generation. Dendrimer construction is fundamentally based on two basic methods, the divergent method, where one branching unit after another is successively attached to the core molecule. Hence, the multiplication of the number of peripheral groups is dependent on the branching multiplicity. Secondly, the method that involves the opposite course is the convergent method. The skeleton is built stepwise starting from the end group towards the inside and is finally treated with a core molecule to produce the dendrimers. Either of the synthetic strategies possesses relative advantages and disadvantages, and their choice depends mainly on the kind of monomer employed and the ultimate target polymer structure.

A typical dendrimer consists of three main structural components: A multifunctional central core (C), branched units (B) and surface groups (S) (Fig. 24.15). The branched units are organized in layers called “generations”, and represent the repeating monomer units of these macromolecules. Dendrimers are characterized by an ideally-branched structure and the presence of a high number of functional groups, which can have a significant effect upon the physical properties both in the solid state as well as in solution. Furthermore, the macromolecular dimensions can be controlled, since the synthesis involves a repetitive sequence of steps. An important area where linear and dendritic polymers exhibit diverse characteristics is their viscosity behaviour. It is well known that the intrinsic viscosity of a linear polymer increases with the increase of molecular weight (MW) according to the Mark-Houwink-Sakurada relationship. However, dendrimers exhibit a linear relationship at lower generation numbers and a maximum that corresponds to the change in shape, followed by a smooth decrease in intrinsic viscosity at higher molecular weight. Another important

characteristic of dendritic molecules is their high solubility in a large number of organic solvents, potentially offering better processability character and rapid dissolution. These distinguishing features of dendritic macromolecules render these novel materials a reliable alternative to traditional polymers in a wide range of applications.

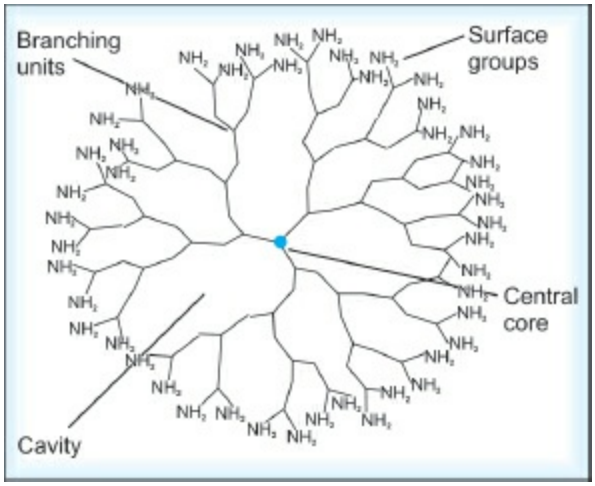



**Fig. 24.15:** Schematic structure of dendrimer

The application of dendrimers in pharmaceutical and medical chemistry is becoming one of the most attractive areas. Being well-defined structure, compact globular shape, size, monodispersity and controllable surface functionalities, dendrimers offer excellent candidature for drug delivery. Dendimers can be utilized as potential drug carriers in two ways: firstly, drug molecules can be physically entrapped inside the dendritic structure, and secondly, drug molecules can be covalently linked onto the dendrimer surface or other functionalities to produce dendrimer-drug conjugates. A variety of applications have been explored especially in gene transfection and medical imaging vis-a-vis as drug delivery systems. Dendrimers provide new platform for the transfection and manipulations of cells. They are structurally-defined and provide very high transfection efficiency in vitro with a wide variety of cell types. These polymeric moieties also furnish a soluble nanometric matrix for drug molecule(s) attachment/incorporation or other materials useful in enhancing transfection.

Polyamidoamine (PAMAM) dendrimer with an ellipsoidal or spheroidal shape is one of the most studied starburst macromolecules. The high density of functional groups ($–NH_2$, −COOH, −OH) in PAMAM dendrimers may be

expected to have potential applications in enhancing the solubility of low aqueous soluble drugs and delivery systems for bioactive materials. Also, these functional groups on the outer shell are responsible for high reactivity, which means that dendrimers can be modified or conjugated with a number of interesting guest molecules. Furthermore, PAMAM dendrimers possess empty internal cavities, which are able to encapsulate hydrophobic guest molecules in the macromolecule interior. Drugs or other molecules can either be attached to dendrimers' end groups or encapsulated in the macromolecule interior. In addition to improving drug properties such as solubility and plasma circulation time, polymeric carriers can also facilitate the passive targeting of drugs to the solid tumours. These factors, in combination, lead to the selective accumulation of macromolecules in tumour tissue—a phenomenon termed the 'Enhanced Permeation and Retention' (EPR) effect. In addition, the presence of reactive groups on the surface of dendrimers makes this carrier a suitable vector for targeted delivery of bioactives to their respective site of action with help of specific homing device attached to the surface of dendrimers.

## Carbon Nanotubes

Over the last decades, carbon nanotubes (CNTs) have become an interesting possible option for drug carrier in the field of biomedicine. CNTs are basically rolled sheet of graphite terminated by two end caps similar to a half $C_{60}$. CNT can be classified into singlewall CNT (SWCNT), and multi-wall CNT (MWCNT) according to the number of layers of the rolled graphite. The type attracting most attention is the single-wall CNT which has a diameter deserving the name of “nano” tube of 0.4 to 2 nm. The length is usually of the order of microns, but single-wall CNT with a length in the order of centimeters has recently been released. The extremities of the CNT are usually closed with lids of the graphite sheet. CNT is made up of a compound containing graphite and carbon and synthesized using transition metals, i.e. iron or nickel as the catalysts. Synthesis from graphite is possible using either an arc-discharge or laser-ablation technique. CNT can also be synthesized from a carbon compound by means of chemical vapour deposition (CVD).

Specific features which make these carriers attractive for drug delivery applications are that their hollow and porous nature which helps in the entrapment of drug molecules and subsequently drug release at the target site in a controlled manner. Biocompatibility of these carriers has been the major issue for using in vivo as delivery vectors until Singh, et al. in their study reported that functionalized CNT (f-CNT) can be rapidly cleared from the systemic circulation through renal excretion. Covalent functionalization using the versatile 1,3-dipolar cyclo-addition is the most widely acceptable approach. The surface is ameliorated functionalized CNTs can deliver the therapeutic molecules to site of action. Similarly, functionalized with polyethylene glycol (PEG) (known as PEGylated CNTs) has also been extensively used to improve the solubility of CNT-conjugates, as well as to impart biocompatibility to the carrier. Wu *et al*., reported the possible use of functionalized CNTs in targeted delivery of Amphotericin B (Amp B) to the Jurkat cells in vitro. In this study they demonstrated that the solubility-related toxicity of Amp B can be decreased on covalent attachments of Amp B to the f-CNT. f-CNT have also been studied for the delivery of macromolecules such as proteins, peptides and vaccine.

## Self-emulsifying Drug Delivery Systems (SEDDS)

One of the most popular and commercially-viable delivery approaches used to improve the solubility and bioavailability of poorly water soluble drugs is self-emulsifying drug delivery systems (SEDDS). SEDDS are defined as isotropic mixtures of drug, oil/lipid, surfactant, and/or cosurfactant. Upon mild agitation, followed by dilution in aqueous media, such as GI fluids, these systems can form fine emulsion/lipid droplets, ranging in size from approximately 100 nm (SEDDS) to less than 50 nm for SMEDDS. Fine oil droplets would pass rapidly from the stomach and promote wide distribution of the drug throughout the GI tract, thereby minimizing the irritation frequently encountered during extended contact between bulk drug substances and the gut wall.

Hydrophobic drugs can often be dissolved in SEDDS allowing them to be encapsulated as unit dosage forms for per-oral administration. When such a formulation is released into the lumen of the gut, it disperses to form a fine emulsion, so that the drug remains as solution in the gut, avoiding the dissolution step which otherwise frequently limits the rate of absorption of hydrophobic drugs from the crystalline state. Generally, this can lead to improved bioavailability, and/or a more consistent temporal profile of absorption from the gut.

Traditional preparation of SEDDS involves dissolution of drugs in oils and their blending with suitable solubilizing agents. These are usually formulated with triglyceride oils and ethoxylated non-ionic surfactants (most commonly Tween 80), usually at surfactant concentrations range of 30–60%. The precise mechanisms of emulsification remain the subject of speculation but there is an empirical interrelation between self-emulsification, liquid crystal formation, oil-water phase-inversion temperature and enhanced solubilization of water by oily formulations, and these phenomena serve as indicators of the efficiency of emulsification. Excipients in the formulation are usually selected from the Generally Recognized As Safe (GRAS) list of ingredients as published by the FDA. If compounds are not listed in GRAS, their potential toxicity is of the utmost importance.

However, SEDDS are normally prepared as liquids that produce some disadvantages, such as high production costs, low stability and portability, low drug loading and few choices of dosage forms. More importantly, the

large quantity (30–60%) of surfactants in the formulations can induce GI irritation. To solve these problems, solid-SEDDS have been investigated as an alternative approach. This system requires the solidification of liquid self-emulsifying ingredients into powders/nanoparticles to create various solid dosage forms including SE tablets, SE capsules, SE solid dispersion, SE pellets, SE sustained-release microspheres and SE nanoparticles. Solid-SEDDS ensures the advantages of both SEDDS (i.e. enhanced solubility and bioavailability), as well as solid dosage forms (e.g. low production cost, high stability and reproducibility, and better patient compliance).

## Aquasomes

Aquasomes are three-layered self-assembling compositions with ceramic carbon nanocrystalline particulate core coated with glassy cellobiose, or alternatively degradable calcium phosphate nanocrystalline particulate core coated with glassy pyridoxal-5 phosphate (Fig. 24.16). Subsequently, drug/enzyme is non-covalently bound to the outer coating. In general, these complex multicomponent particulate delivery systems could be viewed as assemblies of simple polymers, complex lipid mixtures or ceramic materials with a diameter ranging between 30 and 500 nm. As these are solid or glassy particles dispersed in an aqueous environment, they exhibit the physical properties of colloids and their mechanism of action is controlled through their surface chemistry. Aquasomes deliver their contents under combination of multifunctions including specific targeting, molecular shielding, and slow and sustained release processes. Owing to their large size and inherently-active surfaces, a substantial amount of agents through non-covalent processes could be loaded successfully. Moreover, due to their size and relative structure stability, they can avoid clearance by RES or degradation by other environmental challenges.

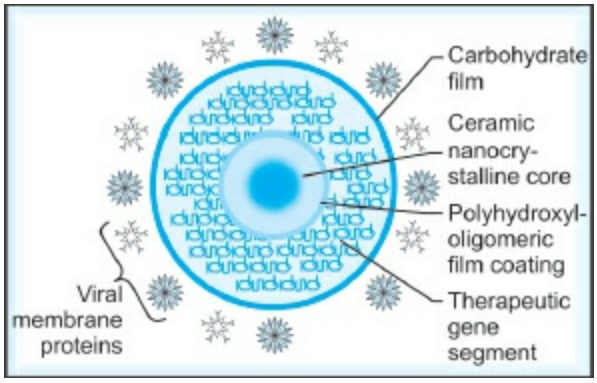



**Fig. 24.16:** Aquasomes for gene therapy

## Emulsome™

Emulsome™ represents lipid-based drug delivery systems with a wide range of therapeutic applications especially for parenteral delivery of drugs, which are poorly water-soluble. Since lipophilic drugs have limited water solubility, they may require an excessive quantity of surface-active agents or co-solvents which often leads to undesirable toxic side effects. Emulsome™ particles basically consist of microscopic lipid assembly with a polar core. The latter contains water-insoluble drugs in the solution form without requiring any surface active agent or cosolvent. These lipid-cored lipid particles are dispersed in an aqueous phase (Fig. 24.17). These systems are often prepared by melt-expression or emulsion-solvent diffusive extraction. However, in place of emulgents, an excessive quantity of phospholipids is used for their stabilization. The latter not only stabilizes the entire dispersion, but excess of phospholipids assemble to form the lipid bilayered membrane similar to liposomes. In other words, this system combines characteristics of lipid spheres and liposomes, i.e. characteristically with apolar core and liposomal crown (surface). These systems can very safely be used for parenteral administration of drugs, adjuvants for vaccines, carriers for targeted-drug delivery to liver, brain or RES-rich organs. The active targeting has also been described to be possible using this system by Pharmos Corporation.

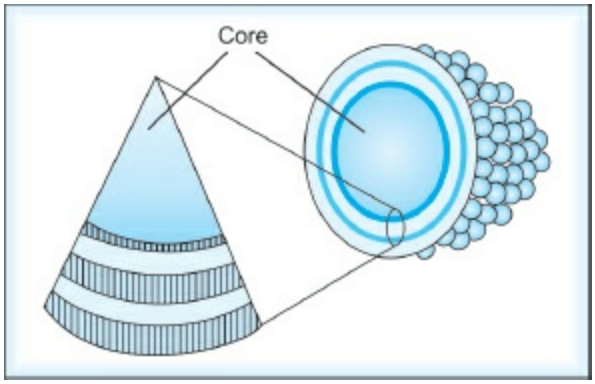



**Fig. 24.17:** An insight into Emulsome™ particles

Emulsomes™ have been successfully applied for delivery of Amp B, zidovidine and methotrexate both through parenteral and intraduodenal route, and of course provide improved clinical response over conventional systems.

## DQAsomes

Dequalinium, a dicationic mitochondriotropic molecule resembles bolaform electrolyte in structural configuration, i.e. it possesses two charge centers separated by a long alkyl chain. These molecules show self-assembling behaviour like aracheal lipids but can be distinguished by the feature that these molecules contain only one alkyl chain connecting two charge centers. Therefore, they are named as “single-chain bola amphiphile”. Similar to liposome, dequalinium and its derivative form vesicle-like aggregates upon sonication in aqueous medium and these vesicles are named as DQAsome since they are derived from deqalinium. The average size of these vesicles has been found to be in range of 70–700 nm. Weissig, et al. who have discovered, DQAsome, successfully use these vesicles as mitochondria-specific delivery system for targeting of DNA molecules towards mitochondria in a number of experiments. Whatsoever molecule is encapsulated or complexed to these vesicles, it is released at the outside of the mitochondrial membrane because they become destabilized upon contact with mitochondrial membrane and contents are left in free form. Subsequently, they are taken up following the process like mitochondrial leader sequence peptide-mediated uptake of the DNA or via the mitochondrial protein import machinery. A potent anti-cancer drug, paclitaxel, which is used in the treatment of many malignancies, has been encapsulated in DQAsome for mitochondria-selective targeting. The results show that at the clinically relevant concentrations, paclitaxel triggers apoptosis by inducing cytochrome C release in a permeability transition pore-dependent manner. These paclitaxel-loaded DQAsomes have been found able to inhibit human colon carcinoma growth in nude mice by 50% over control group. The encapsulation of paclitaxel into DQAsome produces two distinct advantages: first by with DQAsome, solubility problem may be overcome, as they are colloidal carrier system, and secondly, both the barrier in mitochondrial targeting like cell membrane barrier and mitochondrial membrane barrier can be crossed by cationic DQAsome responding to negative inside membrane potentials showing the advantage of “double targeting” by a single carrier system, which seems to be a new conceptual prospect in the case of anticancer therapy.

## Nanosuspension

Nanosuspension is defined as the dispersion of sub-micron colloidal drug particles in water and stabilized by using surfactant. The particle size-distribution of the solid particles in nanosuspensions is usually less than 1 micrometer with an average particle size range of 200–600 nm. The need for nanosuspensions as a dosage form is recognized as a means to administer therapeutic quantities of water-insoluble dosage forms. The reduction of drug particles into the sub-micron range leads to a significant increase in the dissolution rate and therefore, enhances bioavailability. Various mechanisms have been explained which account for the improved bioavailability of these systems. These include adhesion of drug nanoparticles to the mucosa, increased saturation solubility which leads to increased concentration gradient across GI tract lumen and blood, and increased dissolution rate of the drug.

Preparation of nanosuspensions has been reported to be a more cost-effective and technically-simpler alternative, particularly for poorly soluble drugs. It yields a physically more stable product than liposomes, the conventional colloidal drug carriers. The stability of the particles obtained in the nanosuspension is attributed to their uniform particle size which is created by various manufacturing processes. The absence of particles with large differences in their size in nanosuspensions prevents the existence of different saturation solubilities and concentration gradients, consequently preventing the Ostwald ripening effect. Nanosuspension engineering processes currently used are preparation by precipitation, high-pressure homogenization, emulsion and milling techniques. The inherent high loading of this dosage form distinguishes it from liposomes, emulsions, cyclodextrins, and polymeric nanoparticles. These delivery systems can be administered through different routes including parenteral, ocular, pulmonary, oral, etc. Many drugs such as paclitaxel, clofazimine and itraconazole have been administered through parenteral route and resulted into an enhanced efficacy.

The safety profile of drug has been reportedly observed to be improved in many drugs when compared to their respective conventional solution forms. This occurs due to deletion of noxious excipients, change in the pharmacokinetic profile, or regional delivery, thus minimizing systemic

toxicity.

Nanosuspensions have numerous applications such as the delivery of regional anaesthetics, treatment of malignant hyperthermia, fungal treatment, cancer chemotherapy and intrathecal pulmonary/nasal delivery of drugs. Nanosuspensions can also be used for targeted delivery by altering the surface properties of dispersed phase.

## Polymersomes

Polymersomes, block copolymer-based vesicles, constitute a new class of drug carriers. They are spontaneously formed in aqueous media, as unilamellar vesicles up to tens of microns in diameter. However, the size may be in sub-micron range under applied fluid shearing which implies that vesicle size is a kinetically-adjustable parameter. Block copolymers have the same basic architecture as lipids but consist of distinct polymer chains covalently linked in a series of two or more segments (hydrophilic and hydrophobic). There is a clear relationship between the hydrophobic content of polymers and self assembly. Polymers having low hydrophobicity (less than 50%) favour the formation of micelles, however, intermediate level of hydrophobicity (50–80%) favours the formation of vesicles. Polymersomes have been prepared from a variety of block copolymers. Table 24.4 lists the most commonly studied block copolymers for the formation of polymersomes.

**Table 24.4:** Block copolymers used in preparation of polymersomes

| **Synthetic** | **Semisynthetic** |
|---|---|
| **Biodegradable** | |
| PEG-Poly(lactide) | Poly(iso-cyano-L-alaline-L-alanine)-Poly(styrene) |
| PEG-Poly(glycolide) | |
| PEG-Poly(e-caprolactone) | |
| **Non-biodegradable** | |
| PEO-Poly(styrene) | |
| PEG-Poly(styrene sulfonic acid) | |
| Poly(phenyl quinoline)-Poly(styrene) | |
| PEO-Poly(propylene oxide)-PEO | |

*Abbreviation:* PEG: Poly(ethylene glycol), PEO: Poly (ethylene oxide)

Polymersomes possesses several desirable characteristics as a drug carrier. They are as follows:

- They offer enhanced mechanical stability (due to intra-and intermolecular hydrophobic interwinding), whereas liposomes are fragile because the hydrophobic interactions in liposomes is limited to that between adjacent lipids.
- Particle size can be controlled by changing the chain length of the polymers.
- Amphiphilic block copolymers offer a greater flexibility and exhibit membranelike properties.
- Both hydrophilic and hydrophobic drugs can be incorporated.
- Relatively easy to prepare
- Systemic circulation time is relatively longer than lipid vesicles.
- Biodegradable block copolymers are biocompatible and non-toxic to the body.
- Facilitate encapsulation of aqueous solution of proteins and other drugs known to be unstable in organic solvents.

## Nanocapsules

Nanocapsules are polymeric membranes with an oily or aqueous core. They can be defined as vesicular systems in which “drug is confined to an aqueous or oily cavity surrounded by a single polymeric membrane”. They are colloidal carriers applied for drug delivery. Usually, the nanocapsules contain an outer surfactant adsorption layer. Poly-alkylcyanoacrylates and polylactides are polymers used for the outer coating. While the core of the vehicle comprises of “soyabean oil or other triglycerides having long and medium chain fatty acids”. Nanocapsule of polyisobutylcyanoacrylate (PIBCA) finds extensive applications in drug delivery.

## Organogels

Organogels are highly self-structured systems, which are isotropic, thermoreversible, semirigid systems formed by peculiar kinds of small organic molecules. They dynamically tend to self assemble into a three-dimensional network of nanoscale measurement, thereby transforming a liquid into a gel (e.g. lecithin, gelatin, or sorbitan ester based gels). Hence, the movement of the dispersing medium is restricted by interlacing three-dimensional scaffold structure of particle of solvated macromolecules of dispersed phase.

## Lipid Cochleates

The cochleates are representatives of the most versatile technology for the delivery of a wide range of drugs and fragile molecules such as proteins and peptides. They are calcium-phospholipid structures composed of naturally-occurring materials which “wrap” around the drug or nutrient core intended to be introduced into the body. They protect the entrapped molecules from the harsh environmental conditions of GIT. They are reportedly stable on lyophilization and can be reconstituted with appropriate vehicle prior to administration.

Cochleates are formed when small unilamellar vesicles made from negatively-charged lipids mainly phosphatidylserine (PS) fuse into cylindrical rolls, termed as cochleate cylinders, upon addition of $Ca^{++}$ ions. Subsequent removal of $Ca^{++}$ by EDTA or by ion exchange or precipitation, results in the formation of large unilamellar vesicles. Concentration of $Ca^{++}$ require to reproduce such effects vary for the nature of lipids used for PS. It should be slightly above half of the lipid concentration, while CL and especially PG require higher overages. Cochleates have some special features which makes them a suitable carrier for drug delivery as follows:

- Easy and safe production.
- Limited permeability to oxygen prevents the oxidation of lipids, which imparts stability to nanocochleates.
- Hydrophobic or amphiphilic molecules and tissue impermeable drugs can be delivered via cochleates.
- Lipids used in cochleates also posses nutritional value and have an effect on enhancing the brain functions in elderly people, e.g. phosphatidylserine.
- Lipid matrix-based subunit vaccines can be used to produce custom-designed vaccines that elicit the desired immune responses targeted to specific parts of the pathogen that are relevant to protection.

Nanocochleates can be used to deliver proteins, peptides and DNA for vaccine and gene therapy applications. They have showed potential to deliver Amp B orally and parentally with a good safety profile and reduced cost of treatment. The prepared cochleates of Amp B have demonstrated improved stability and efficacy even at low doses.

## APPLICATIONS OF NOVEL DRUG DELIVERY SYSTEMS

### Sustained and Controlled-drug Delivery

Controlled release of the drug or encapsulated bioactives could be achieved using NDDS. Desired release pattern will definitely improve the pharmacokinetics and hence pharmacodynamics of drug. The controlled delivery of antibiotics in the treatment of *H. pylori* via NDDS is an effective process compared to conventional one. Similarly, slow and sustained release of drug from implants avoids regular administration of drug and hence ensures patient compliance. Numerous applications of NDDS in sustained and controlled delivery of drugs are enumerated. Some of them have already been discussed in preceeding sections.

Depot formulations of short-acting peptides have been successfully developed using microparticle technology. Such peptides include leuprorelin acetate and triptoreline, both luteinizing hormone releasing hormone agonists. Leuprorelin polylactided acid co-glycolide microspheres may be used as monthly and three-monthly dosage forms in the treatment of advanced prostate cancer, endometriosis and other hormone responsive conditions. These microspheres effectively halt the progression of prostate cancer or endometriosis in patients and are currently marketed as Prostap SR.

Other peptides formulated as sustained release microparticles include the angiotensin receptor-antagonist, L-158809, for the treatment of hypertension, thyrotropin releasing hormone for central nervous system stimulation, salmon calcitonin for the treatment of hypercalcemia and postmenopausal osteoporosis and the immunosuppressant drug cyclosporin A. There are a number of products available in the market for clinical studies as listed in Table 24.5.

**Table 24.5:** List of various marketed formulations based on novel drug delivery systems

| Product name | Drug | Carrier system | Indication | Company name |
|---|---|---|---|---|
| Doxil® (CAELYX) | Doxorubicin | Liposomes | Kaposi's sarcoma | SEQUUS |
| DaunoXome® | Daunorubicin | Liposomes | Advanced Kaposi's sarcoma | NeXstar |
| AmBisome® | Amphotericin B | Liposomes | Systemic fungal infection | NeXstar |
| Amphocil® | Amphotericin B | Liposomes | Systemic fungal infection | SEQUUS |
| Lupron Depot® | Leuprolide acetate | Microsphere | Prostate cancer | Takeda-Abott |
| Nutropin Depot® | Recombinant human growth hormone | Microsphere | Growth hormone deficiency | Genentech-Alkermes |
| Dacapeptyl® | Triptorelin | Microsphere | LHRH agonist | Novartis |
| Eligard® | Leuprolide acetate | Microsphere | Prostate cancer | Sanofi-Synthelabo |
| Abraxane® | Paclitaxel | Nanoparticle | Breast cancer | Abraxis Bioscience, Inc. |

## Targeted-drug Delivery

A targeting approach may provide maximum possible clinical response from a drug with minimum side effects. One of the major applications of novel technology is in the selective and effective delivery of various anticancer agents. The microvasculature of healthy tissue varies by tissue type, but in most tissues including the heart, brain and lung, there are tight intercellular junctions with less than 10 nm intercellular space, however tumour microvasculature contains pores ranging from 100 to 1000 nm in diameter. Therefore, tumor within these tissues can be selectively targeted by creating drug-delivery systems of size greater than the intercellular gap of the healthy tissue but smaller than the pores found in tumor vasculature. Therefore, the carriers of size range from 10–1000 nm are appropriate for the targeted delivery of various anti-cancer drugs.

Drug delivery to the central nervous system still remains to be a challenge because of the unique barrier nature of BBB, which segregates the brain from the circulating blood and effectively limits drug entry to the brain. As a result, high dose is required for the therapeutic benefit, which, at the same time, increases risks of adverse side effects. Drug delivery to the brain thus requires further advances and innovations in both drug discovery and drug delivery. Only small, lipophilic drugs are believed to transverse across the BBB. Small colloidal nanocarriers such as liposomes and nanoparticles have been suggested as possible carriers for selective delivery of drugs to the brain.

Now-a-days, cardiovascular diseases constitute vulnerable target diseases group in developed countries. For better treatment of these diseases, the drug should reach at the target site in a sufficient concentration for the desirable period of time while at the same time its systemic toxicity should be minimum. To fulfill such requirements, nanotechnology-based carrier systems appear as appropriate options. Small size of nanocarriers allows rapid incorporation of drug into the thrombus interior. Encapsulation of immunogenic thrombolytics such as streptokinase in the liposomes may also decrease the immunogenicity of streptokinase. Thus, by using engineered novel carrier-based formulations, it could be possible to treat such diseases more effectively.

Surface modification of novel drug delivery systems provides platform

for targeted drug delivery to the desired site. These functionalized carrier systems have been implicated in targeted delivery of anti-cancer agents to cancerous cells, brain targeting, colon targeting, macrophage targeting, etc. (*see* chapter “Targeted Drug Delivery” for more details). Since these carrier systems could be made effective by suitable manipulation of size, hydrophobicity, ligand anchoring, chemical modification or surface modification with muco-bioadhesive material, it is said that NDDS are an advanced version of pharmaceutics capable of targeted drug delivery.

## Vaccine Delivery

New vaccine delivery technologies are clearly needed both to remedy the limitations of existing immunization, regimens and to allow for the development of new or improved vaccines. Novel vaccine delivery systems, distinct from classical adjuvants, are now being investigated as a means to modulate the immune response following vaccination. Several such systems are under clinical evaluation with promising results.

It is appreciated that novel delivery systems can play an important role in the development of new strategies for improved vaccine delivery and such systems could have the major benefit of ease of preparation. Novel carrier systems act as potent and non-toxic immunological adjuvant was first observed in 1974 by Allison and Gregoriadis using diphtheria toxoid as a model antigen. Both humoral and cell-mediated immunity was observed to be induced. This was attributed to the ability of liposomes to slowly release antigens at the site of injection, to migrate to the lymphatic as well as to their avid uptake by antigen presenting cells, usually macrophages. Recently, the first liposome-based vaccine Epaxal Berna® against hepatitis A has been licensed for use in Switzerland.

The carrier-encapsulated antigens are capable of inducing systemic and mucosal immune responses following administration. They can be freeze-dried and therefore, avoid the need for cold storage. The incorporation of antigens into biodegradable microspheres has several advantages, including, the protection of antigen from proteolysis and possible co-incorporation of immunological adjuvants that may further enhance the immune response. Recent advances in the field of nanotechnology and drug encapsulation procedures has demonstrated that nanocarriers such as liposomes, PLGA nanoparticles, etc. stabilize the antigen, against harsh environment of gastric fluid and increase the uptake of antigen, especially through M-cells for presentation to the antigen-presenting cells (APCs).

## Gene Therapy

The ability to transfect genes into cells and to cause their expression is leading to the practical emergence of human gene therapy, wherein, functionally-active genes are putatively inserted into the somatic cells of a person requiring the expression of a given protein. A novel adaptation of gene therapy is the transfection of cells with non-resident genes in order to accomplish in situ expression of a pharmacologically-beneficial protein or create a site for further therapeutic intervention. In other words, genes would act like “drugs”, generating a product with a specific pharmacological effect. In simple terms, gene therapy involves the insertion of genetic material into a patient’s cells to make them capable of producing the therapeutic protein.

Due to the sensitivity of DNA to enzymatic degradation genetic material can not be introduced in vivo unprotected. To achieve effective transfection, various vectors are employed as a means of delivering the genetic material. At present, the majority of approved clinical trials on gene therapy in human subjects have involved viral transfection using viral vector-mediated transfer. Among non-viral carriers, cationic liposomes, cationic polymers-based carriers etc. have excellent transfection activity and hence are accepted as useful tool for gene therapy.

Cationic liposomes, as the name suggests, consist mainly of a positively-charged lipid. Felgner and collaborators proposed the use of cationic liposomes (composed of the cationic lipid 2,3-bis(oleoyl)oxipropyl trimethylammonium chloride and dioleoyl-phosphatidylethanolamine (DOPE), which are commercially available as a trans-fection reagent, Lipofectin®, as efficient carriers for the intracellular delivery of DNA. DOPE can fuse with endosomal membrane. DNA, on mixing with cationic liposomes, produces a condensed DNA along with tubular structure and liposome aggregate. The mechanism by which DNA-cationic liposome complex delivers the DNA is understood to be as follows. The complex first interacts with the cell membrane, followed by endocytosis and finally disruption of endosomes. These lipid-based transfection agents have been employed successfully for genetic loading of a variety of cell lines in vitro.

Similar to cationic lipids, cationic polymers, such as polybrene and DEAE-dextran, have been studied in the past for transfecting cells in vitro. However, their use was restricted in vitro due to their low efficiency, cyto-

toxicity and non-biodegradability. Recently, nonlinear polycationic polymers have been synthesized and proposed for non-viral gene delivery.

So far, various cationic polymers like poly-L-lysine, polyallylamine, polyamidoamine dendrimers, polydimethylaminoethylmethacrylate, and PEI have been investigated in gene delivery. Polymers like PEI possess inherent endosomolytic activity and protonated at physiological pH. These kinds of polycations are believed to escape the endosomes via the proton sponge mechanisms.

Starburst$^{TM}$ polyamidoamine (PAMAM) dendrimers belong to a new type of synthetic polymer class characterized by a branched spherical shape and a high density surface charge. These high-density surface charges help bind to various forms of nucleic acid through electrostatic interactions. This ability has been investigated for the effective delivery of ‘antisense oligonucleotides’ and ‘antisense expression plasmids’.

For many gene therapy applications, it is essential to achieve high localized expression of a DNA of interest. Receptor-mediated gene transfer provided by molecular conjugate vectors has been used to achieve this goal. Commonly used conjugate vectors include the asialoglycoprotein to target gene to the hepatocytes. Such ligands can be easily anchored on NDDS and can be utilized for correct gene therapy.

## COMMERCIAL DEVELOPMENT OF NOVEL CARRIER: LIPOSOME

Several liposomebased products are either approved or under the process of approval in various parts of the world. However, scaling-up them as a conventional market product still remains to be established and implicated. Problems generally encountered in the development of pharmaceutical liposomes are:

- Poor quality of the raw materials mainly the phospholipids
- Poor characterization of the physicochemical properties of the liposomes
- “Payload” is too low
- Circulation half life is too short
- Scale-up problems
- Absence of any data on safety of these carrier systems on chronic use.

However, in recent years several pragmatic solutions are being proposed and tested to overcome the abovementioned process related problems, which include:

- High-quality products with improved purification protocols and validated analytical techniques are now available.
- Quality control assay can be performed using sophisticated instruments and batch-to-batch variability can be checked.
- Payload problems can be sorted out using either lipophilic drug/lipophilic prodrug of hydrophilic drugs or using active (remote loading) techniques.
- Shelf-life can be improved using cryoprotectant and lyoprotectant and can be successfully freeze-dried.
- Scaling-up can be improved by carefully selecting the method of preparation (high shear homogenizer), sterilization by autoclaving or membrane filtration (0.2 μm) coupled with aseptic processing and pyrogen removal using properly validated LAL test.
- By choosing candidate potent drugs with narrow therapeutic window (e.g. cytotoxic drugs and fungicides), the drug related safety problems can be alleviated.

After many encouraging pre-clinical data and safety studies, the

liposomal formulations entered in the clinical trial stages. In parallel, the manufacturing process was scaled up and all the quality controls and analysis are now well defined and standardized. In a typical large-scale method, production lipids are mixed and dissolved in officially-approved organic solvent, i.e. chloroform, methylene chloride, methanol, tertiary butanol and some other solvents. These solvents can be removed either by evaporation, vacuum drying or lyophilization (in case of t-butanol). The dry lipid film, paste or cake is then hydrated.

In the case of Doxil$^{TM}$ production, lipids are mixed and dissolved in organic solvent (FDA and safety committee approved), which can be injected into an aqueous solution of ammonium sulfate. Upon hydration the organic solvent is diafiltered away and large multilamellar vesicles are sized down, preferably by an extrusion technique. After the exchange of the external solution empty liposomes are loaded with doxorubicin which is added as a concentrated solution in the system. Loading is so sufficient that free drug removal is not necessary. Some buffers may be added and the product is then purified, sterile filtered, filled into vials, sealed and labeled.

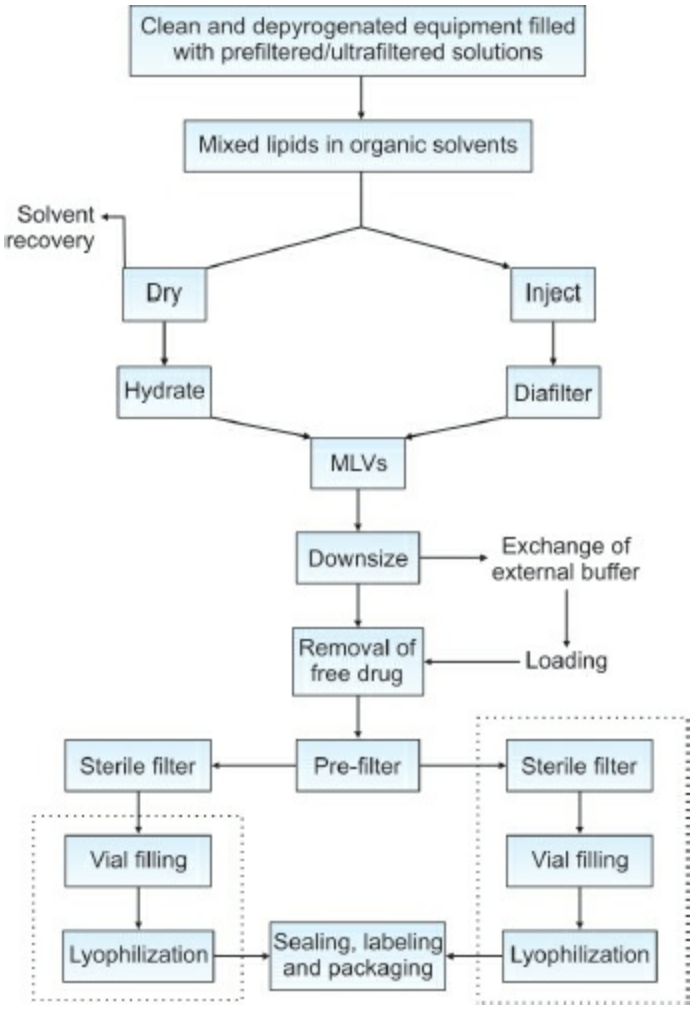



**Fig. 24.18:** Various steps for the commercial manufacturing of liposomes (Doxil$^{TM}$). Steps indicated in the box with discontinuous lines should be performed under the aseptic conditions

Fig. 24.18 schematically presents the manufacturing process and scaling up of liposomal doxorubicin (Doxil™) as an example.

# 25:
# Targeted Drug Delivery Systems

The therapeutic response of a drug depends mainly upon the interaction of drug molecules with cell on the cell membrane-related biological events at receptor sites in concentration-dependent manner. To achieve this goal, the correct amount of drug has to be delivered to the site of action along with simultaneous control of the drug input rate. Generally, after administration of drug/drug-loaded carriers through any route, undesired biodistribution of drug molecules to the healthy tissues takes place, which otherwise limits their clinical utility. This distribution to other tissues, therefore seems unnecessary, wasteful and a potential cause of toxicity. Furthermore, systemic administration of drug molecules encounters limitations viz. rapid clearance from the systemic circulation either by metabolism and/or excretion, inability to access and penetrate the target tissues, or undesired non-specific uptake by sensitive normal tissues and/or cells.

These limitations can be circumvented by “targeting” the drug to the site of action in the required amount by employing various purpose-specific drug delivery systems. These delivery systems, if appropriately engineered, can provide the desired therapeutic response without or with limited side effects associated with conventional drug dosage form. In 1902, Paul Ehlrich proposed the concept of magic bullet. He posulated that therapeutic molecules like drugs; vaccines or macromolecules such as DNA, etc. could be successfully delivered to the desired therapeutic site in the optimal quantity. Targeted drug delivery thus needs successful engineering of a drug carrier or drug per se so that drug molecules could exert their effects with the significant and prompt in vivo response.

One ever sought after, yet unattained goal in clinical medicine remains to be the successful designing and development of site-specific/targeted drug delivery systems. Obviously, the practical realization of the concept shall offer a great breakthrough in medical sciences. Not only will it ensure site specificity, it will in addition, mitigate toxicity of drug(s) on non-target site(s) and well-controlled and effective drug’s blood level(s) (Fig. 25.1). Selective

drug delivery or targeting therefore seeks to improve upon the benefit/risk ratio associated with drugs. Ideally, a drug intended for clinical use should have a high therapeutic index, which is the ratio of drug efficacy (therapeutic effect) and drug toxicity (side effects). This chapter deals with the various types of targeting strategies exploitable for aiming delivery of therapeutic moieties to different body compartments/organs and/or even to intracellular organelles such as nucleus and mitochondria, etc.

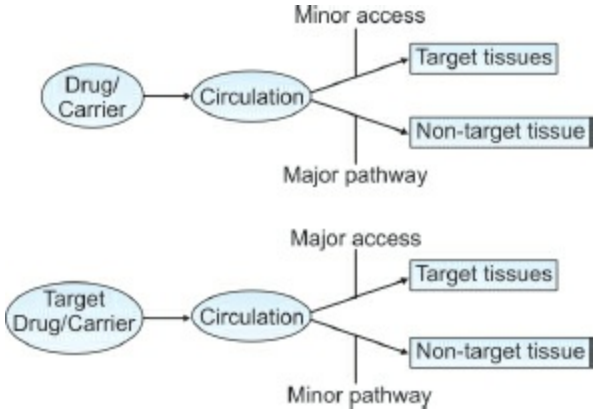



**Fig. 25.1:** General concept of targeted drug delivery

## IMPORTANCE OF DRUG TARGETING

Targeting can be achieved, if the target compartment is distinguishable from the other body compartments, convincingly, on the basis of its biochemistry, surface characteristics or etiological variations, and also if the active drug could be placed predominantly in the proximity of target site. The restricted distribution of the parent drug to the non-target site(s) with effective accessibility to the target site(s) could apparently maximize the benefits of targeted drug delivery.

Many diseases are being treated by cytotoxic agents which only demand for controlled drug delivery but also drug release rate and pattern which is desired to be specific, precise and defined at quantitative levels. The cell-related biological events iteratively occurring in high order of specificity and precision in diseased conditions offer the basis for quantitative targeted drug delivery. A number of essential involved bioligands for physiologic need of cell and these in biosignaling are taken up by the cells. This uptake is executed through some bioports commonly referred to as receptors. The ligand-receptor interactions are highly regiospecific. Thus, ligands or receptors could judiciously be exploited for site-/cell-specific drug delivery, quantitatively, in a well defined manner. A number of essential aspects which should be considered for the designing of drug delivery systems include target, carrier, ligand(s) and physically modulated components. It is therefore inferred that targeted drug delivery essentially implies for selective and effective localization of the pharmacologically-active moiety at preidentified (preselected) target(s) in therapeutic concentration, while restricting its access to non-target normal cellular linings, thus minimizing toxic effects and maximizing the therapeutic index.

The common approaches of advanced drug delivery which have been adopted include (1) controlling the distribution of drug by incorporating it in a carrier system, (2) altering the structure of the drug at the molecular level and (3) controlling the input of the drug into the bioenvironment to ensure a programmed and desirable biodistribution. Applications of the developments in molecular genetics are enabling both, the diagnosis of disease and understanding pathogenesis to the level of clockwork-precision, while advents of novel drugs and delivery systems offer options and opportunities for radical cure of diagnosed diseases.

## DRUG TARGETING STRATEGIES

Drug targeting can be achieved by using carrier systems exploiting both, the intrinsic pathway(s) that these carriers follow, and the bioprotection that they can offer to the loaded bioactive molecule enroute to the target. The various targeting strategies used to carry the drug to the selective site are broadly termed as passive, inverse, active, dual, double and combination targeting (Fig. 25.2).

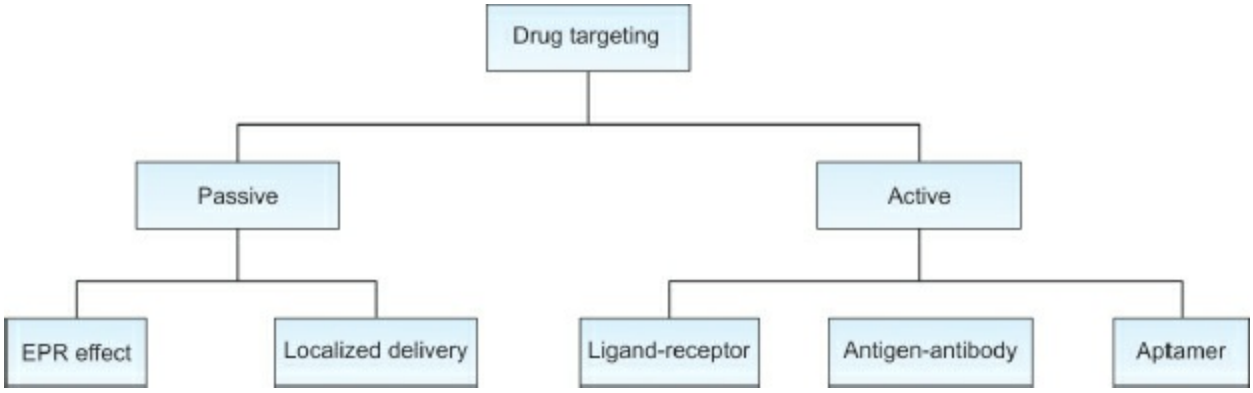



**Fig. 25.2:** Classification of various type of targeting strategies

## Passive Targeting

In general, “passive” delivery stems up as a result of the body’s natural response towards a drug-carrier system. Classically, passive targeting utilizes the natural course of biodistribution of the carrier system through which, the drug may eventually accumulate in an organ or compartment(s) of the body. Some colloids are taken up by the reticuloendothelial system (RES) and hence, emerge as ideal vectors for passive targeting of drugs to RES predominant compartments. Similarily, passive capture of colloidal carriers by macrophages offers therapeutic opportunities for the delivery of anti-infective agents to the wondering and fixed macrophages, in case of oppurtunistic infections such as leishmaniasis, brucellosis, candidiasis and tuberculosis.

## Inverse Targeting

Inverse targeting is a result of the avoidance of passive uptake of colloidal carriers by the RES. This process leads to the reversion of biodistribution trend of the carrier and hence, the process is referred to as inverse targeting. One strategy applied to achieve inverse targeting is to suppress the function of RES by pre-injection of a large amount of blank colloidal carriers or macromolecules like dextran sulphate. This approach leads to RES phagocytic saturation and as a consequence, impairment of host defense system. Some other strategies are based upon the modification and defined manipulation of the size, surface charge, composition, surface rigidity and hydrophilicity characteristics of carriers for desirable biofate (Fig. 25.3).

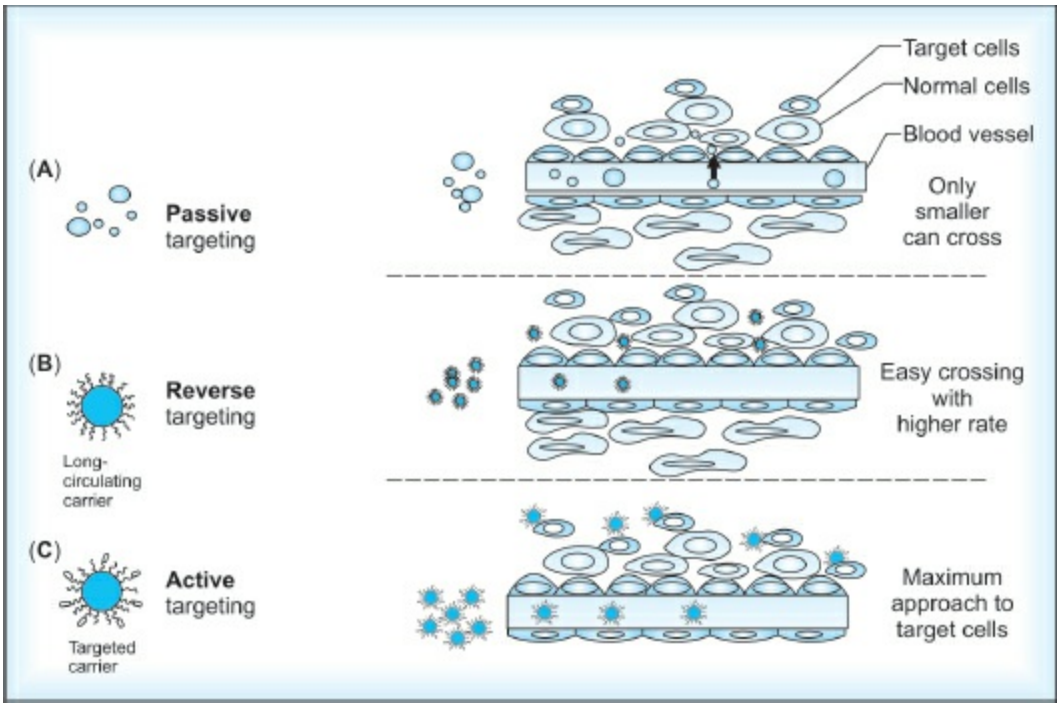



**Fig. 25.3:** Schematic representation of the biofate of different types of drug-loaded carriers in tumour vasculature: (A) Non-targeted drug-loaded novel carrier system; (B) Long-circulating, but targeted drug-loaded novel carrier system; (C) Long-circulating ligand-appended targeted novel carrier system

A few polymers such as Poloxamine 908 and Poloxamer 308 have been successfully used to evade RES uptake and impart prolonged systemic circulation to the carrier. Surface hydrophilicity presumably reduce or even

eliminate the adsorption of opsonin or HDL on to the surface of carrier, which is believed to negotiate and initiate the process of phagocytosis that is ultimately responsible for the uptake of the carrier by RES system.

## Active Targeting

Active targeting involves the modification or functionalization of drug carriers so that the contents are delivered exclusively to the site corresponding to which the carrier is architected. The natural distribution pattern of the drug carrier is modified using chemical, biological and physical means, so that it approaches the identified biosite(s) selectively. Drug carrier binding to target cells is facilitated by using cell receptor-specific ligands or determinant appended with engineered homing devices to increase receptor-mediated (or in some cases receptor-independent but epitope-based) localization of the carrier-drug composite and target-specific delivery of drug(s). Active targeting can be affected at different levels and, accordingly, referred to as first order targeting (organ compartmentalization), second-order targeting (cellular targeting) and third order targeting (intracellular organelles targeting).

### *First-order Targeting*

First-order targeting refers to size-dependent distribution and retentive accumulation of the drug-carrier system on to the capillary bed of a pre-determined target site, organ or tissue. Targeting to lymphatics, peritoneal cavity, plural cavity, cerebral ventricles, lungs, joints, eyes, etc. represents first-order targeting.

### *Second-order Targeting*

The selective delivery of bioactives to a specific cell type such as tumour cells and omitting/excluding the normal healthy cells is referred to as second order drug targeting or cellular targeting. Selective drug delivery to the Kupffer cells in the liver exemplifies the referred order of targeting.

### *Third-order Targeting*

Third-order targeting is defined as drug delivery specifically to the intracellular organelles of the target cells. For example, the receptor-based ligand-mediated entry of a drug complex into a cell by endocytosis, lysosomal degradation of carrier followed by release of bioactives intracellularly to the cytosol, mitochondrion, etc. or gene delivery to nucleolus can be categorized as intracellular drug delivery of third order.

## Ligand-mediated Targeting

Most of the carrier systems are colloidal in nature and can be specifically functionalized using various biologically-relevant molecular ligands including antibodies, polypeptides, oligosaccharides, viral proteins and fusogenic residues. The ligands offer specific avidity to the drug carrier. The engineered carrier constructs selectively deliver the drug to a cell or group of cells generally referred as the target. The cascade of events involved in ligand-mediated site-specific drug delivery is discussed in the latter part of this chapter.

## Physical Targeting

Drug delivery programmed and monitored at the external level (ex vivo) with the help of physical means is referred to as physical targeting. In this mode of targeting, some characteristics of the bioenvironment are used either to direct the carrier to a particular location or to affect selective release of its contents. The release of drug from temperature-sensitive liposomes in the vicinity of a tumour (temperature level higher or equal to the phase transition temperature of constitutive lipids) is brought about by serum components, mostly the lipoproteins, which at phase transition induce release of the entrapped drug. Alternatively, physical means including light, laser beams, ultrasound, temperature, or magnetisin as per the properties of the system may be used to modulate drug release from the carrier.

## Dual Targeting

This classical approach of drug targeting employs carrier molecules, which have their own intrinsic biological effect, thereby synergising with the effect of the loaded active drug. Based on this approach, drug conjugates are prepared with fortified activity profile against viral replication. A major target is the virus replication process that can be attacked at multiple stages and hence the possibility multi-drug resitance (MDR) associated with antiviral therapy can be minimized.

## Double Targeting

As new emerging trends, drug targeting may be combined with another strategy, other than passive and active targeting for effective drug delivery. The combination is made between spatial control and temporal control of drug delivery. The temporal control of drug delivery has originally been developed in regard to control drug release prior to the emergence of drug targeting concept. If spatial targeting is combined with temporal control of drug release, it results in an improved therapeutic index by mainly following two effects. Firstly, in case drug release, or activation occurs locally at therapeutic sites, selectivity is increased by multiplication of the spatial selectivity with the local release/activation. Secondly, the improvement in the therapeutic index by a combination of a spatially-selective delivery and a preferable release pattern for a drug, such as zero-order release for a longer time period of the drugs. When these two concepts are combined, the phenomenon is called as “Double targeting”.

## Combination Targeting

Similar to double targeting, the combination targeting systems for the site-specific delivery of proteins and peptides are equipped with carriers, polymers and homing devices with molecular specificity. The latter provides a direct access to the target site. Modification of proteins and peptides with natural polymers, such as polysaccharides, or synthetic polymers, such as poly(ethylene glycol), may alter their intrinsic physical characteristics which play a crucial role in the targeting of a specific compartment, organs or their tissues within the vasculature. Further, vectorization of these modified proteins and peptides through or by using vesicular or microparticulate carriers. The intrinsic (through homing devices) properties of carrier, i.e. dynamic recognition of target, followed by interaction in order to achieve a site-specific active drug delivery of the encapsulated contents are therefore designed and explored.

## CELLULAR BIOCHEMISTRY AND MOLECULAR EVENTS

A large number of specific and non-specific cellular mechanisms have also been explored to facilitate the internalization and uptake of endogenous and exogenous ligands or ligand-carrier composites. Most of them are based on cell-surface receptors (transmembrane proteins) or cell-surface epitopes (determinants) that help arbitrate trafficking of ligands with the help of receptor mediated bioevents, mainly receptor mediated endocytosis (RME). Cell-surface biochemistry and molecular portals have been realized as delivery modules and exploited for site specific (spatial) and temporal (controlled) drug delivery. Cell-surface markers, such as antigenic determinants or specific sequence of receptor sub-units which serve as ligands, have extensively been studied and found to constitute the basis of biochemical and molecular biology-assisted delivery units.

## Molecular Targets for Cellular Targeting

A cell or group of cells can be described as target identified to be in the need of treatment. Some distinctive cellular elements associated with the cell surface are considered to be essential components/determinants in designing of carriers for targeting, as shown in Fig. 25.4. Cell surface-antigens, have judiciously been used for rising and clonal expansion of cell specific and non cross-selective monoclonal antibodies; cell-surface receptors, which recognize and internalize the macromolecular ligands and associated carrier; and cellular oncogenes or oncogenic viruses, serve as molecular targets. Molecular targets represent the most specific form of proteins and peptide drug targeting. Molecular targets provide a means for "active" drug delivery, i.e. the site to which the drug is ultimately needed to be delivered depends on the molecular specificity and composition of the targeting agent rather than the physical characteristics of the drug or delivery system.

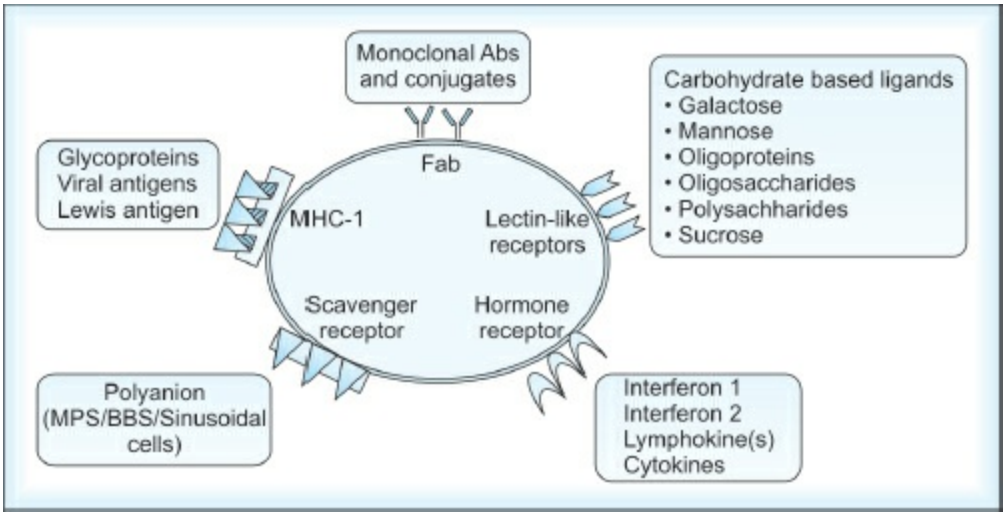



**Fig. 25.4:** Various types of ligands expressed by different cells within body that can be utilized for targeted drug delivery

## Receptor as Delivery Ports

Cell-surface receptors are complex transmembrane proteins, which mediate highly specific interactions between cells and their extracellular milieu. Receptors however, are cellular markers and play an integral role in the regulation of cellular functions including, growth, differentiation, metabolism, secretion, contraction and migration. Two most common functions of receptors are to mediate the transmembrane transport and trafficking of bio-ligands or eluents, and to transduce and regulate transmembrane signaling. Thus, their level of expression may quantitatively affect the cellular uptake/transportation of drug and carrier. Furthermore, cell membrane associated receptors are differentially expressed in various cell types and tissues and they may offer a basis for targeted drug delivery. Biochemical and physiological properties of receptors vary depending on both receptor type and cellular background. It is likely that some receptor systems may be more suitable than others for receptor-mediated drug delivery. Receptors, therefore, could be appreciated as the sensors expressed on target cells (or its intracellular domains) by which cells detect endogenous ligands/macromolecules. Receptors have highly specialized recognition sites with rigid structural requirements for binding of the signaling ligand. Receptors are primarily embedded in the cellular or sub-cellular membrane extending further as glycosylation (carbohydrate) pendent chains exposed on the extracellular side. In the classical sense, an endogenous signaling molecule/ligand binds to its corresponding receptor (ligand binding under affinity and specificity) resulting in to the activation of receptors and a transmembrane signal reaches to the cellular interiors to negotiate the desired response at the intracellular signal reception point. Exogenous ligands could be engineered with specific avidity towards the receptor(s) expressed on the target cell(s) and hence, this strategy could translate the concept of cellular as well as intracellular targeting into an adaptable practicing therapeutic strategy. Receptors are often identified and characterized according to their ligand-binding properties. The regulation of receptor expression and ligand-binding allows cells to modulate the hormonal and metabolic factors to meet the dynamic needs and interact with the surroundings in order to initiate a desired and specific effect.

Selective, defined and saturable ligand binding is a hallmark of receptors.

These receptors are distributed as traffics between the plasma membrane and various intracellular compartments, such as endocytic vesicles and the Golgi apparatus. Certain types of receptors posses an intrinsic tyrosine kinase activity, which plays an essential role in their signaling, while other receptors initiate their post-receptor signaling by coupling with cytosolic enzymes. Areas of potential importance to receptor-mediated drug delivery include various receptor functions and cellular events like endocytosis, transcytosis, ligand-receptor complex, disposition and receptor regulation.

## Ligands as Delivery and Targeting Tools

In drug delivery, reference ligands are carrier surface group(s), which can selectively direct the carrier to the pre-specified site(s) housing the appropriate receptor units to serve as 'homing device' to the carrier/drug. The ligand-appended carrier bio-interaction serves to assist the presentation of ligands to their respective receptors localized on the cellular surface. Various ligands exploited for selective drug targeting include antibodies, polypeptides, oligosaccharides, viral proteins, endogenous hormones and fusogenic residues etc. The ligands, therefore, confer recognition and specificity upon drug-carrier and endow them with an ability to approach the respective target selectively and deliver the drug. Ligands are either covalently anchored or noncovalently associated with the surface of the carrier such that the receptor-interacting determinants remain well-exposed and accessible for receptor recognition.

### *Ligand-driven Receptor-mediated Drug Delivery*

Design and development of potential carriers for cell-specific delivery of therapeutics are immensely dependent on the selectivity of the ligands with the cellular receptors distributed variably at intracellular sites and on the surface of cellular systems. Other crucial factors include the anatomical and pathological barriers that have to be circumvented, en route before recognition site(s) are reached. Intra-cellular mesogenic constraints as well as physiologic constraints are also encountered following receptor recognition. Similarly, cellular internalization and subsequent trafficking are equivocally critical for intracellular routing, degradation and release of carrier and its contents. Table 25.1 summarizes some important receptors expressed over the different cells.

**Table 25.1:** List of various cell types and their subsequent receptors

| Cell types/expression | Receptor(s) |
|---|---|
| Monocytes | Mannose-6-phosphate-t(n)GP, β-glucan |
| Hepatocytes | Galactose-t(n)GP (high-density), HDL, LDL, EGF, IgA, transferrin |
| Enterocytes | Maternal IgG, dimeric IgG, |

| | |
|---|---|
| | transcobalamin II |
| Macrophages | Mannose-6-phosphate-t(n)GP, galactose-t(n)GP, Man-Glc NAc-t(n)GP, fucosyl-terminated glycol-conjugates |
| Kupffer cells | Mannose-t(n)GP, galactose-particles, polymeric negatively-charged proteins, complement factors, fucose, LDL, fucosyl-based glycoconjugates |
| T4 and other cells of immune system | Galactose-t(n)GP (low-density), interleukin, transferrin, CD4, complement |
| Fibroblasts | Mannose-6-phosphate-t(n)GP, transferrin, transcobalamin II, LDL, EGF |
| Endothelial cells | |
| i. Liver | Monomeric negatively-charged proteins, Man-GlcNAc--t(n)GP, Fc receptors |
| ii. Blood brain | Transferrin, insulin |
| iii. Lung, diaphragm and heart | Albumin |
| Mammary acinar cells | Growth factor |
| Renal tubular cells | Low molecular weight proteins (cationic) |
| Tumor cells | Transferrin, folate, EGF receptor, fucose, TNF receptor, cell-adhesion components |
| Leukocytes | Integrin, selectin, cell adhesion molecules (CAM) |

## Cellular Transmembrane Transport System(s)

### *Endocytosis*

Endocytosis is defined as the process of internalization of plasma membrane with concomitant engulfment of extracellular cargo/fluid. The process serves to selectively retrieve and assimilate various macromolecules from the extracellular fluid needed for a variety of cellular functions. It is the main cellular activity that operates in the internalization of the extracellular fluid and their vesicular coat proteins, which are subsequently processed via different pathways to appropriate cellular and intracellular targets. Endocytosis can be classified as phagocytosis and pinocytosis. Phagocytosis is the engulfment of the endogenous and exogenous particulate materials, such as bacteria, erythrocytes, latex beads, colloidal particles and immunoglobulin molecules, a process. It is performed by the phagocytic cells of the RES including Kupffer cells of the hepatic sinusoids, tissue fixed macrophages (histocytes) and blood macrophages or monocytes. The process involves sequential steps of “recognition” (mediated by coating of blood components, mainly by opsonins and high density lipo-proteins), “adhesion” (attachment of the particle to the macrophages of the RES) and “digestion” (whereby, the particles are transferred to phagosome, phagolysosome and finally to digestive vacuoles). Multiple attachments of particle associated ligands with membrane receptors is an essential stimulus for phagocytic capture of particles (Zippering). Endocytosis processes are schematically shown in Fig. 25.5.

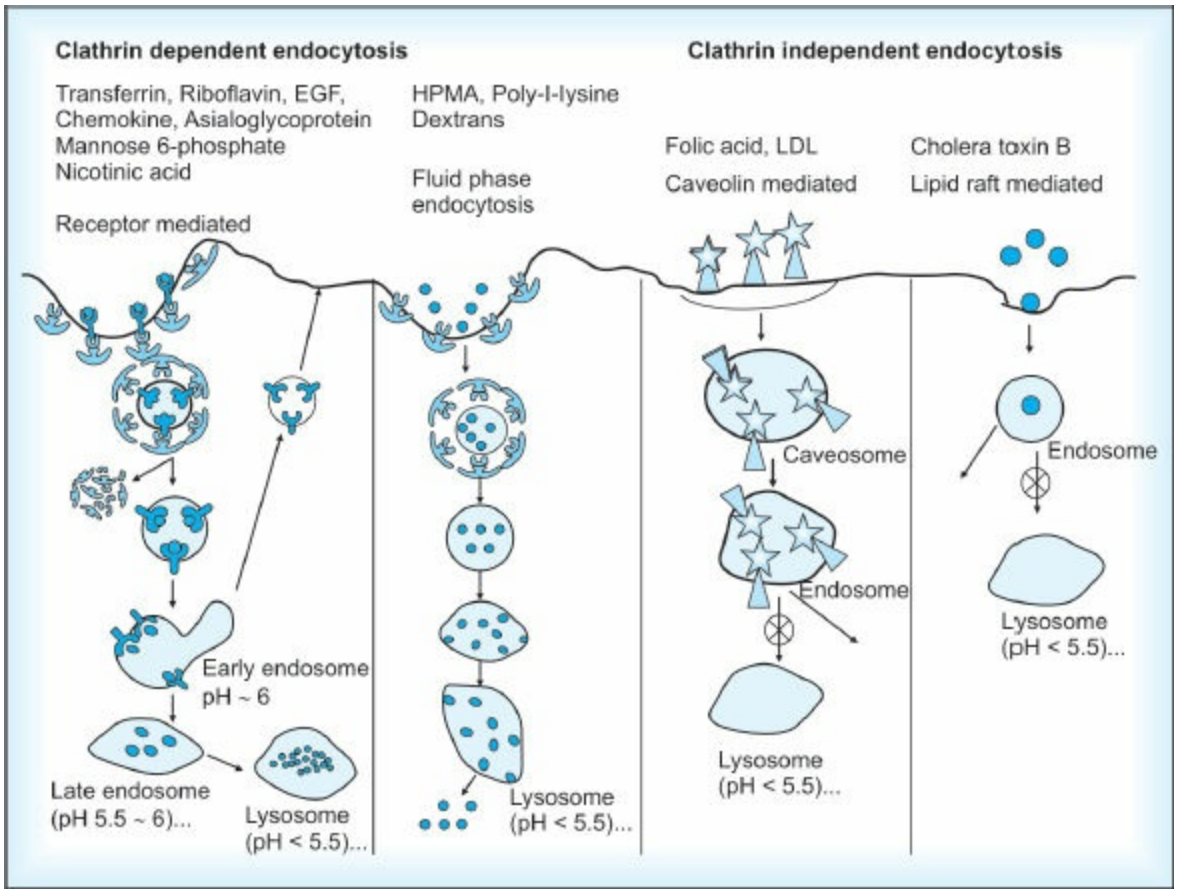



**Fig. 25.5:** Schematic presentation of endocytosis processes commonly involved with the ligands used for drug targeting

On the other hand, pinocytosis is the uptake of small solutes and small droplets of extracellular fluid, shown as cell eating, process known as ‘cell drinking’. It is similar in many aspects to phagocytosis. It too involves the formation of intracellular vesicles from the plasma membrane, like ‘phagosomes’. These ‘pinosomes’ can carry extracellular materials for digestion to the lysosome. It occurs in nearly all nucleated cells. Table 25.2 enlists different lectin receptors and their subsequent ligands.

**Table 25.2:** Various lectin receptors and their subsequent ligands used for targeted drug delivery

| Lectin receptor | Ligand(s) | Cellular expression | Cell event |
|---|---|---|---|
| Asialoglycoprotein receptor/GlcNAc receptor | Asialo-fetuin, asialo-transferrin ceruloplasmin, asialo-lactoferrin, orosomucoid, lac-BSA, α2-macro-globulin and hepatoglobulin | Hepatocytes | Endocytosis |
| ASGP receptor | Oxidized LDL-terminated ligands | Hepatocytes | Endocytosis |
| Galactose particle receptor | Gal-particle (Tris-gal-chol) | Macrophage (Kupffer cells) | Endocytosis |
| Mannose/GluNAc receptor | agalacto-orosomucoid, ovalbumin, β-glucuronidase, mannan | Monocyte derived sinusoids/mature macrophages | Endocytosis |
| Phosphomannosyl receptor | | | |
| 1. Cation independent | Man-6-P (phosphorylated mannose (Insulin like growth factor-II) | Freshly isolated ligands) | Endocytosis monocytes |
| 2. Cation dependent | Man-6-P (phosphorylated mannose ligands) | Freshly isolated monocytes | Lysosomal enzyme trafficking |
| Mannosyl-fucosyl receptor | | Macrophages, Hepatic endothelium | Endocytosis, secretion of glycoproteins |
| Fucosyl receptor | L-fucosyl glyco-conjugates, Gal-BSA, fucosylamine-conjugates | Rat kupffer cells, murine leukemia cell L1210 | Endocytosis, over expression |
| β-Glucan | β-3 D-Glucan on heat killed yeast, zymogen glucan particles | Peritoneal macrophages, human monocytes | Endocytosis |
| Lymphocyte homing receptor | Vascular addresins | Lymphocytes | Recirculation of Lymphocytes to lymphoid organs |

## *Receptor-mediated Endocytosis (RME)*

Based on the kinetics and other biochemical characteristics three cellular transmembrane transportation and internalization mechanisms have been proposed. Therefore, fluid phase pinocytosis; adsorptive, receptor mediated pinocytosis, and adsorptive, non-receptor mediated pinocytosis. As already discussed, pinosomes only capture liquid and the soluble solute(s) present in the extracellular fluid (fluid phase pinocytosis). However, endocytic uptake is a prominent feature. This may proceed through adsorptive pinocytosis, whereby a solute initially binds to the exterior surface of the plasma membrane which is then drawn into the cell interior by a pinosome. The pinosome sacs encapsulate a solute concentration higher than that in the ambient liquid. Adsorptive pinocytosis can be highly efficient and could facilitate higher uptake rate as compared to substrate captured through fluid

phase. Further, adsorptive pinocytosis can be substrate-specific (receptor-mediated) as well as non-specific (non-receptor-mediated). In the former, the cell surface recognizes and internalizes a liquid of defined chemical composition, whereas in the latter, substrate specificity is much broader, and vague.

Free drug and drug-carrier composites enter the cells following different modes and mechanisms. Free drug enters the cell interior via transmembrane diffusive transport or adsorptive (non-specific) pinocytosis, while cellular uptake of drug-carrier composites is mostly through receptor-mediated endocytosis (RME). Exogenous particulate material is also taken up by adsorptive non-receptor mediated endocytosis. RME is a well defined biological process of cell that operates for the uptake of many endogenous and exogenous ligands. Specialized receptor proteins have been identified to operate for nutrients (LDL-cholesterol, Tf-iron), growth factors (EGF, insulin), viruses (influenza), toxins (diphtheria), glycoproteins (galactose-terminating or mannose-terminating glyco-conjugates) and negatively charged macromolecular ligands.

## *Ligand-mediated Transcytosis*

Transcytosis is the process by which intracellular ligand or extracellular particulate(s) are internalized at one domain of a polarized cell and transported by vesicular carrier intermediates to the contralateral plasma sides of cell membranes. It is the principle mechanism that governs receptor-mediated uptake of ligand(s). Ligand transcytosis suggests that exclusive localization of particulate does not occur necessarily within sinusoidal cells, but the particulate ligands may also be taken up in hepatocytes (in case, hepatocytes are not capable of endocytosing initially). Colloidal gold particles coated with either lactosylated serum albumin or mannan are taken up initially by liver macrophages and endothelial cells. However, subsequent redistribution to hepatocytes that confirms ligand transcytosis operates and is responsible for the redistribution of captured particulates. Some of the specific proteins, rarely TAPs (transcytosis-associated proteins) are particularly found on transcytotic vesicles and are believed to be required for vesicle fusion with the target membrane.

## Intracellular Processing and Disposition of Drug-carrier Composites

### *Receptor Recognition and Ligand-receptor Interactions*

Receptor recognition of carrier is a prerequisite in ligand-mediated targeting. The concept involves cell-specific recognition of the carrier binding of drug conjugate, followed by internalization and intracellular release and retention of the active drug. These recognition sites (receptors) serve to support the endocytosis of receptor-ligand conjugate (complex) resulting in intracellular delivery of drugs and genetic materials from ligand-associated carriers. Although such receptors often provide help for internalization and subsequent intracellular transport to degradative compartments (lysosomes), the relative kinetics of intracellular transportation processes in various cell types vary markedly. Wherein, only external binding occurs. In the latter case, local release (bystander) of drug from carrier in the micro-bioenvironment of the cell membrane provides for a concentration gradient-dependent diffusive driving force for cellular delivery/uptake of drug. Both receptor and its density as well as affinity to a ligand substrate in the presence or absence of a competing endogenous ligand determine the degree of carrier receptor occupation (interaction) and the consequent extraction of carrier-associated drug complex by the target tissue(s).

Ligand-receptor interaction in association with endocytosis at the cellular level results into site specific delivery of drug(s). Advents in molecular biology has succsecced in the identification of functionalized receptor(s) and corresponding ligands which have been assessed as novel navigators for selective delivery of drug(s). A drug conjugated to a ligand (neo-ligand) may obviously be synthesized considering that functional integrity of the active species is retrived and preserved after the ligand is dissociated or released. The ligand, however, may be a small molecule or a macromolecule such as protein. Thus, it can conclusively inferred that drug delivery strategies are based upon the ligand immobilized on a drug carrier and ligand-receptor interactions with the carrier.

### *Intracellular Processing of Receptor-ligand Complex*

Intracellular transport and processing after receptor-mediated endocytosis and

transcytosis vary markedly between different receptor-ligand systems and different cell types. The ultimate distributive fate of the drug-carrier composites to the specific intracellular destinations is determined by the receptor involved in endocytosis. Subsequent to ligand-receptor dissociation, recycling of receptor to the plasma membrane and subsequent transportation of ligand to the lysosomal compartment is the most widely executed cellular pathway of endosomally trapped carriers. In this case, the ligands dissociate from their receptors in acidic environment of the endosome and presented to the lysosome, while the receptors are recycled via transport vesicles back to the cell surface (Fig. 25.6). This enables the target cell to endocytose the extracellular ligand at a constant rate for hours. On an average taking 10–15 min for each round endocytic cycle comprising of the pathway beginning from plasma membrane to endosomal compartment to uncurl compartment, where receptor is cycled back to the plasma membrane.

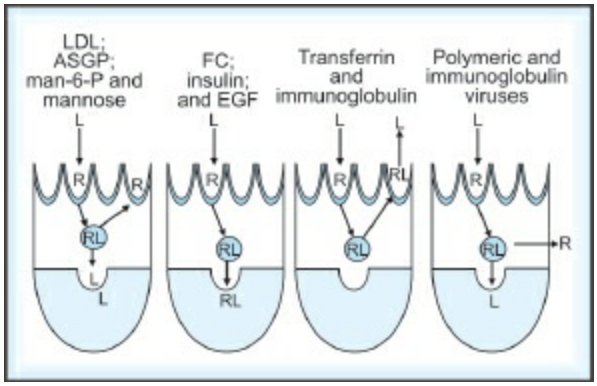



**Fig. 25.6:** Pathways of receptor internalization and recycling. Subsequent to entry into acidic endosomes, ligand and receptors are sorted and trafficked independently which may result into degradation, recycling or transcytosis of either molecule. L = Ligand, R = Receptor

## *Delivery of Drug-carrier Composites to Acidic Endosomal and Lysosomal Compartments*

Ideally, target-oriented drug carriers should efficiently deliver the drug to the target cells. They are needed to be rapidly internalized, through endocytic compartments to the lysosomes, wherein free drug may diffuse into the cytosol. Internalization is the key step in receptor-mediated trafficking of the contents to the cellular compartments and organelles. If the target receptor is a molecule that is not internalized, or the internalization is a ligand-

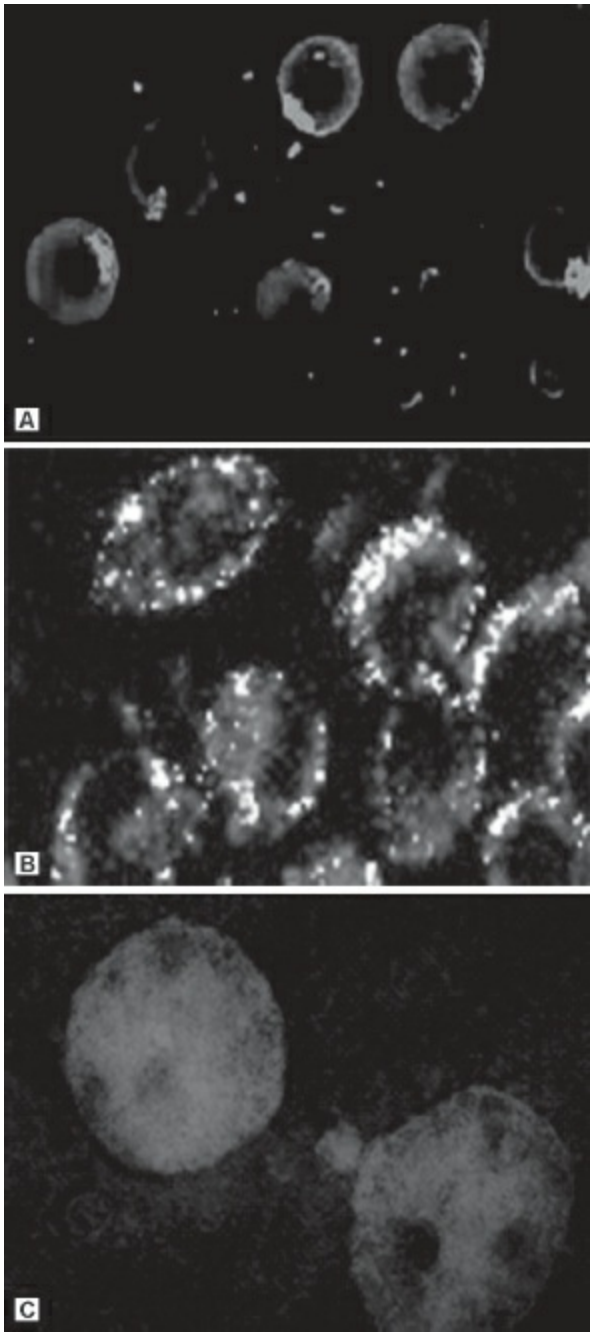
independent non-induced process, then, the targeted drug composite is likely to possess a low targetability potential.

Followed by receptor-mediated internalization, the ligand-receptor complex is routed through an acidic compartment and fuses with its other cellular vesicles for processing of the carrier and contents. This prelysosomal sorting compartment is referred to as compartment of uncoupling of receptor and ligand (CURL), endosome, or receptosome. The endosomal compartment is biologically a complex structure consisting of tubules and vesicles. The vesicles tend to fuse either with interconnected tubular elements called trans-Golgi, where they are biologically scrutinized and selectively modified for subsequent trafficking and targeting; or with the components of CURL. The mechanisms by which endocytic substances are transported from early to late endosomes and finally to lysosomes remain to be vague because the subcellular fate of ligand is dependent on its intrinsic property. Some ligands are internalized, subsequently transferred into endocytic vacuoles, then into the fine anastomoses, i.e. tubulo-vesicular structures, and finally into lysosomes. Others proceed directly from endocytic vacuoles to lysosomes. Two models have been postulated. The ‘vesicle shunt model’ that assumes that early and late endosomes are pre-existing compartments which communicate through vesicle-mediated transport, while the ‘maturation model’ assumes that early endosomes mature on their own gradually into late endosomes.

## *Delivery of Drug-carrier Composites to Cytosolic Compartments*

There have been various attempts to target the drug directly to the cytoplasm by scavenging the endosomal and lysosomal compartments or facilitating their disposition from endosomes to cytosol. Different endosomolytic agents (choloroquine) or lysosomal membrane-active pH-responding amphipathic helices based on fragments of the influenza virus haemag-glutinin are used to promote translocation of immunotoxins or polymer-DNA complexes from endosomes to the cytosol. Studies on virus fusion are consistent with pH-dependent events occurring within a subset of endosomes, which have an appropriate pH value. Instead of showing pH induced sorting (receptor-ligand dissociation), the virus coat proteins undergo a conformational change at pH 6.2, exposing its fusogenic components resulting into the penetration of nucleocapsid into the cytosol (“microinjection”) and forms the basis of

various delivery strategies of viral origin (Fig. 25.7). Vesicle disruption can also be achieved by exposing the cells simultaneously to adenovirus and an immunotoxin. Adenovirus escapes into the cytosol by disrupting the membrane of the endocytic vesicles, and as a consequence, the contents of the vesicles could simultaneously be released into the cytosol.




**Figs 25.7A to C:** Fluorescence micrographs of J774 A. 1 cells treated with various formulations: (A) Palin liposome; (B) Mannosylated liposome; (C)

Mannosylated virus mimicking liposomes (Man-VMLs)

## *Receptor-mediated Delivery of Gene Expression System*

The preferred routes for internalization of the DNA-ligand or DNA-lipid particles (or, otherwise receptor-mediated targeting through virus bearing target DNA) have not been clearly elucidated. However, several studies are carried out on ligand-receptor mediated delivery have been for efficient internalization and delivery of the DNA-bearing particles. The delivery of exogenous DNA or DNA particles to a target tissue by using viruses (retrovirus and adenoviruses) occurs through a series of events ultimately leading to receptor-mediated internalization, followed by appropriate trafficking and expression of the viral DNA (Fig. 25.8). The mechanisms include:

1. Endocytosis into an enveloped vesicle.
2. Escape to the cytoplasmic compartments by fusogenic mechanisms.
3. Dissociation of the viral core and nuclear translocation of DNA.
4. Expression of specific genes included in the virus.

On the other hand, DNA particles targeted through receptor-mediated endocytosis are subjected to the degradative pathways and ultimately transported and digested in the lysosomes. However, a proportion of endocytosed ligands may escape into the cytosol during normal bioevents (diacytosis) or may recycles back to the membrane along with the receptor or may accompany the receptor to the Golgi apparatus where some of the contents may spill into non-degradative compartment of the cell. It is, therefore possible that the association of ligand-attached poly-l-lysine or PLL with the DNA may permit the DNA particle to escape the endosomal and/or lysosomal degradative pathways. This could be happen either due to higher diacytic rate or by reduced degradation of the DNA particle while normal diacytosis occurs, simultaneously leading to a more specific binding with the cells, which bear receptor for the ligands. After internalization, the DNA particles are transported to the nucleus and the genes are expressed. Gene transfer via receptor mediated endocytosis can be augmented by the use of pharmacological agents that disrupt the endocytic trafficking of DNA particles. Chloroquine has been used to increase the expression of transgene delivered via transferrin receptor. Chloroquine acts by increasing the pH

inside the lysosomes and therefore, inhibits the activity of hydrolytic enzymes. However, chloroquine and other lysosomotropic agents, like monensin, may interfere with efficient transfection by inhibiting the recycling of receptor to the cell membrane from pre-endosomal vesicles.

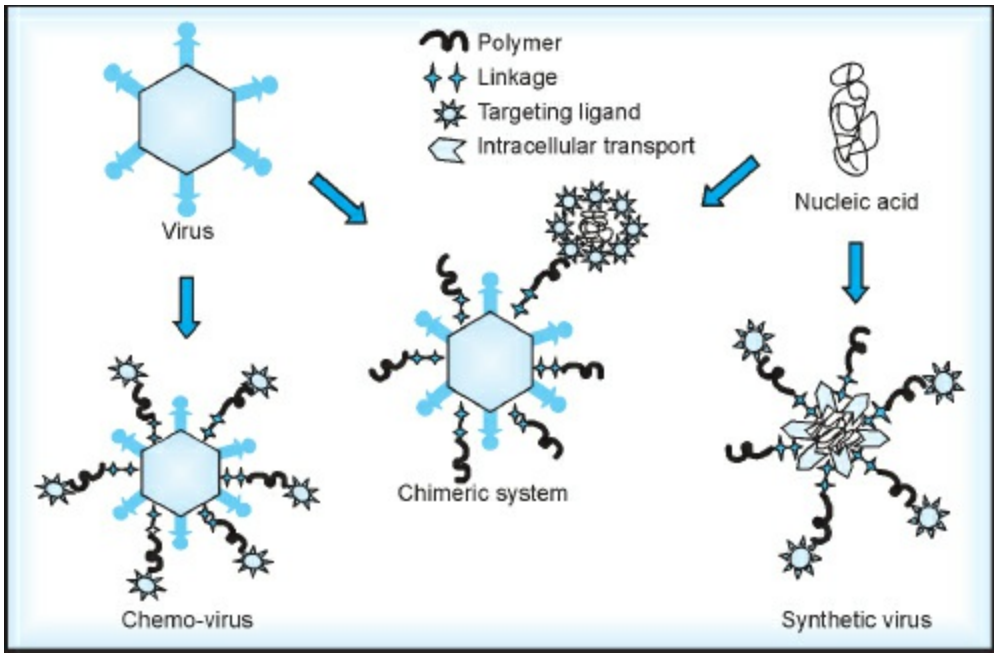



**Fig. 25.8:** Scheme showing different types of viral vectors and synthetic virus systems used for targeted gene delivery

# OTHER APPROACHES OF DRUG TARGETING

## Magnetically-modulated Drug Targeting

An interesting approach of targeting carrier systems has been to magnetize the carriers so that these particles can be retained at or guided to the target site by the application of an external magnetic field of appropriate strength. Retention of magnetic carrier at the target site will delay reticuloendothelial clearance, facilitate extravasation and thus prolong the systemic action of drug (Fig. 25.9). Magnetic fields are believed to be harmless to biological systems and adaptable to any part of the body. Up to 60% of an injected dose can be deposited and released in a controlled manner in selected non-reticuloendothelial organs. Magnetic targeting has several advantages, which include:

1. Therapeutic responses in target organs at only one tenth of the free drug dose.
2. Controlled drug release within target tissues for intervals of 30 min to 30 h, as desired.
3. Avoidance of acute drug toxicity directed against endothelium and normal parenchymal cells.
4. Adaptable to any part of the body.

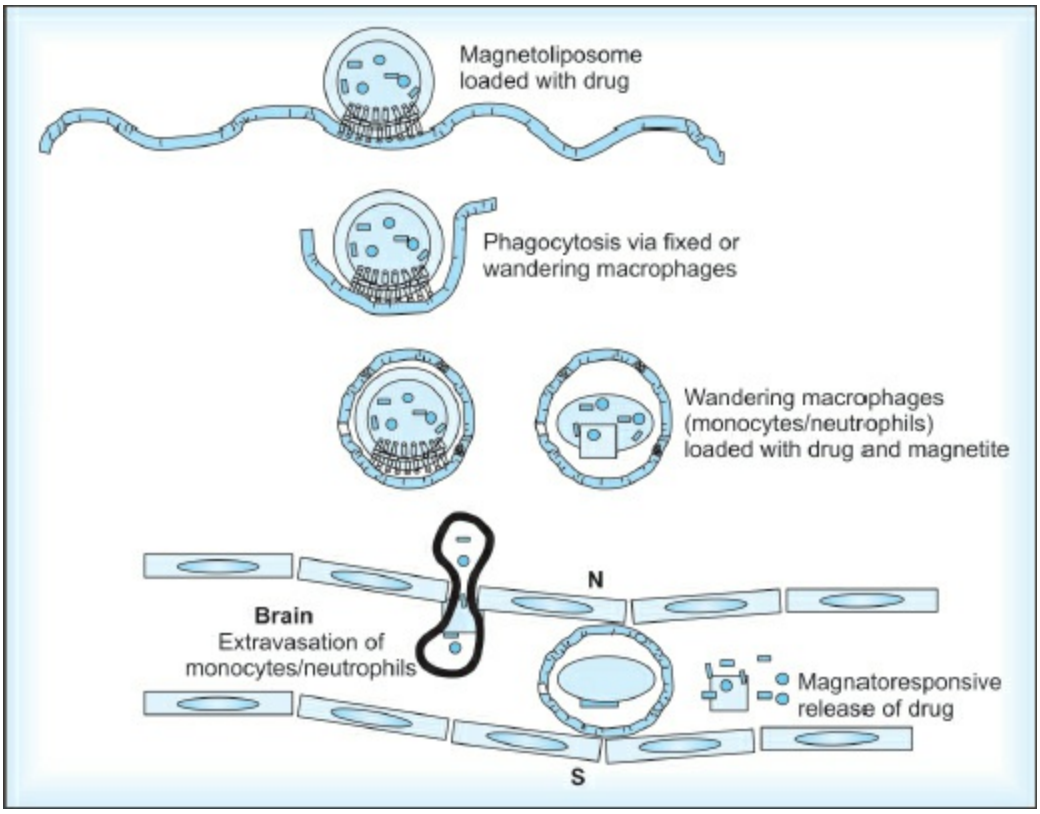



**Fig. 25.9:** Mechanism of drug release from a magnetoreponsive carrier in the brain under the influence of suitable magnetic field

However, this novel approach suffers from certain disadvantages also, as given below:

1. Magnetic targeting is an expensive, technical approach and requires specialized manufacture and quality control system.
2. It needs miniaturized specialized magnet for targeting, advanced techniques for monitoring, and trained personnel to perform procedures.
3. Magnets must have relatively constant gradients, in order to avoid focal overdosing with toxic drugs.
4. A large fraction (40–60%) of the magnetite, which is entrapped in carriers, may be deposited permanently in target tissues.

Magnetic drug delivery by using magnetoresponsive particulate carriers is an efficient method of directing and delivering a drug to the localized diseased site. Magnetic drug transport technique is based on the fact that the drug can be either encapsulated into a magnetic microsphere (or nanosphere)

or conjugated on to the surface of the micro/nanospheres. When the magnetic carrier is intravenously administered, the accumulation takes place within area to which the magnetic field is applied and often augmented by induced magnetic agglomerates. The accumulation of the carrier at the target site allows them to deliver the drug locally. Efficiency of accumulation of magnetic carrier on physiological carrier essentially depends on the physiological parameters, i.e. particle size, surface characteristics, magnetic field strength, and blood flow rate to or through the site. The magnetic field helps to extravasate the magnetic carrier to enable it to reach the target site. An excessively high concentration of chemotherapeutic agents can be estimated near the target site as a consequence of effective targeting, while toxic effects to normal surrounding tissue or to whole body could be minimized. It is, thus possible to deliver large amounts of drug magnetically to the target site, exceeding desirable concentration, which may even reach several fold compared to drug levels achieved following conventional chemotherapy.

## *Magnetically-modulated Microcarriers*

Magnetic microcarriers are nanoparticles that are small enough to circulate through capillaries without producing embolic occlusion (<4 μm). They are, however sufficiently susceptible (ferromagnetic) to become trapped in microvessels and dragged in to the adjacent tissues by magnetic fields of 0.5–0.8 tesla (T). These microcarriers include microspheres, liposomes, cells, nanoparticles, etc. When magnetic force was evaluated for the site-specific drug delivery by using albumin microspheres containing magnetite, it was observed that in the presence of a suitable magnetic field, the microspheres are internalized by the endothelial cells of the target tissue in healthy as well as tumour-bearing animals. Similarily, the magnetically guided microspheres and liposomes have been employed to selectively transport curare like drugs (pyrocurine, diadonium) to the muscles of one of the limbs of the test animals. The use of magnetic microspheres produced no changes in the systemic arterial blood pressure, local blood flow, EEG or ECG. The method could successfully prevent respiratory depression produced by the curare like substances when they are conventionally used for body muscle relaxation.

## *Magnetic Liposomes*

The magnetic liposomes have been prepared and used in cellular sorting

successfully. Liposomes bearing anti-fibronectin antibodies and associated with ferromagnetic particles bound firmly to the surface of mouse embryo fibroblasts. Upon binding magnetoliposomes, the cells could be sorted with the application of a magnetic field. The feasibility of magnetic liposomes as a targeting device for drugs was also explored. Ultrafine magnetite particles encapsulated in the vesicles composed of egg-phosphatidylcholine and tocopherol. The liposomes were targeted to sarcoma implanted in the footpad of rats. A very small amount of the liposomes, but significantly more than the control, was found trapped within the tumour tissue.

Magnetically-responsive polymerized liposomes as potential oral delivery vehicles were prepared to protect complex molecules such as proteins or peptides (e.g. vaccines) from the harsh gastrointestinal environment and targeting them to the Peyer's patches. The intestinal transit of liposomes is relatively fast in mice (1–2 h for small intestine). During this time, liposomes are continuously pushed downward by intestinal flow as well as intestinal motion. As a result, most liposomes fail to have sufficient time to reach the Peyer's patches before they are swept further down. In order to slow down the liposomal transit so as to retain them in the intestine for a prolonged period of time, magnetically responsive liposomes were prepared. These liposomes were localized at the desired areas by exposing them to an external magnetic field. In vivo absorption of the magnetite ($Fe_3O_4$) containing liposomes was examined in mice. After liposomal administration, the mice were given an access to food and water for 40 min before they were restrained and exposed to the magnetic field. In this way, when external magnetic field was applied, the majority of liposomes were localized within the intestine instead of the stomach as measured by the amount of radioactivity retained in various tissue samples.

## *Magnetic Resealed Erythrocytes*

Resealed erythrocytes as drug carriers have been explored for a number of purposes. Magnetically-driven RBCs loaded with aspirin have been developed for the treatment of thrombosis. It was found that local thrombosis in animal arteries could be prevented by means of magnetic targeting of aspirin loaded erythrocytes. It was also noted that a completely occluding red thrombus developed inside the vessel was cleared by placing an external magnet to one of the arteries. The constant magnetic field produced by the

magnet had no influence on clot formation. Autologous red cells loaded with ferromagnetic colloid compound and aspirin after IV administeration, completely aborted arteriothrombosis on magnet application side with no deterio-ratory effect on clot formation in the control artery.

Similarily, magnetically responsive ibuprofen-loaded erythrocytes in optimum concentrations were prepared. Erythrocytes could tolerate ibuprofen as no appreciable detrimental effects were noticed or recorded on cell morphology, osmotic fragility and turbulence shock, when compared with normal erythrocytes. The study suggested the potential of diclofenac sodium-loaded magnetic erythrocytes on the active delivery of drug to the painful inflammed joints, for possible physical modulation of carrier and contained drug biodistribution.

## *Magnetic Emulsions*

Besides magnetically modulated microcapsules/microspheres, magnetic emulsions have also been developed as drug carrier for chemotherapeutic agents. A magnetically responsive oil in water type of emulsion bearing chemotherapeutic agent 1-(2-chloroethyl)-3-(trans-4-methyl cyclo-hexyl)-1-nitrosourea (methyl-CCNU) was developed which could selectively be localized by applying an external magnetic field to a specific target site. The magnetic emulsion consisted of ethyl oleate based magnetic fluid containing CCNU, an anticancer agent, as the dispersed phase and casein solution as the continuous phase. The emulsion exhibited high retention under the effect of a magnetic field in vitro. After IV injection in rat, the magnetic emulsion was mainly localized in the lungs by application of an electromagnet over the chest. Therefore, it was proposed that such magnetic emulsion appears to have potential in site specific delivery of chemotherapeutic agents.

## pH and Temperature-responsive Drug Targeting

Some of the interesting strategies used for enhancing liposome-mediated drug delivery in vivo include the enhancement of stability and circulation time in the bloodstream, targeting to specific tissues or cells and facilitation of intracytoplasmic delivery. The pH-sensitive liposomes have been developed to mediate the introduction of highly hydrophilic molecules or macromolecules into the cytoplasm Fig. 25.10. These liposomes destabilize under acidic conditions of the endocytotic pathway and usually contain phosphatidylethanolamine (PE) and titratable stabilizing amphiphiles. Encapsulated compounds are thought to be delivered into the cytoplasm probably through destabilization of/or fusion with the endosome membrane. Incorporation of a low mole percentage of poly (ethylene glycol) (PEG)-conjugated lipids into pH-sensitive liposomes prolonged blood circulation times of liposomes, which are otherwise cleared rapidly. Antibodies or ligands are appended on the surface of liposome for defined cell-surface receptors ligand interaction for site-specific targeting. Such pH-sensitive liposomes have been used to deliver anticancer drugs, antibiotics, antisense oligonucleotides, ribozymes, plasmids, proteins and peptides to cells in culture as well as *in vivo*.

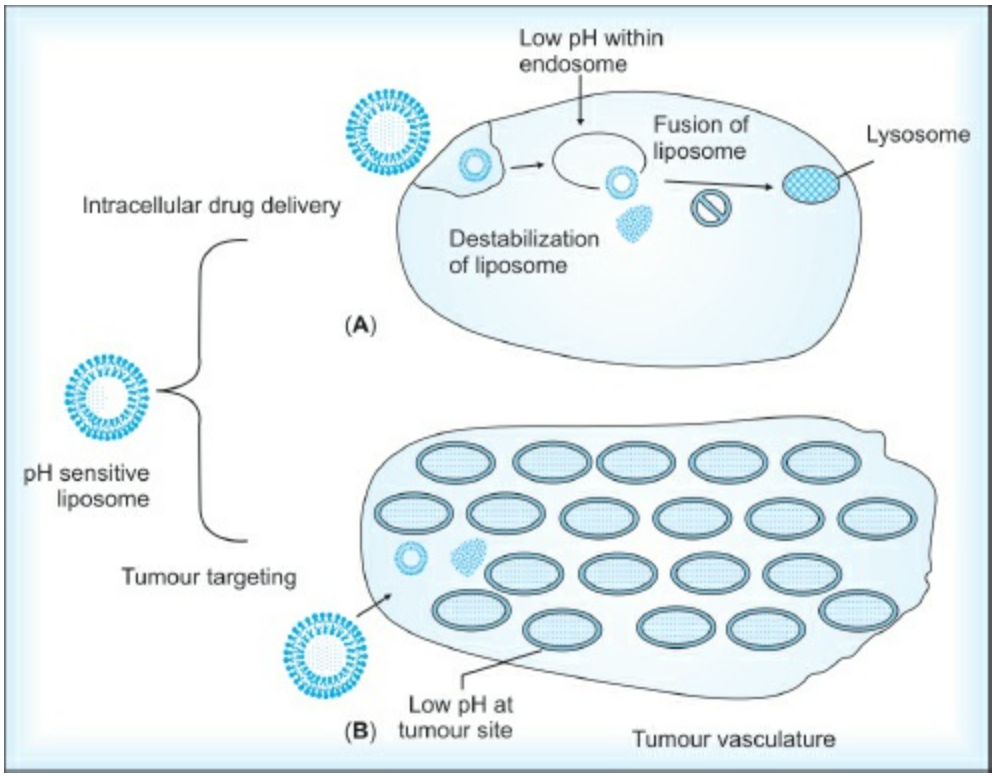



**Fig. 25.10:** Schematic presentation of targeted drug delivery from pH sensitive liposome: (A) Mechanism involved in intracellular drug delivery; (B) Localized delivery of anticancer drug to tumor site

The versatility of pH-sensitive liposomes has been well-illustrated in a wide range of applications. These include (a) the transport of fluorescent probes that allowed not only the evaluation efficacy of different liposome compositions, but also clarification of mechanisms involved in intracellular carrier trafficking; (b) efficient delivery of anti-neoplastic drugs or recombinant proteins; (c) intracellular transport of antigens, aiming at clarifying the intracellular pathways involved in processing and presentation of antigens as well as in enhancing the immune response to tumour cells and (d) the intracellular transport of genetic material for application in gene and antisense therapies.

## Monoclonal Antibody-based Targeted Drug Delivery

Immunoglobulin developed against a surface determinant after immobilization on particulate or colloidal delivery systems could be targeted to the cells bearing Fc receptors via their Fc region or to the target antigen on the cell surface through Fab' region followed by

internalization and release of the encapsulated bioactives. Immunoglobulin fragments Fab' and $F(ab')_2$ or chimeras IgG developed against surface antigens or antigenic determinants are well known strategies employed in Fc receptor-mediated drug/gene therapy. Mabs are now available for all the Fc receptors. Instead of using complete IgG portion, immunologically active fragments [$F(ab)_2$ and Fab'] have been used for improved access to the receptor bearing cells. Recent developments in liposome technology could explore therapeutic applications involving Mabs for site specific delivery. Anti-target monoclonal antibody anchored on liposomes (immunoliposomes) with specific avidity to the cellular determinants or receptor(s) have been investigated to deliver the drugs including idoxuridine and acyclovir in the treatment of Herpes Simplex Virus (HSV)-infected cell lines.

### *Immunoconjugates*

Numerous cell surface markers are now known to proliferate in solid tumours. The possibility of rasing monoclonal antibodies against these markers allows tumour site targeting discretely. However, to make these antibodies a "magic bullet", conjugation with toxic or cytotoxic components is required that leads to the formation of immunoconjugates, which can be another antibody molecule (bispecific antibody), an enzyme (antibody-enzyme conjugates), toxins (immunotoxins), chemotherapeutic agents (drug conjugates), biological modulators such as lymphokines or growth factors. Some immunoconjugates utilize intermediate carrier systems consisting of polymeric molecules such as polysaccharides. Anti-receptor (target) monoclonal antibodies are well established for their role in tumour targeting. The conjugation of these targeting ligands (antibodies developed against a specific tumour determinant) with another recognition component provides them dual specificity to target the drug or toxin intracellularly.

### *Bispecific Antibodies*

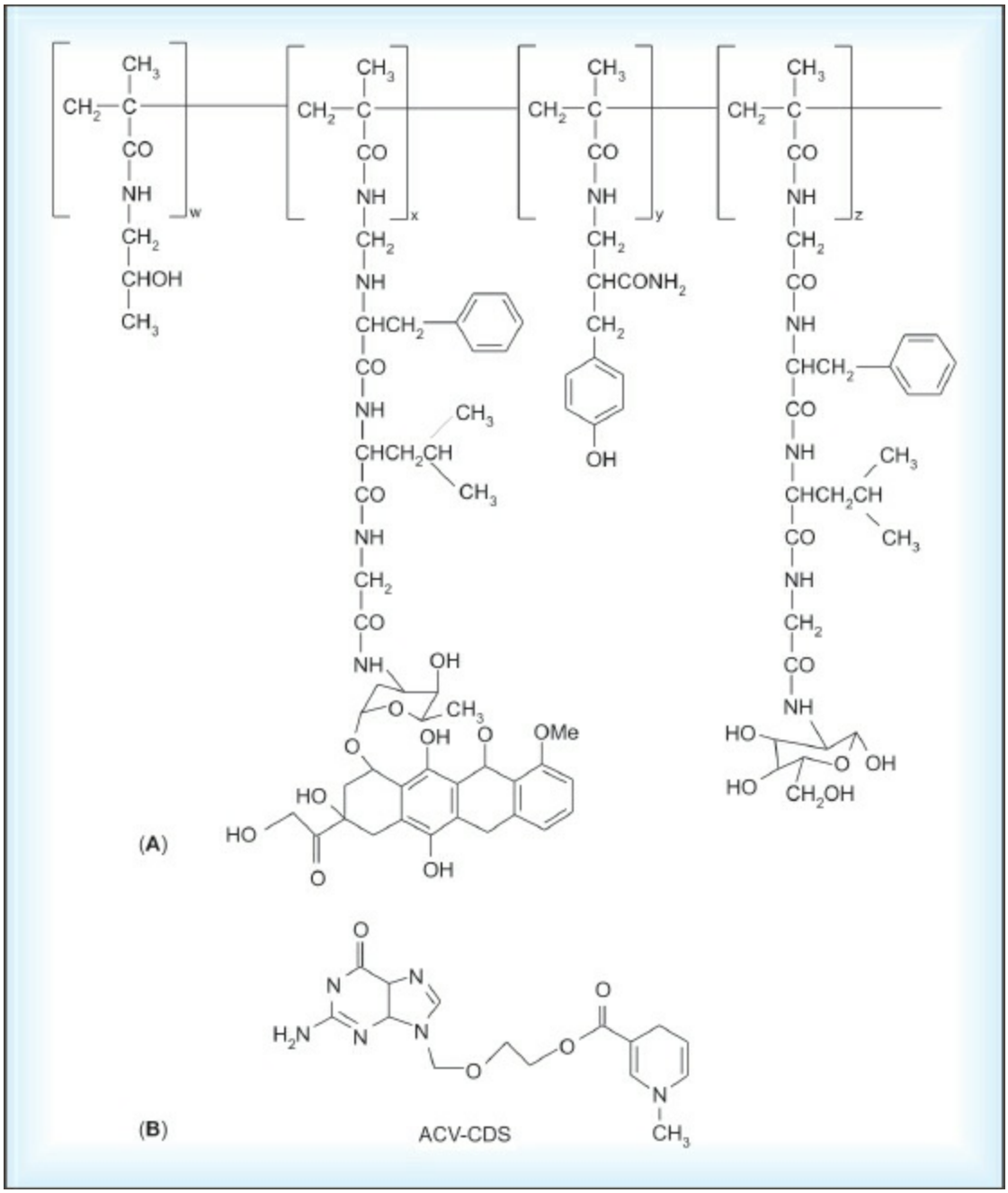
This is an attractive strategy for target-oriented site-specific drug delivery. The approach has been mainly suggested for immunotherapy of immunological disorders especially these related to lack of MHC restricted recognition by immune effector cells. Bispecific antibodies against tumour endothelium on one hand and tissue factors (the initiator of the extrinsic pathway of blood coagulation) on the other hand have also been proposed for the synergistic effects.

## *Immunotoxins*

Immunotoxins are conjugates of antibody (Mab and Fab) fragments and toxins, in which cell binding moieties of the toxins are replaced with specific binding chain of the antibodies. The immunotoxins may be anchored to the surface of carrier systems which provides better projection to them, and effectively presents and places them to recognition site, leading to receptor-mediated endocytosis and eventually killing of cells which express Fc receptors. For the targeting of HIV infections, immunotoxins directed to specific viral antigen or aimed at the highly conserved CD4 receptors on T lymphocytes and several other cell types have been developed.

## Chemical Delivery Systems

A more advanced version of prodrugs is chemical delivery system (CDS) in which the drug is transformed into an inactive derivative, which will then undergo sequential enzymatic transformations to deliver the drug to its site of action. CDSs are utilized for sustained, as well as site-specific targeted drug delivery. These chemically modulated systems can be designed to target specific enzymes or carriers by considering enzyme-substrate specificity or carrier-substrate specificity in order to overcome various undersirable drug properties. This type of chemically-modulated targeted drug design requires considerable knowledge of particular enzymes or carrier systems, including their molecular and functional characteristics. In chemically-modulated prodrug design, enzymes can be recognized as presystemic metabolic sites or promoiety in vivo reconversion sites. CDS is the approach which has been used for targeted delivery of drug to almost every compartment/organ of the body including eye, brain, colon, lymph, intestine, liver, kidney, lung, skin and tumour. Fig. 25.11 shows the structure of some CDS developed for targeting purpose.




**Figs 25.11A and B:** (A) HPMA-doxorubicin galactosamine drug conjugate (PK2); (B) Chemical delivery system of antiviral agent acyclovir for brain targeting

# TARGETED DRUG DELIVERY TO ORGANS

## Targeted Chemotherapy to Tumour

The therapeutics of tumour must catego-rorically is largely confined to directly encounter the tumour cells. The current cytostatic agents are those which interfere with the processes involved in cell growth, while the aim in immunotherapy is to make immune effector cells to selectively attack the tumour cells. A recent approach to fight tumour is to impede and interfere with its blood supply, i.e. turning off angiogenesis or neovascularization.

Chemotherapy is the most widely used therapy of tumor. To lower cytotoxicity and increase therapeutic effects, targeted drug delivery of anti-tumour drugs is highly desirable. The drug delivery systems meant for tumour specific drug delivery mainly includes polymeric carriers or lipoidal colloidal carriers like liposomes, immunoliposomes, microspheres often directed towards the epitopes present on tumour cells and/or receptors expressed specifically or differentially on tumour cells. The target-oriented carriers carry drugs which either interfere with tumour growth or kill them. In all cases, the bioactives are required to cross the tumour blood vessel wall consisting of endothelial cells and basement membrane. Especially in drug delivery strategies in which polymeric, macromolecular or particulate carriers are used to increase treatment selectivity, the endothelial barriers constitute a major obstacle. The focus of research shifts from the manipulation in barrier function or designing of carrier systems that can cross the tumour vasculature, to identification of recognition elements on tumour cells, which could be targeted using suitable ligands. The vasculature endothelium, basement membrane and tumour stroma may contain potential tumour specific targets or molecular determinants. Strategies directed to these identified targets are aimed at interfering with blood vessel permeability, angiogenesis or tumour blood supply, or at manipulating endothelial cell-mediated immune effector cell movement and recognition, thereby, into the tumour tissue. The tumour vasculature is indifferent from the normal vasculature in the following respects:

1. Permselectivity of the tumour vessels is less than that for normal vessels presumably due to the large pores in the vessel wall and hence they are prominently leaky. This could be due to for the absence of basement membrane adjacent to endothelial cells.

2. Hampered and altered expression of adhesion molecules by tumour vasculature endothelium.
3. Turnover time of normal endothelial cells is estimated in the range of 1000 days or more, whereas tumour endothelial cells grow with a turnover time of only 4–5 days.
4. Heterogeneity in angiogenic peptide expression resulting in a heterogenic endothelial cell population.
5. Altered vascular permeability across the tumour mass, the centre being denser and poorly vascularized compared to peripheral region.

## *Vascularization and Localization of Drug Carriers in Tumours*

Overall, the tumour vasculature within tumour tissue is highly disordered with numerous vascular shunts. Vascular haemodynamics and microvessel permeability are of utmost importance in determining the uptake of circulating drug-carrier complexes by tumours. The density of the functional vessels is lower and vascular diameters are irregular and slightly higher compared to normal tissues. Total blood flow, rates of perfusion and vessel permeability may vary significantly in different regions of the same tumours and in tumours of different types, like, spontaneous, transplanted and metastases.

## *Scheme of Site Retention*

## *Barrier(s) Offered by Tumour Vasculature*

The anatomy of the microcirculation in different tissues and organs can reasonably considered to be of prime importance particularily in determining whether drug carriers or antibody conjugates can escape into the extravascular tissue. Similar to extravasation of circulating blood cells, extravasation of delivery systems can presumably be expected to be restricted to capillaries and small-diameter post-capillary venules. The major barrier posed by the tumour vasculature is the basal lamina (basement membrane) on which the endothelium rests and the endothelium itself. Permeability of the tumour blood vessels for transport of small molecules and macromolecular drug-carrier is only sufficient in the blood vessels at the tumour host interface.

In tumour therapy drugs or drug-carrier constructs have to cross the

endothelial linings of the tumour blood vessels to reach the interstitial fluid. This is followed by travelling of carrier through the extracellular matrix towards the tumour cells. In addition to physical barrier posed by the endothelial cells, basement membrane and ECM, high interstitial pressure and low microvasculature pressure may additionally interfere with extravasation of molecules or carriers and cells into tumour tissues. Angiogenesis related factors influence both vascular permeability and immune cell recruitment. Permeability of tumour blood vessels is appreciably higher than vessels in normal tissue.

## *Molecular Targets for Tumour Therapy*

The molecular targets on the surface membrane of malignant cells may conveniently be divided in to following categories and could be targeted using thier respective counter ligand or specially designed antibodies.

1. Altered expression of cell adhesion molecules and their ligands.
2. Altered expression of certain receptors otherwise expressed by all eukaryotic cells, like insulin receptors and MHC Class I associated compound receptors.
3. Exquisite expression of receptors during certain stages of cellular differentiation, like transferrin receptor (TfR), folate receptors, apo-lipoprotein receptor, c-kit receptors, haemopexin receptor and MHC Class II associated compound receptors.
4. Altered expression of certain growth factors (epidermal growth factor receptor, EGFr) and certain vasoactive and angiogenic peptides.
5. Expression of tumour vasculature epitopes, either of the endothelial cells or of the basement membrane supporting the endothelial cells or tumour stroma components (30.5 kD antigen; CD19, CD34; endosialin; endoglin; F19 cell surface glycoprotein; fibronectin; fibrin; myosin and histone).
6. Expression of surface determinants on malignant cells, like Ia antigens and tumour associated antigens (TAA).

*Altered/over expression of cell specific receptors:*

It is appropriate with the ligand-receptor mediated targeting approaches to consider the receptors, which are either over-expressed or down regulated in the malignant target sites. Depending upon the over-expression (folate,

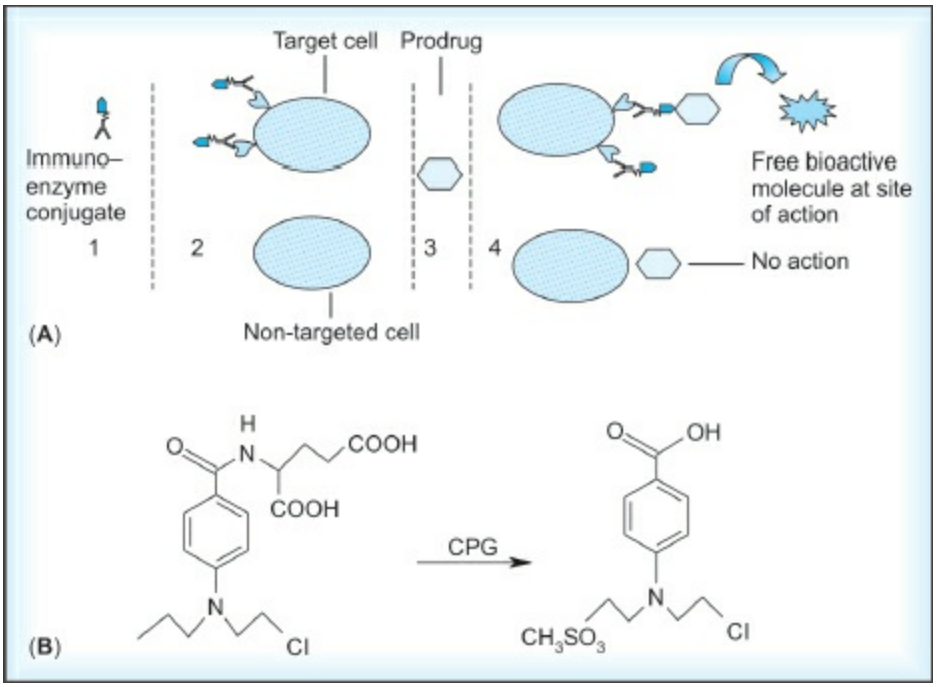
transferrin, fucose and lipoprotein) or down regulation (asialoglycoprotein receptor, mannose receptor), effective cellular targeting approaches could be engineered. For example, $\alpha_v\beta_3$ integrins are overexpressed on actively proliferating endothelium on and around tumor tissue and identified as a promising determinant on angiogenic endothelium. It can interact with various RGD (Arg-Gly-Asp) sequence containing extracellular matrix components. Synthetic cyclic RGD peptides have been shown to bind selectively to $\alpha_v\beta_3$ integrin inhibiting the angiogenesis and thus have been used as targeting ligand for angion-genic entdothelial tumor cells. Targeting strategies encounter with several problems which impede and partially or fully disable itself particularly, in vivo. These include:

- Rapid clearance of targeted systems specially antibody targeted carriers.
- Immune reactions against intravenously administered systems.
- Target tissue heterogeneity.
- Bystander Effect and Binding Site Barrier Concept.
- Down regulation and sloughing of surface epitopes.
- Diffusion and redistribution of released drug.

*Prodrug apparoch for tumor targeted delivery:* Prodrug is an inactive pharmacological moiety developed/modified to optimize pharmacokinetics or site selectivity of a drug. The site selective delivery to destination/target is of great importance particularly in cancer chemotherapy which always demands for reduction in adverse effects of cytotoxic agents. For example, prodrugs have been used to target hypoxic cancer cells, through the use of redoxactivation. They utilize the large quantities of reductase enzyme present in the hypoxic cell to convert the drug into its cytotoxic form. Since the prodrug has low cytotoxicity prior to its activation, it has very few chances of its encounter with healthy, noncancerous cells which reduces the side-effects associated with these chemotherapeutic agents. Some of the important enzymes which involved in prodrug activation in terms of tissue distribution, up-regulation in tumor cells and turnover rates include the following: aldehyde oxidase, amino acid oxidase, cytochrome P450 reductase, DT-diaphorase, cytochrome P450, tyrosinase, thymidylate synthase, thymidine phosphorylase, glutathione S-transferase, deoxycytidine kinase, carboxylesterase, alkaline phosphatase, beta-glucuronidase and cysteine

conjugate beta-lyase, etc. On the basis of enzyme activity, a number of prodrugs have been synthesized or designed for tumor-selective activity of various chemotherapeutic agents including cyclophosphamide, ifosf-amide, cisplatin, mercaptopurine, thioguanine, doxorubicin, daunorubicin, epirubicin, paclitaxel and etoposide, ironotecan.

In general, selective enzyme expression, hypoxia, and low extracellular pH at tumor site is utilized for prodrug activation. However, some distinctive tumor-specific delivery techniques, including activation of prodrugs by exogenous enzymes delivered to tumor cells via monoclonal antibodies (ADEPT), or generated in tumor cells from DNA constructs containing the corresponding gene (GDEPT) have also shown promising results. ADPET has been investigated in the treatment of tumors, whereby an antibody-enzyme conjugate is administered systemically, where it clears from the circulation and localizes to its target by virtue of the antibody binding to its specific biomarker on the tumor. Subsequently, a suitable nontoxic prodrug is then administered, which can be converted to its active cytotoxic form by the enzyme attached to the antibody. The activated drug penetrates the tumor cells and exerts its lethal effect in a localized manner (Figs 25.12A and B). A list of enzymes and prodrugs developed for this purpose is given in Table 25.3. Off course, prodrugs have been relatively successful but they could not achieve complete selectivity. Therefore, more advances in this area are needed to make its complete successful strategy particularity differentiating between tumor and non-tumor cells and to develop optimal substrates in terms of substrate affinity and enzyme turnover rates for prodrug-activating enzymes resulting in more rapid and selective cleavage of the prodrug inside the tumor cells.




**Figs 25.12A and B:** (A) Schematic representation of ADEPT: (1) Development of enzyme-antibody conjugate specific for target site; (2) Administration of enzyme-antibody conjugate, its recognition and binding to target cell while it remains unattached in case of non-target cells; (3) Administration of prodrug as inactive enzyme-specific substrate and (4) Conversion of prodrug into active biomolecule (may be cytotoxic agent in case of tumor) in the vicinity of antibody-specific cells, whereas on contact with non-targeted cell they will not exert any pharmacological action; (B) Cleavage of nitrogen mustard prodrug 4-[N-(2-chloro-ethyl)-N-[2-(mesyloxy) ethyl] amino] benzoyl-L-glutamic acid (CMDA) by carboxypeptidase G (CPG) to cytotoxic 4-[N-(2-chloro-ethyl)-N-[2-(mesyloxy) ethyl] amino] benzoic acid at tumour site

**Table 25.3:** List of enzymes and prodrugs used for ADEPT

| Enzyme | Prodrug | Reactivity |
|---|---|---|
| Alkaline phosphatase | • Doxorubicin<br>• Etoposide<br>• Mitomycin<br>• Phenol mustard | Hydrolysis of phosphate group |
| β-lactamase | • Doxorubicin | Cleavage of lactam |

| | | |
|---|---|---|
| | • Mitomycin<br>• Nitrogen mustard-cephalosporin p-phenylenediamine<br>• Paclitaxel | ring |
| β-glucuronidase | • Anthracycline-glucuronide<br>• Cyanophenylmethyl-β-D-gluco-pyrano-siduronic acid<br>• Epirubicin-glucuronide<br>• Mustard-glucuronide<br>• p-hydroxyaniline | Hydrolysis of β-D-glucuronic acid residues |
| Carboxypeptidase A | • Antifolates<br>• Methotrexate | Hydrolysis of peptide bonds of C-terminal residues |
| Carboxypeptidase G | • Benzoic acid mustards<br>• Nitrogen mustards | Cleavage of l-glutamyl residues |
| Nitroreductase | • 5-(azaridin-1-yl 2,4-dinitrobenzamide | Reduction of nitro groups |
| Penicillin-V-amidase | • Adriamycin-N phenoxy acetyl melphalan | Cleavage of various phenylacetamides |
| | • N-p-hydroxyphenoxy acetamide | Hydrolases, those acting on carbon-nitrogen bonds |
| Cytosine deaminase | • 5-fluorocytosine | |
| Galactosidase | • 5-fluorouridine | Hydrolysis of galactosides into monosaccharides |

The prodrugs that are unable to enter normal cells but are proteolytically activated by peptidases secreted by cancer cells are called extracellularly

tumor-activated prodrugs (ETAP). The ideal ETAP compounds are pharmacologically inactive due to its inability to reach the intracellular compartments of cells but they become active cytotoxic agent at the tumor site. Entry of cytotoxic agent into the cell interior can be restricted through conjugation to a non-cell penetrating, blood-stable peptide. These prodrugs converted back to active moiety by the action of extracellular peptidases secreted by tumor cells. Selective delivery of doxorubicin to cancer cells or their associated angiogenesis is the most promising approach to reduce the side-effects. Small RGD-containing peptides have been used for targeted delivery of ETAP of doxorubicin as shown in Fig. 25.13.

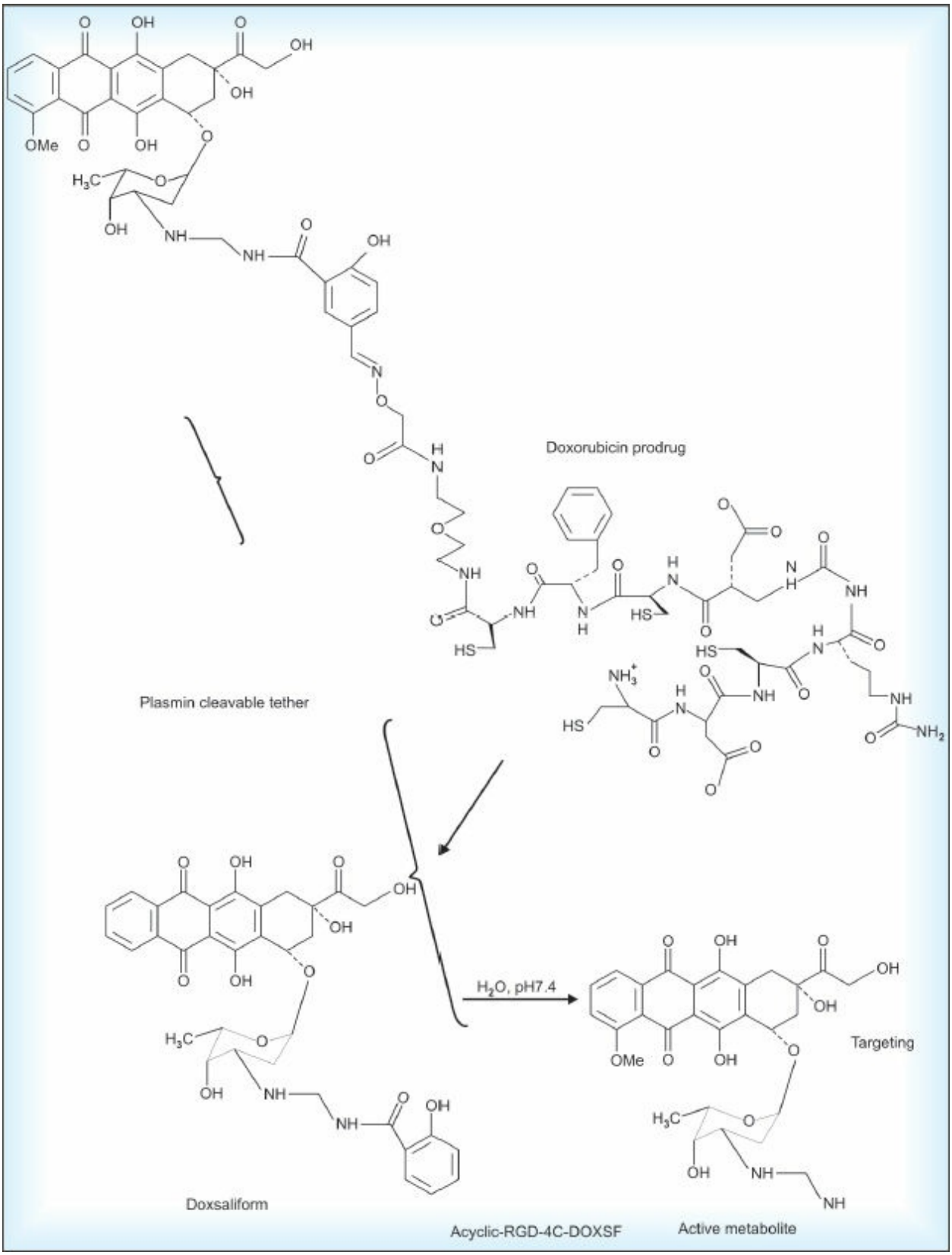



**Fig. 25.13:** Targeted delivery of doxorubicin as ETAP

## Brain Targeting

The advances in our understanding enable us aware of mechanisms involved in pharmacodynamic activity of neuroactive agents, physiopathology and etiology of neurogenerative disorders as well as impediments and limitations to effective therapeutics. The drug accessibility to the central nervous system (CNS) is mainly limited by the blood-brain-barrier (BBB). In the treatment of diseases or conditions that result from the lack of simple hormones and peptides, the administration of these compounds in a controlled fashion could provide effective management of diseases and therapy. Conditions such as diabetic neuropathy, amylotropic lateral sclerosis (ALS) and Huntington's disease and Parkisonism's disease may be treated with better pharmacodynamic effects using targeted drug strategies. Brain related diseases of diversed etiology are the major causes of debilitation, agony and death. The management of brain related diseases with present available therapeutic systems is very difficult, as insufficient amount of drug reaches to the brain, due to highly lypophilic nature of blood-brain-barrier. Drug delivery to the brain requires advances in both, drug delivery technologies and drug discovery. Due to the presence of the blood-brain-barrier only small, lipid soluble drugs in the circulation are ultimately delivered to the brain cells. Therefore, practical strategies are required for mediating drug transport across blood-brain-barrier. The emerging strategies for selective delivery to the brain arise from investigation that critically reveals the physiological mechanisms involved with the solute transport across the blood-brain-barrier. They are schematically shown in Fig. 25.14.

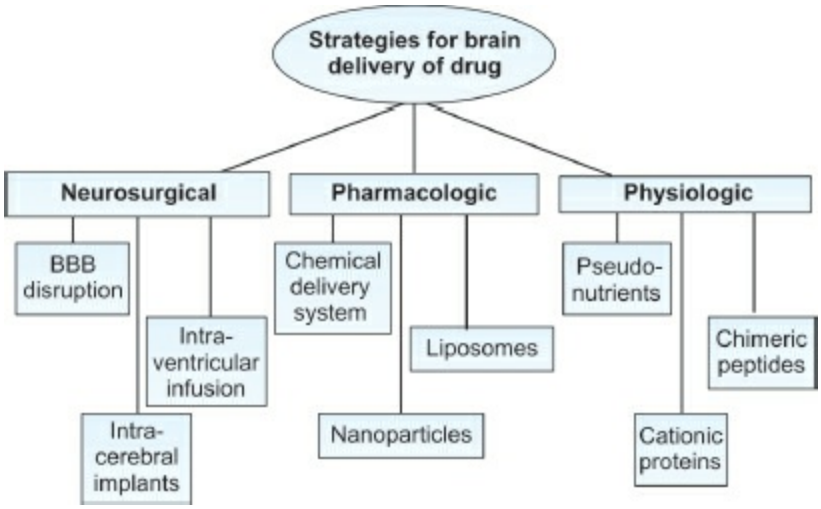


**Fig. 25.14:** Carriers/Strategies of crossing blood-brain-barrier for brain targeting

## *Limitations in Brain Uptake of Drugs*

A number of drug peptide, biological response modifiers and monoclonal antibodies and Fc fragments are presently available which have been proven valuable in inhibiting a variety of malignant infectious diseases and rectification of neurotransmitter and enzyme imbalance in tissue culture systems. However, their in vivo therapeutic efficacy is frequently compromized due to their inability to reach and maintain active concentration at the diseased site located in brain. There are few more cases where inadequate pharmacokinetics can limit drug therapy when a disease is located within the central nervous system. This problem is commonly encountered in patients with acute cerebral bacterial or viral infections, as well as with neurodegenerative diseases, such as Parkinson's, Hunttington's or Tay-Sach disease. However, the most extreme cases are encountered by neurooncologist in treating patients with brain tumours. Transport mechanisms operating at the BBB for peptides and proteins can be classified into the following categories and schematically are shown in Fig. 25.15.

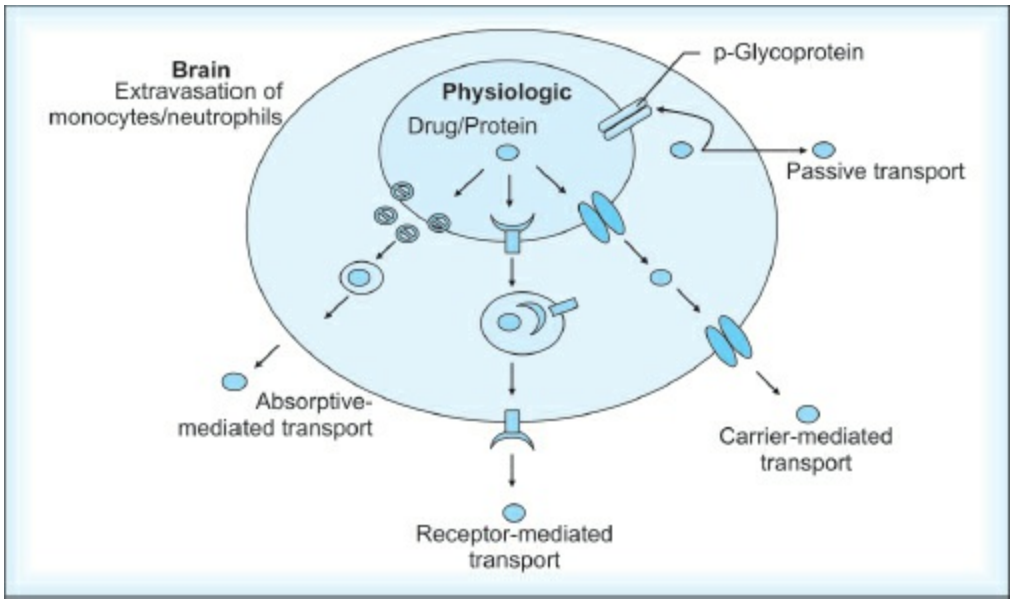



**Fig. 25.15:** Schematic diagram of various types of mechanisms associated with BBB that can be used for targeted delivery of drugs to brain

- Transport via carrier mediated systems
- Receptor mediated transcytosis (RMT)
- Absorptive mediated transcytosis (AMT)

The factors that govern the permeation of a drug across the normal BBB and determine its time dependent concentration within the brain following its systemic administration have been examined by a number of investigators. The factors identified include the followings.

- The time dependent plasma concentration profile of the compound, this is related to its distribution and elimination process.
- The binding of agent to plasma constituents and tissue, and binding off rates from them (plasma clearance).
- The permeability of the BBB to the agent.
- Local cerebral blood flow.

The cerebral availability of peptide/drugs therefore depends on the route of administration. Intracerebral, intraventricular or intracarotid arterial administration provides high concentrations in CSF and diffusion of the peptide drug from CSF through interstitial fluid (ISF) is significantly limited.

## Bone Marrow Targeting

The bone marrow is basically a connective spongy tissue found inside bones. There are two types of bone marrows in the human body, yellow marrow and red marrow. Yellow marrow consists mostly of adipose tissue, and red marrow consists of haematopoietic or blood forming tissue that produces red and white blood cells. The yellow marrow is normally located in the shafts of long bones and the red marrow is found at the ends of long bones and short, flat and irregularly shaped bones.

The bone marrow in the chest, bone, skull, hips, ribs and spine contains stem cells also know as “mother” cells as mature blood cells evolve from these stem cells. The most primitive of these stem cells is the pluripotent stem cell that is believed to be the progenitor or origin of all blood cells. In contrast to a unipotent cell, which differentiates into a single cell type, a haematopoietic stem cell is pluripotent. It is able to differentiate following number of pathways and thereby generate erythrocytes, granulocytes, monocytes, mast cells, lymphocytes and megakaryocytes. Pluripotent stem cells differ from other blood cells in that they are capable both, i.e. unlimited self-renewal and differentiation. Self-renewal is the ability of the cell to reproduce another cell identical to it thus maintaining a steady number of these types of cells in the body. Differentiation is the process of generating one or more subsets of more mature cells that eventually evolve into components of blood either as erythrocytes, neutrophils, eosinophils, basophils, lymphocytes, monocytes or platelets.

The clinical advantages that would arise from a system with bone marrow homing specificity are quite obvious as explained for the under mentioned reasons. Therapeutic approaches might include gene therapy, selective delivery and local release of antimicrobials, as well agents that could induce self-renewal, proliferation and maturation of stem and progenitor cells. The rapid clearance of foreign particles from the circulation by the macrophages lining the sinusoids in the liver, spleen and bone marrow is one of the most important mechanisms in host defense against infections and the same has been exploited for localization of drug delivery systems in the RES-rich organs. Among these are serious parasitic and systemic fungal infections as well as RES-localized lysosomal storage diseases.

Traffic across the MBB is not limited to cellular elements; the

endothelium of bone marrow sinusoids is unique among postcapillary venules in that it removes particulate materials from the blood. The endothelium is not only capable of phagocytic uptake and storage of particles (thus functioning as part of body's reticuloendothelial system) but also provides for transmural passage of particulate material to the extravascular space where the particles are then phagocytosed by central macrophages or by other cells. The successful delivery of particulates to the RES-rich organs focused interest on the possible delivery to other sites in the body, such as various extravascular tissues. This is not only a consequence of the rapid clearance of particulates by the RES but also of natural anatomical barriers in the blood vessels. The anatomy of the microcirculation in different tissues and organs can reasonably be expected to be of prime importance in determining whether drug carriers or antibodies conjugates can escape into the surrounding extravascular tissue. In common with extravasation of circulating blood cells, extravasation of delivery systems can presumably be expected to be restricted to capillaries and small diameters postcapillary venules. Based upon the architecture of the lining endothelium and the underlying sub-endothelial basement membrane (basal lamina), three groups of capillaries are described as continuous, fenestrated, discontinuous or sinusoidal.

Proliferation and differentiation in the haemopoietic system is regulated by haemopoietic cytokines produced by a network of stromal cells in the bone marrow and other cells in the marrow microenvironment. Various studies conducted on rat bone marrow suggested that the luminal surface of bone marrow sinus endothelium possesses both exposed sialic acid moieties (mostly made from sialoglycoproteins) and some unidentified and poorly characterized non-neuraminidase-sensitive anionic material with a pKa higher than sialic acid. These later anionic sites are located throughout the abluminal surface to capture circulating particulate materials for endocytosis across the sinus endothelium of bone marrow, which may be utilized for targeting purpose.

The regulation of expression of specific adhesion molecules by bone marrow endothelium or mature myeloid or erythroid elements may allow for selective exit of these cells out of the marrow parenchyma. Adhesion molecules like CD34, PECAM and thrombospondin have been identified by bone marrow microvascular endothelial cells monolayers. However, in majority they function as a transmigratory bridge permitting the exit of mature cells out of the bone marrow. The selective homing of transplanted

progenitor cells to the bone marrow is also likely to involve expression of specific adhesion molecules. The expression of CD34 on haemato-poietic progenitor cells has been demonstrated. Because mucin can bear lineage-specific carbohydrate determi-nants, CD34 can act as skeleton for the attachment of lineage-specific glycans, allowing progenitor cells to bind to lectins expressed by bone marrow components.

## Macrophage Targeting

Successful delivery of drugs to the desired biological compartment of the host usually depends on the intrinsic properties of the drug carrier. However, it can always be manipulated by appropriate designing of the carrier/delivery system, as little can be done to alter/modify the target particularily macrophage and its surroundings. Various carrier systems have emerged to deliver drugs to macrophages, albeit the efficacy, reliability and selectivity of these carriers are still in question. To date, the most extensively studied carriers are liposomes and microspheres. In fact, physicochemical properties of these carriers can alter their efficacy and specificity to a great extent. These properties include hydrophilicity, surface charge, composition, concentration, and presence of various target specific ligands on their surface. Incidentally, the particulate nature of these vehicles may facilitate passive delivery/homing of the entrapped drug molecules to the macrophages, which may harbour many of the important pathogens in their intracellular compartments, such as Mycobacterium sps, Leishmania and Dengue virus, etc. Moreover, macrophages upon interaction with particulate drug delivery vehicles may act as secondary drug depot, thus helping in localized delivery of the drug at the infected site.

## Lung Targeting

Most drugs targeted to the respiratory tract are used for their local action. For example, ephidrine for nasal decongestion, beta-2 agonists for bronchodilatation, and inhaled steroids to suppress the inflammation seen in asthmatic airways. Since the drug is delivered directly to its required site, only a small quantity is needed for an adequate therapeutic response, and consequently the incidences of systemic side effects tended to be low as compared with oral or intravenous administration. More recently, it has become apparent that the lining of the respiratory tract, from nasal mucosa to airways and alveoli, may be used for the absorption of a drug for its systemic effect. This route of administration may be particularly attractive to avoid the metabolic destruction encountered when a drug under reference is administered through alternative routes (for instance, peptides and proteins are rapidly destroyed by peptidases when given by the oral route). The physiology of lungs makes it possible to target them via two different routes, i.e. through the circulation or through the respiratory tract. In other words by parenteral route or inhalatory administration, respectively. The same drug carriers might be used for either route of administration. Another interesting problem is that lungs can be both the ‘designation point’ for a given drug (the delivery route is not important in this case) or the ‘intermediate stop’ for a drug on its way to the circulation (for the inhalatory route only). In the first case lung pathology is supposed to be treated, while in the second, the lungs might be healthy, however they can serve as reserve for secondary destination delivery of the drug.

Lungs provide an enormous adsorptive area (35–140 $m^2$) capable of delivering even large compounds to the circulation, which opens wide opportunities for inhalatory delivery of many drugs in the form of aerosols. Aerosol therapy using particulate drug carrier systems is becoming a popular method to deliver therapeutic or diagnostic compounds either locally or systemically. This is due to the large alveolar surface area suitable for drug absorption, the low thickness of the epithelial barrier, extensive vascularization and relatively low proteolytic activity in the alveolar space compared to other routes of administration and the absence of the first-pass metabolism. In general, nanoparticle delivery to the lungs is an attractive concept because it can cause retention of the particles in the lungs

accompanied with a prolonged drug release if large porous nanoparticle matrices are used. On the other hand studies have shown that nanoparticles uptake by alveolar macrophages can be reduced if the particles are smaller than 260 nm. Both effects combined might improve local pulmonary drug therapy. However, the particle size of medically used nanoparticles is too small to be suitable for direct lung delivery. A prerequisite for deep lung delivery is the design of proper carrier systems. Successful delivery of inhaled particles depends mostly on particle size and particle density, and hence, the mass median aerodynamic diameter.

## Liver Targeting

The pathogenesis of liver injury and fibrosis involve complicated interactions among different cell populations of the liver, soluble factors, such as cytokines, and the extracellular matrix. Hepatocytes are injured in a number of pathologic processes (chemical, biological and immunological) and the damage initiates fibrogenesis. Kupffer cells probably are primarily responsible for mediating the second phase of hepatocellular necrosis, induced by reactive oxygen intermediates, and for activating Ito cells, the major cell type responsible for enhanced extracellular matrix production during the fibrogenic process. A number of cytokines and other soluble factors, such as tumour necrosis factor-a (TNF-a), interleukin 2 and 6 (IL-2 and IL-6), and nuclear factor-KB (NF-KB) (4) participate in the damage of hepatocytes and sinusoidal endothelial cells; whereas, transforming growth factor-p (TGF-P) and platelet-derived growth factor (PDGF) have been considered to be fibrogenic cytokines which contribute to the activation of cells.

In order to improve therapy to the hepatocellular injury, vesicles that are capable of delivering hepatoprotective agents to the liver constitute an ideal approach that increases local concentration of the agent and also reduces adverse effects in order to achieve maximal therapeutical efficiency. A great number of chemicals or drugs have been evaluated to reduce hepatocyte damage or necrosis and to inhibit liver fibrogenesis, and some of them are highly promising. However, none of them is liver-specific or cell type-specific. Since the majority of liposomes administered intravenously are endocytosed by the reticuloendothelial system (RES), the liposomes are often found highly concentrated in Kupffer cells. Thus, it may be possible to improve the efficiencies of anti-fibrotic agents by reducing the role of Kupffer cells in the fibrogenic process. For examples, inhibition of cytokine expression by specific antisense oligonucleotides entrapped in liposomes could be one alternative for interfering with Kupffer cell effects on injury and fibrogenesis. Such a use of liposomes provides for the potential of therapeutic benefits to many forms of liver diseases. Liposomes are prepared from a variety of lipids and lipid mixtures, with phospholipids the most commonly used. Generally, two kinds of liposomes (oligolamellar and sterically stabilized small unilamellar liposomes) can be classified according to their

composition, chemical features and pharmacokinetics.

The oligolamellar liposomes transport substances through membrane-association, whereas small unilamellar liposomes frequently encapsulate substances in the intraliposomal aqueous phase if the substance is water-soluble. Nevertheless, both types of common liposomes are cleared primarily by the RES and the Kupffer cells in the liver have been shown to have high concentrations of delivered/sequestered contents. Thus, prolongation of liposome circulating time, as well as reducing Kupffer cell uptake and enhancing hepatocyte uptake, are challenges of research in liposome liver targeting. The applications of targeted therapy in liver diseases include attenuation of drug side effects, study of Kupffer cell function, reduction of hepatotoxin-induced liver damage and gene therapy of gentic liver diseases. The modifications of carrier for hepatocyte targeting can be done as shown in Fig. 25.16.

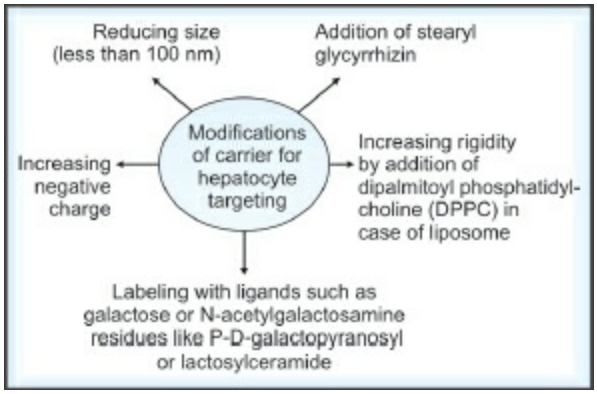



**Fig. 25.16:** Possible mode and mechanism of liver targeting of a carrier

## Lymphatic Targeting

Lymphatic system has a large and extensive system of drainage of waste unwanted content in its fluid throught the body. Drug targeting to intestinal lymphatic system has gain renewed interest for delivery of hydrophobic drugs, macromolecules like peptides, proteins and vaccines. This approach of passive targeting offers several advantages which include avoidance of first pass metabolism, selective treatment of diseases and infections of the mesenteric lymphatics, enhancement of the absorption of the large molecules such as peptides/particulates and inhibition of cancer cell metastasis. A major function of the intestinal lymphatics is to facilitate the absorption of long chain fatty acids following re-esterification and reassembling them into chylomicrons within the enterocytes. Exogenous compounds absorbed via the intestinal lymph are generally transported in association with the lipid core of intestinal lipoproteins as chylomicron. For which they require co-administeration of lipid to stimulate lipoprotein formation and to deliver the content in the systemic circulation through well known transcellular mechanism of lipid transport. Several approaches including prodrug synthesis, use of permeation enhancers, surface modification, complex formation and more recently colloidal lipid carrier based strategies can be used for the delivery of drugs to intestinal lymphatics.

Lymph from the intestinal lymphatic system (as well as hepatic and lumbar lymph) drains through the thoracic lymph duct into the left internal jugular vein and then to the systemic circulation. Thus, the transport of drug through intestinal lymphatic system may increase the percentage of drug that can gain access to the systemic circulation. In addition, the process of intestinal lymphatic drug transport often continues over time periods longer than typically observed for drug absorption through portal vein. Consequently, drug transport through the lymph may be utilized to prolong the time course of drug delivery to the systemic circulation. Since drugs transported to or through the gastrointestinal lymphatic system bypass the liver and avoid potential hepatic first-pass metabolism, therefore, lymphatic delivery of immunomodulatory agents and low therapeutic index drugs used in the treatment of cancer cell metastases and HIV presents an opportunity to maximize therapeutic index of cytotoxic drugs (Fig. 25.17) shows the lymphatic uptake of different class of drugs.

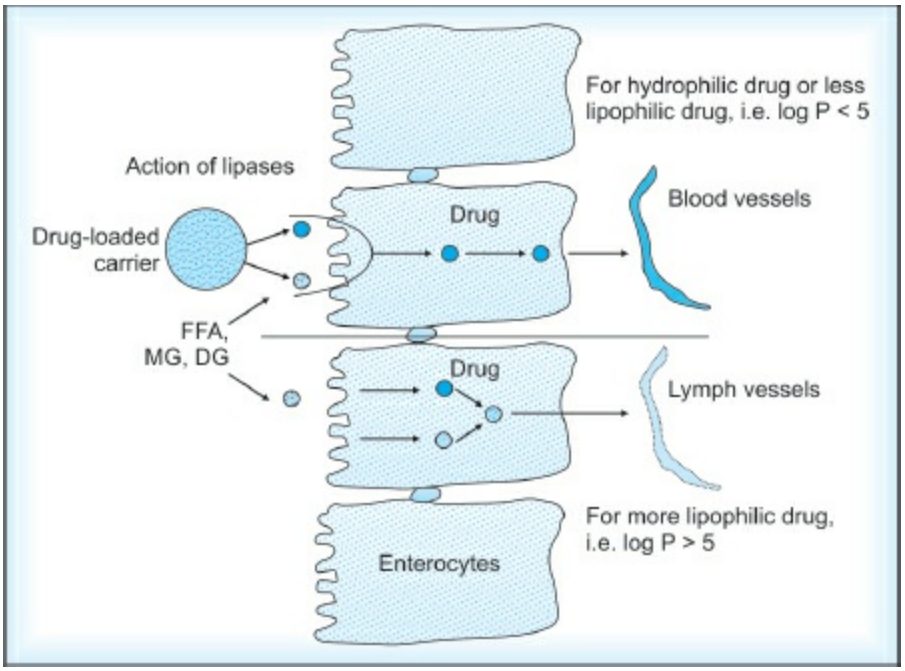



**Fig. 25.17:** Passive targeting of drug molecules to and through intestinal lymphatics after oral administration of drug-loaded carrier system

## Colon Targeting

Targeting drugs to the colon is therapeutically desired and beneficial strategy in a variety of disorders, and the colon has proven to be a potential site for local as well as systemic administration of drugs. Colon targeting is of clinical potential for local action in a variety of diseased conditions, such as inflammatory bowel disease, irritable bowel syndrome, and colonic cancer. Aminosalicylates, corticosteroids, immunosuppressive agents, cationized antioxidant enzymes, genetically engineered bacteria to produce cytokines, probiotics, nicotine, and other drugs have sucessfully tested and exhibited significantly enhanced efficacy when delivered in splendid isolation to the colon. Targeting drugs to the cancer cells using receptors and ligands have opened up new possibilities avenues in the treatment of colonic cancer. Colon targeting is apparently useful for systemic action of protein-peptide drugs such as insulin, calcitonin, and met-enkaphalin and even for other nonpeptide drugs such as cardiovascular and antiasthmatic agents (Fig. 25.18) describes approaches used for colon targeting.

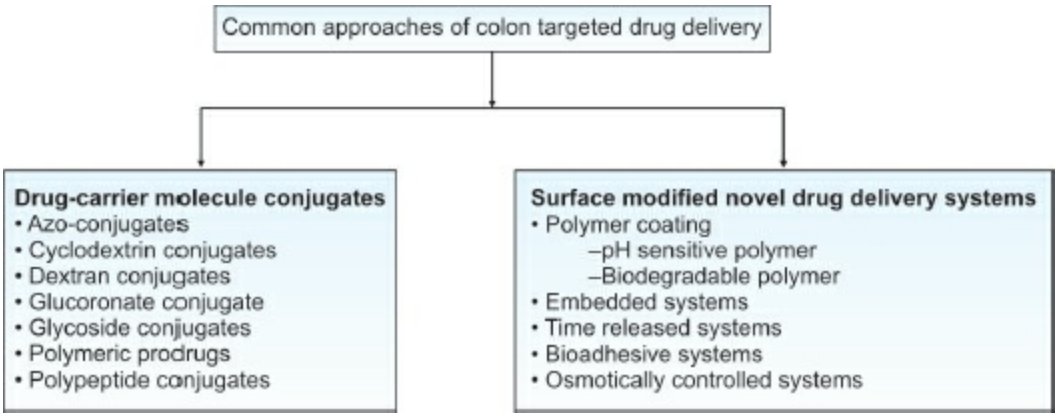



**Fig. 25.18:** Novel methodologies used for colon-targeted drug delivery

Drugs, which are destroyed by the acid mentle of stomach and metabolized by pancreatic enzymes, are minimally effective in the colon. Sustained colonic release of drugs can be useful in the treatment of certain diseases (Table 25.4). The colonic delivery is also beneficial in the systemic absorption of drugs like nifedipine, theophylline, isosorbide, etc. Further, colon was found to be a promising site for systemic absorption of peptides and proteins because of the less hydrolytic hostile environment in comparison with stomach and small intestine as well as the existence of specific

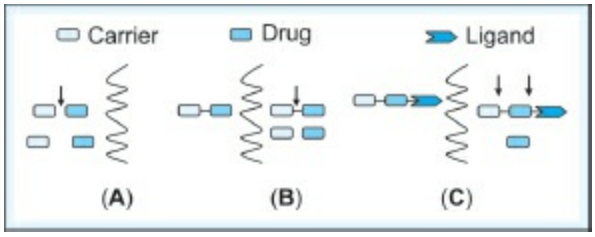
transporters. Additionally, the colon is a highly responsive site for the absorption of poorly absorbable drugs. Furthermore, the treatment of colon diseases such as ulcerative colitis, colorectal cancer and Crohn's disease is more effective with direct delivery of drugs to the affected area. Like-wise, colon delivery of diagnostic agents requires relatively low smaller doses. The successful delivery of drugs to the colon via the gastrointestinal (GI) tract requires the protection of a drug from being released in stomach and small intestine. This might be achieved by the use of special drug delivery system that can protect the drug during its passage to the colon. On the other hand, the drug must be released in the colon from the carrier. Targeting relies on exploiting a unique feature of the intended site and protecting the active agent until it reaches that site. Several approaches have been developed for targeted colon drug delivery. Most of them utilize the following four main properties of the GI tract and colon (1) approximation of transit time of the small intestine, (2) different physiological conditions in different branches of GI tract, (3) specificity of bacterial enzyme localized in the colon and (4) targeting of carrier to colon utilizing targeting moieties specific to colon.

**Table 25.4:** List of drug moieties used for colonic delivery for local and systemic effects

| Criteria for colon targeting | Category | Therapeutic molecule | |
|---|---|---|---|
| | | Non-peptide drugs | Peptide drugs |
| Drugs used for local effects in colon against GIT diseases | Anti-inflammatory drugs | • 5-Amino-salicylic acid<br>• 5-Flourouracil | • Amylin |
| Drug poorly absorbed from upper GIT | Antihypertensive and antianginal drugs | • Bleomycin<br>• Bromophenaramine | • Antisense oligonucleotide |
| Drugs for colon cancer | Antineoplastic drugs | • Dexamethasone | • Calcitonin |
| Drug degrading in stomach and small intestine | Peptides and proteins | • Diclofenac sodium<br>• Doxorubicin | • Cyclosporin |
| Drugs undergoing extensive first-pass metabolism | Nitroglycerin and corticosteroids | • Hydrocortisone<br>• Ibuprofen | • Desmopressin |
| Drugs needing targeting | Antiarthritic and antiasthmatic drugs | • Isosorbides<br>• Metoprolol | • Epoetin |
| | | • Nicotine<br>• Nifedipine<br>• Nimustine<br>• Oxyprenolol<br>• Prednisolone<br>• Pseudoephedrine<br>• Theophylline | • Glucagon<br>• Gonadoreline<br>• Insulin<br>• Interferons<br>• Leuprolide<br>• Filgrastin<br>• Molgramoatim<br>• Protinelin<br>• Sermorelin<br>• Saloatonin<br>• Somatropin<br>• Urotoilitin<br>• Vasopressin |

The use of transit times to targeted drug delivery specifically to the colon is limited by significant variations of this parameter between different patients, dependence of gastric emptying time on pathological, nutritional and other physiological factors. Similarly, large variations in pH in the GI tract limit the effectiveness of colon targeted delivery systems based on the peculiarities of colon pH. Natural polysaccharides are now extensively being used for the development of solid dosage forms for delivery of drug to the colon. The rationale for the development of a polysaccharide based delivery system for colonic drug delivery is the presence of high levels of polysaccharidases of microbial origin in the human colon. A variety of colon bacteria are known to secrete many enzymes, e.g. beta-D-glucosidase, beta-D-galactosidase, amylase, pectinase, xylanase, beta-D-xylosidase, dextranase, etc. Polysaccharides may be utilized for colon-specific delivery in various forms: fermentable coating of the drug core, embedding of the drug in the biodegradable matrix, formulation of drug-saccharide conjugate (prodrugs).

A large number of polysaccharides have already been studied for their potential as colon-specific drug carrier systems, such as chitosan, pectin and its salts, chondroitin sulfate, cyclodextrin, dextrans, guar gum, inulin, amylose and locust bean gum. More complex polymer-based system utilizing colon cleavable spacers containing azo bond, p-aminobenzoic acid (PABA) and p-aminohippuric acid (PAH) have also been used for colon drug delivery. Delivery systems and targeting agents, which are being developed for the delivery of drugs, may also be exploited for the delivery of vaccines, since many of the delivery problems are common to both areas. Prodrug based colon targeted drug delivery is shown schematically in Fig. 25.19.




**Figs 25.19A to C:** Common mechanism of targeted drug delivery to colon through pro-drug approaches: (A) Drug-carrier or prodrug break down in colonic environment, i.e. either through azo-reductase enzyme or pH sensitivity; (B) Transport of the prodrug across the colonic epithelium followed by subsequent release of the drug; (C) Ligand directed delivery of prodrug through colonic epithelium into cells

## M Cell Targeted Delivery of Antigens

To deliver vaccines across mucosal surfaces the epithelial barrier must be overcome. In addition to restricted absorption at mucosal surfaces, fluid secretion may flush away applied delivery vehicles. It has, therefore, been proposed that mucosal vaccine delivery could be improved by the use of appropriate bioadhesins that bind to mucosal surfaces and by exploiting the normal antigen-sampling mechanisms within mucosa-associated lymphoid tissues (MALT). Some of them are shown in Fig. 25.20. Of the potential bio-adhesins available, the most studied to date are probably lectins, a structurally diverse group of proteins and glycoproteins which bind reversibly to specific the carbohydrate residues. Since cell surface carbohydrate expression is a widespread phenomenon, it may be possible to exploit the bioadhesive properties of lectins to enhance delivery at a variety of sites. In addition, since cell surface carbohydrate expression exhibits regional and cell type specificity, lectins may also permit targeting to selected areas within a single mucosal tract.

The organised mucosa-associated lymphoid tissues (O-MALT) are antigen-sampling and inductive sites of the mucosal immune system. At these sites, antigens are transported across the mucosal epithelial barrier and prime up underlying lymphocytes for a subsequent immunological response. Specialised cells termed microfold membranous epithelial (M) cells are responsible for this transepithelial antigen transport, and putative M cells have been identified at multiple mucosal sites. Intestinal M cells are located in the follicle-associated epithelium (FAE) overlying the isolated and aggregated lymphoid follicles of the small and large intestines.

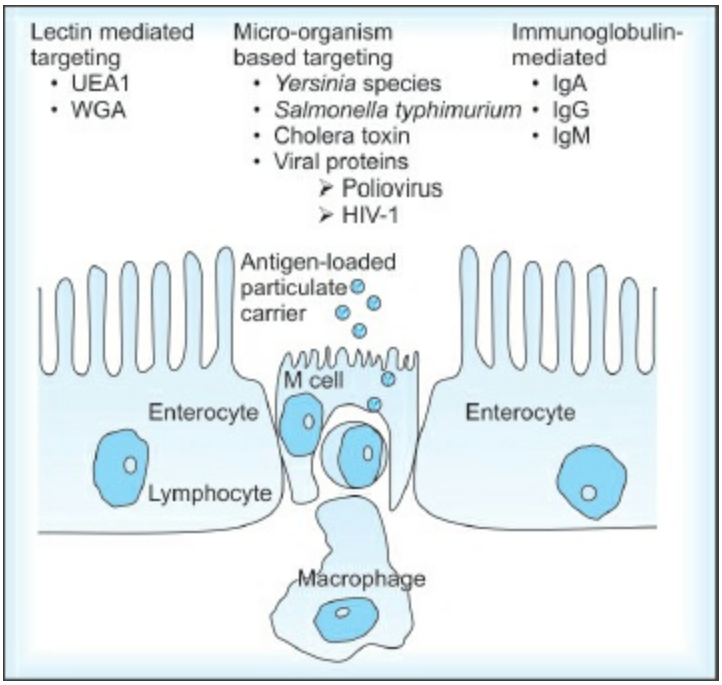



**Fig. 25.20:** List of ligand used for M cell targeted delivery of vaccine and their subsequent presentation to macrophages

M cells are characterized by the presence of irregular cell surface microvilli and by a basolateral cytoplasmic invagination that creates a pocket containing lymphocytes and occasional macrophages. Consistent with their antigen-sampling role, experimental studies suggest that M cells internalise and transcytose a wide variety of soluble tracers and inert particulates. M cells also represent a weak point in the intestinal epithelial barrier which is exploited by a wide variety of pathogens as a route of host invasion. The interaction of microorganisms with M cells is thought to involve both non-specific adherence mechanisms and interaction with specific receptors either of these mechanisms might be exploited for vaccine delivery through this cell type. Local factors associated with the lymphoid tissue environment are believed to control Peyer's patch and M cell development.

It has been proposed that, under the influence of lymphocytes, M cells develop from undifferentiated dome-associated crypt cells or, alternatively, fromfully differentiated enterocytes within the FAE. An algorithm of experimental evidences, including recent detailed microscopical analysis, supports the former hypothesis. However, this does not exclude the possibility that, under appropriate stimulation, a population of enterocytes within the FAE may retain the ability to convert into M cells. Consistent with

this proposal, some reports suggest that antigenic stimulation may induce FAE enterocytes to rapidly convert into M cells, thereby increasing the size of the M cell population. For example, Salmonella typhimurium caused an increase in the number of alkaline phosphatase negative cells (presumed M cells) in mouse Peyer's patch FAE 12 h after infection. Such uptake of particulate carrier by M cell has been utilized for vaccine delivey agianst Hepatitis B, anthrax, tetanus toxoid and influenza.

## Intracellular Targeting

Recent advances in molecular and cellular biology have led to the development of new classes of therapeutic agents, which are required to be delivered at their cellular/subcellular targets. Most of the newer agents developed have their site of action in the cytosol or cellular organelle of the cells, e.g. glucocorticoids such as dexamethasone; enzymes for the lysosomal compartment and various anticancer proapoptotic drugs to the mitochondria. While bioproximity for pharmacodynamic action of bioactive substance could be arrived by various currently developed transport strategies, the plasma membrane of the cellular target provides a formidable obstacle for large and charged molecules and hence getting a drug across the plasma membrane into the cytosol is considered as one of the stern rate limiting steps, as the majority of cells are not phagocytic and fusion of carriers with target cells is a very remote and rare phenomenon.

To develop an effective and successful carrier to deliver the therapeutic agent at the cellular level, an understanding of differences in membrane function, properties, and structure among cellular organelles as well as the basic mechanism(s) by which cells internalize extracellular material is essential. Following endocytosis, the lysosome is encountered as a major obstacle to the delivery of proteins, peptides, and DNA to the cytosol owing to lysosomal membrane permeabilizing properties and the lysosomal degradation pathway. The delivery of bioactive molecules/macromolecules to the intracellular site can be achieved by various strategies, which are termed ‘cytosolic approaches’. These include both direct entry to the cell cytosol and entry by endosomal escape (Fig. 25.21). Cell-penetrating peptides follow the first pathway and various approaches such as vesicle membrane destabilization or buffering of the compartment have been used to avoid the degradation in the lysosomes. The carriers that have been used successfully are pH-sensitive liposomes cationic liposomes, cationic lipids, polymeric nanoparticles, etc. For more efficient and specific delivery, various targeting moieties can be attached to the surface of the delivery system, e.g. folate and transferring for tumor cells, polyschacharides for hepatoma cells. Various endogeneous ligands and their respective receptors have been discussed elsewhere.

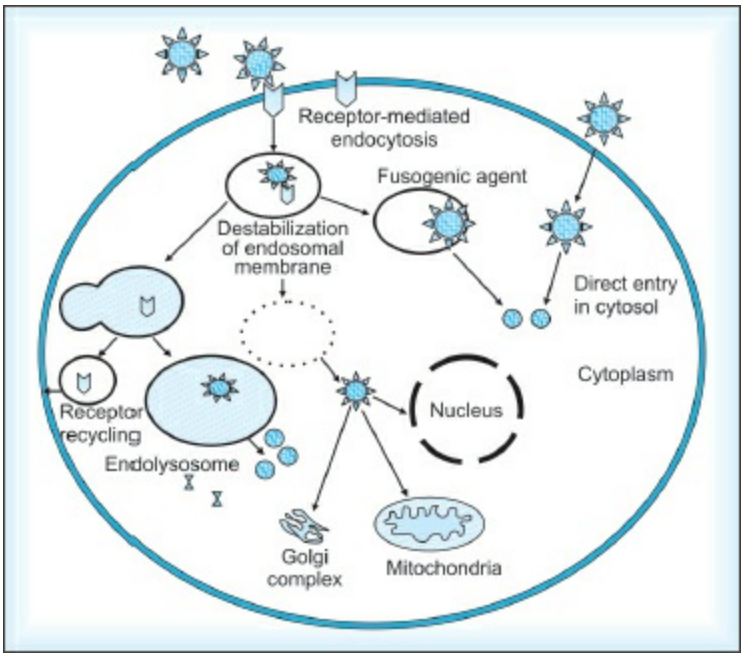



**Fig. 25.21:** Mechanistic approaches used for intracellular drug delivery and organelle-specific targeting

## Cytosol Specific Drug Targeting

A cellular targeted drug carrier should deliver the drug not just to the target cells but also to the particular cellular compartments or microenvironments where it is most required. Degradation of ligand coupled drug-carrier complexes in lysosomes along with the encapsulant, as well as the inability of the encapsulant to cross the lysosomal lipidic membrane, which constitute a 'lysosomotropic' approach to drug targeting, are the major barriers to achieving effective cellular delivery. These barriers must be overcome before the targeted delivery of cellular biomolecular therapeutics to specific cells and tissues becomes an accepatble clinical reality. An ideal drug carrier for cellular targeting should (a) bypass the anatomical barriers, (b) be recognized only by the target cells and (c) release the drug at or inside the target cells and not elsewhere. Targeting to the appropriate cells is not enough for therapy involving drugs acting intracellularly. The initial extracellular recognition step is usually followed by internalization via receptor-mediated endocytosis. The efficacy of several important proteins and DNA therapeutics is subsequently limited by nonproductive intracellular trafficking. For example, the release of immunotoxins and gene therapeutics from endosomes can represent the dose-limiting step in getting the drug to the site of action. The current challenge is to manipulate or circumvent the dominant non-productive trafficking pathways, such as routing of drug(s), proteins and DNA to the lysosomes where they are degraded. The formulation attributes for the successful design of target oriented carriers are as follows:

1. Cell specific specialized ligand/recognition moieties.
2. Cellular compartment specific module/strategy to achieve further specificity, e.g. cytosol, cell organelles.
3. Should not cause toxic or immune reactions, and
4. The drug-carrier complex should be capable of being manufactured under sterile and apyrogenic conditions.

There are various strategies which can be exploited and utilized to deliver the drug through carrier complexes directly to the cytoplasm scavenging the late endosomal and lysosomal compartments. These strategies can be subdivided into two broad classes: (a) Direct entry to cytosol and (b) Entry by escaping the endosomal compartment.

## *Direct Entry to Cytosol*

Various macromolecular drugs such as proteins and peptides, DNA and a number of drugs which are unstable at endosomal/lysosomal pH, are required to bypass the endocytic pathway of internalization for efficient activity in the cytosol or other organelles. The lipidic nature of biological membranes (plasma membranes) restricts the entry of such drugs to the cell.

A novel approach for transferring molecules directly to the cytosol bypassing the endocytic pathway is based upon the use of peptides which possess cell penetrating property (CPPs). The most actively studied peptides are derived from HIV-TAT, HSV-VP22 and antennapedia (Antp). These peptides are structurally similar in that they all contain a short sequence of less than 20 amino acids with a positively charged arginine and lysine residues. This sequence is called the “protein transduction domain” (PTD) and is significant for contact with the cell membrane. The mechanism of internalization of PTD is not well understood, but it excludes the classical endocytic pathway and occurs efficiently at both 37°C and 4°C. Tat mediated intracellular delivery of protein and nanoparticles via energy-dependent macropinocytosis followed by escape from the endosome into the cytosol have also been reported. This has been used for the delivery of various large and small molecules of drugs. The uses of CPPs for the efficient intracellular delivery of various carrier systems like nanoparticles, liposomes and Quantum Dots have also been studied. Recently, it has been reported that plain and PEGylated liposomes of 200 nm size could be delivered to various cells by attaching multiple TAT molecules to the surface of the liposomes (TATp-liposome) via a p-nitro-phenylcarbonyl PEG-phosphatidylethanolamine spacer group and they have been found to be intact in the cytosol for 1 h after translocation.

However, there are some limitations to using PTDs as carriers since they require crosslinking. Some of them, such as those derived from HIV-1 TAT proteins require denaturation of the protein before delivery to increase the accessibility of the PTD domains. Recently a short synthetic amphipathic carrier Pep-1 has been developed which delivers protein and peptide intracellularly without requiring any crosslinking or denaturation. The Pep-1 peptide carrier has been shown to be extremely efficient in the targeting of proteins into cells independently of endocytosis.

## Entry by Escaping Endosomal Compartment

### *Fusogenic Peptides or Proteins*

Following uptake of a bioactive loaded carrier system by receptor-mediated endocytosis, it reaches the endosome. The next step, of releasing the bioactive substance from the endosome before degradation, is the rate limiting step. When a virus reaches the endosome by cell adsorption it releases its genome in to the cytosol by activation of capsid protein at the acidic pH of the endosome. This active protein fuses with the endosomal membrane resulting in membrane rupture or prominent pores formation. A number of viral fusion peptides have been identified, e.g. C503-I609/ADM-1 proteins of C. elegans, G1-G29/HA-2 influenza virus, D149-D166/E2 glycoprotein of rubella virus, M1-Q16/Sprotein of Hepatitis B virus, G524-E540/Ebola virus, INF7/influenza virus. Fusion peptides have been identified, e.g. C503-I609/ADM-1 proteins of C. elegans, G1-G29/HA-2 influenza virus, D149-D166/E2 glycoprotein of rubella virus, M1-Q16/Sprotein of Hepatitis B virus, G524-E540/Ebola virus, INF7/influenza virus. Hemagglutinin was embedded in a phospholipid cholesterol bilayer. Sendai virus may directly fuse with preformed liposomes to yield virosomes with improved in vitro intracellular delivery. Sendai virus F protein is known to effect intracellular delivery through two independent mechanisms: (i) Galactosylated F protein is a ligand for the cell surface asialoglycoprotein receptor and (ii) F protein as a membrane fusogen. The fusogenic activity of F protein containing liposomes was found to be abolished by brief heat treatment without affecting the galactose mediated endocytic pathway.

## Targeted Delivery to Nucleus

Macromolecules and supramolecular complexes are frequently required to enter and exit the nucleus during normal cell function, but access is restricted and exchange to and from the nucleus is strictly controlled. Opportunities exist based on to make use of natural pathways for delivery of therapeutic entities, in particular to develop safe and effective methods for gene therapy, although past attempts to design non-viral nuclear delivery systems have met with limited success. To increase the likelihood of success scientists need an elucidation and appreciation of the mechanisms by which viruses deliver their genomes to the nucleus, and need a commitment to control the architecture of non-viral delivery systems at the molecular level. Effective delivery systems require several attributes in order to facilitate endosomal escape, microtubular transport and uptake through the nuclear pore complex. The published literature provides a strong foundation for design of nuclear targeting systems. The challenge faced by delivery scientists is to assemble a system which is as effective as, for example, the adenovirus however which lacks its immunogenicity.

## Mitochondrial Targeting

Mitochondrion, the energy generating organelle of the cell, is a subject of interest for selective targeting of therapeutic molecules, both Drugs and DNA, since it is involved in numerous disorders like diabetes, neurodegenerative disease, cancer, etc. Mitochondrial therapy can be made possible if the bioactive molecule is selectively delivered to the mitochondria of correct cell type, using cell specific ligands and mitochondriotropic molecules in designing of cell selective mitochondria specific carrier system.

Complete knowledge of pharmacological response of mitochondria for a drug molecule and applying this strategy with modification using pharmaceutical approaches for the treatment of disease associated with mitochondrial dysfunction, will torch in new direction and this knowledge can be used for cure and/or minimize the disease state by targeting drugs (acting on mitochondria) to mitochondrial sites or *via* gene therapy of mitochondrial DNA (Table 25.5). Since desired effect of any targeted drug or gene delivery can be achieved only if bioactive molecule is delivered to the destined organ and/or cell type, and also targeted to the correct location within the cell. To achieve this specificity of target, more efficient and selective delivery vehicles are designed Engineering of such mitochondria-selective homing devices is now the subject of current interest for controlled delivery of bioactives to mitochondria.

**Table 25.5:** List of physiological conditions/events of mitochondria responsible for diseases

| **Physiological event of mitochondria** | **Developing disease** |
|---|---|
| Free radical generation | Degenerative disease |
| Calcium homeostasis disturbance | Cell death |
| Release of cellsignaling pathways | Generation of programmed cell death |
| Mutation of mitochondrial DNA | Progressive neurodegenerative disease like aging |
| Dysfunction of uncoupling proteins | Thermogesis and fat storage |

Mitochondrial therapy may be successful by constructing such vehicles and/or carrier systems, which deliver the drug/DNA to the mitochondria of specific cell that is in pathological state. This can be achieved only when cell-specific homing ligand are attached on to the drug carrier in order to make it cell-specific. Successful mitochondrial targeting strategies include three important steps as follows:

1. Delivery of cell-selective and mitochondria-specific carrier system to the target cell surface.
2. Intracellular trafficking and cytoplasmic delivery of mitochondriotropic delivery system.
3. Targeted delivery of bioactive(s) to mitochondria via mitochondriotropics.
   a. Mitochondriotropic-drug conjugates (mitoconjugates).
   b. DQAsomes (mitochondria selective vesicular systems).

## Future Prospects

Targeted drug delivery is the major focus of current research. After the concept of magic bullet, only a few targeted formulations could reach the market. The discovery in the area of molecular biology, biotechnology and phar-macogenomics regularly demand the practical key issues of targeting of biomolecules to the therapeutic site and hence remain to be the center of attention. Tumour targeted drug/gene delivery is the most demanded therapeutic requirement of the coming future.

## Section IV:

# Product Processing, Packaging Evaluation and Regulation

# 26:
# Pilot Plant Scale-up and Production Management

# PILOT PLANT SCALE-UP TECHNIQUE

Research and development personnel expend a considerable amount of effort in developing drug dosage forms with exact specifications, so as to guarantee adequate physical and chemical stability. These products, designed to deliver and release a drug according to specific criteria, have, up to this stage, been manufactured on a laboratory scale, or in an intermediate-sized pilot plant equipment. Such equipment is usually fairly standard and available in most laboratories. In addition to the obvious requirements of clinical efficacy and safety, the ability of the experimental formulation to be reproducibly manufactured on high-speed production equipment in a cost-effective manner, often becomes a differentiating factor between a successful product and one that is simply a research curiosity. Further, the product must also be capable of being processed and packaged on a large-scale, often with equipment that remotely resembles that used in the development laboratory. In the pilot plant, a formula is transformed into a viable, robust product by the development of a reliable and practical method of manufacture that allows for the orderly transition of the product from the laboratory to routine processing in a full-scale production facility.

The pilot plant is a ‘hybrid’ development facility and manufacturing unit, which integrates development, which integrates early development activities, clinical supply manufacture, technology evaluation, scale-up, and transfer to production sites. Pilot plant studies must include a current good manufacturing practices (cGMPs) environment, a flexible highly trained staff, a close examination of the formula to determine its ability to withstand batch-scale and process modification, equipment to support multiple dosage form development and equipment at multiple scales based on similar operating principles to those in production. During this process, the availability of raw materials that consistently comply with the specifications required to produce the product must also be determined. Production rates and their relationship to immediate and future market requirements must be considered. The physical space required and the layout of related functions should be taken into consderation during the pilot plant phase itself with the intent to provide short-term and long-term efficiencies. The requirements, training, reporting relationships, and responsibilities of personnel are also factors, which must

be determined for a successful product scale-up. During the scale-up efforts in the pilot plant, production and process controls are evaluated, validated, and finalized. In addition, appropriate records and reports are issued to support Good Manufacturing Practices (GMPs) and to provide information about the historical development of the production formulation, process, equipment train, and specifications.

Often, meaningful product reprocessing procedures can only be developed and validated at pilot plant scale. All the critical features of the process must be identified so that, as the process is scaled-up, it can be adequately monitored to provide assurance that the process is under control, and that the product produced at each level of the scale-up maintains the specified attributes originally intended.

The pilot plant plays a critical role in the technology evaluation, scale-up, and transfer activities of new products. These activities begin early in the development cycle and include technical aspects of process development and scale-up, organization and determination of responsibilities of technology transfer teams, documentation of the transfer process, and often preparation for a FDA preapproval inspection. A properly designed and operated pilot plant can aid in enhanced collection of scientific data necessary for the support of internal transfer activities, as well as regulatory submissions. Key technical aspects that must be addressed during scale-up in the pilot plant includes:

1. Identification and control of critical component and formulation variables early in development.
2. Pilot plant equipment simulates as closely as possible to the equipment that will be used at the manufacturing site.
3. Identification of critical process parameters and operating ranges for the pilot plant equipment through the use of statistically designed experiments.
4. Collection of product and process data to adequately characterize each unit operation.

The facility design and the pilot plant staff involved in manufacturing operations plays a key role in ensuring the smooth and timely transfer of process technology to the manufacturing site.

In the past, the transfer of formulation and manufacturing technology was

sometimes done discretely from the development staff to manufacturing staff with little interaction or foresight. Nowadays, however, it is commonly recognized that the interaction of these groups at an early development stage is critical in obtaining an efficient and successful transfer. A team oriented approach to the manufacture of pilot or large-scale batches in the pilot plant will allow key production site personnel to view and comment on the process and make specific recommendations for improvement based on knowledge of the manufacturing site. An example of the team approach to the transfer of manufacturing technology, key development, production milestones, and the activities related to these milestones, is presented in Fig. 26.1. Interactions between the pilot plant staff that begin during product development stage and carry through production start-up ensure a smooth transition of the product from the lab to the marketplace.

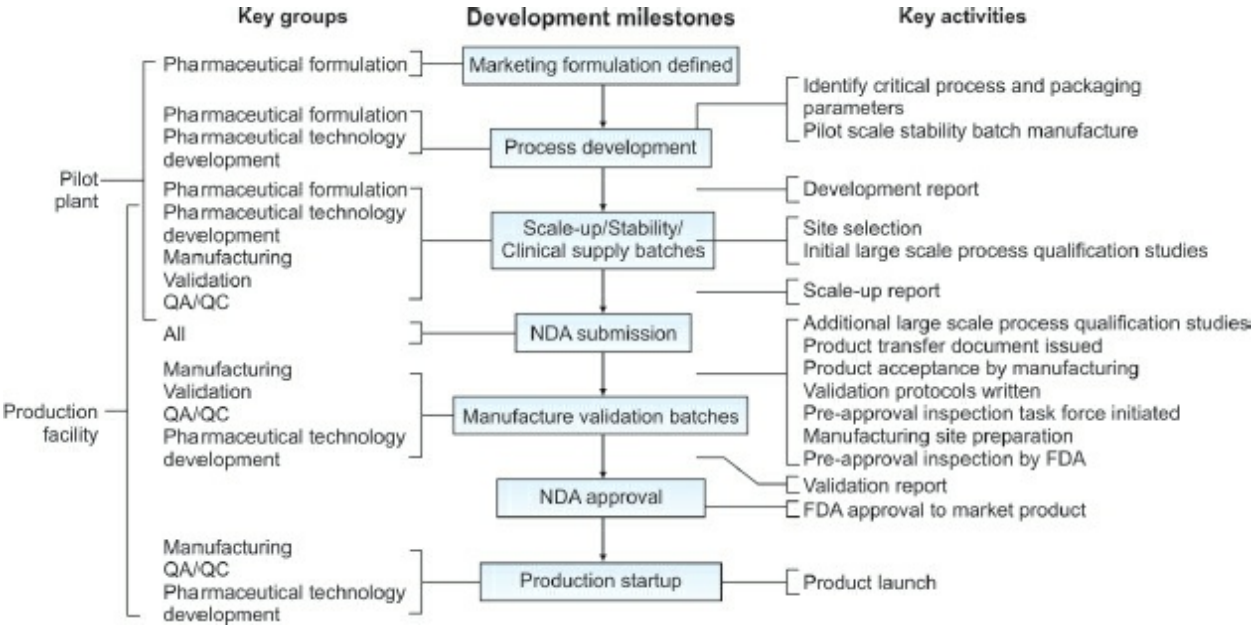



**Fig. 26.1:** The relationship of key groups, development milestones and activities, typically experienced during technology transfer from the pilot plant to the production facility.

# GENERAL CONSIDERATIONS

## Reporting Responsibilities

The question of who should be responsible for pilot plant studies has always been debated at great length with no clear resolution, as can be seen from the various reporting relationships in use within major pharmaceutical companies. Pilot plant functions can be part of a research and development group with separate staffing. This arrangement is designed to provide a hierarchy of responsibility to scale-up formulations that have been developed by other formulators within research and development, thereby providing an opportunity for critique of formula/process that is independent of the initial formulation function. Alternatively, the formulators who developed the product can take it into production and continue to provide support even after the transition into production has been completed. Proponents of this system cite the advantage of historical continuity that comes from this approach. Both of these structures arise from the premise, that if the pilot plant is a research responsibility, greater consideration should be given to the preformulation and formulation experiences obtained during initial development of the particular drug and its dosage form.

Some companies prefer that the pilot plant and the technical service group are organizationally separate from research and reporting instead to the operations side of the business. The advantage here may be that such a group would be more operations-oriented, more pragmatic, and more receptive to the operations priorities. Management philosophy, the nature of a company's products, and the background experience of the personnel involved in pilot plant studies help determine which arrangement is best for any one particular organization at any point in time. A few companies have adopted a composite of both approaches in hope of achieving the best attributes of both the research and development and the operations-oriented systems.

Whatever the reporting arrangement be, the goal of a pilot plant is to facilitate the transfer of a product from the laboratory into production. The effectiveness of the pilot plant is determined by the ease with which new products or processes are brought into routine production. This can best be achieved if a good relationship exists between the pilot plant group and the other groups with which they interact, namely, research and development, processing, packaging, engineering, quality assurance/control (QA/QC), 

regulatory, and marketing.

## Personnel Requirements

The qualifications required for a position in a pilot plant organization include a blend of good theoretic knowledge of pharmaceutics and some practical experience in the pharmaceutical industry. In addition, the ability to communicate well, both in speaking and in writing, and the ability to develop good relationships with other people are important since experience and knowledge are most useful when adequately and effectively communicated. Practical experience in pilot plant operations is invaluable. Experience within the group should encompass both the formulation experience and the process and equipment experience in the actual production environment. Personnel in the pilot plant must recognize the intent of the formulator, and at the same time, understand the perspective of production personnel. For these reasons, successful pilot plant organizations frequently include scientists with experience in both areas.

The type and level of education within the group is important. Pharmaceutically trained scientists contribute fundamental strength to the function in their ability to assimilate the complex interrelationship between pharmaceutical processes and the potential impact on chemical, physical, biochemical, and medical attributes of dosage forms. It is also important, however, that the group possess some engineering capability since the scale-up of many of the processes involves engineering principles, which are usually not well-covered in the normal pharmaceutical training. In addition, it is becoming increasingly important for the group to contain individuals who are knowledgeable in both electronics and computers.

The number of people in a pilot plant group depends on the number of products being supported and on the level of support required. An experienced scientist with a knowledgeable technician should be able to handle one major project, or two major projects simultaneously depending on their complexity, while at the same time providing technical support for an additional group of marketed products.

Many established companies have production operators who have had many years of experience with particular dosage forms and equipment, and supervisors who have had considerable research experience. In such cases, the support effort required of a pilot plant group would be much reduced. In such a company, the introduction of new products involving existing

technologies would require less support than the introduction of a process involving new technology. On the other hand, companies where production facilities are operated without the support of extensively trained technical personnel, and in which the supervisory function is mainly administrative, can be expected to require considerably more support during introduction and scale-up of new products or processes and also to require substantially more continuing technical support.

The establishment and maintenance of a satisfactory system of quality assurance and the correct manufacture of medicinal products relies upon people. For this reason there must be sufficient qualified personnel to carry out all the tasks which are the responsibility of the manufacturer. Individual responsibilities should be clearly understood by the individuals and recorded. All personnel should be aware of the principles of the Good Manufacturing Practice that affect them and receive initial and continuing training, including hygiene instructions, relevant to their needs.

Importance of a sufficient number of qualified, trained personnel is clearly reflected in the statement "*the quality of a product ultimately depends on the quality of those producing it*" (Sir Derek Dunlop, 1971).

References to the 4 P's (Personnel, Premises, Plant and Procedures) as the essential elements in any quality-orientated industrial enterprise are commonplace, and such generalizations serve to focus attention on the basic requirements. There can be no doubt that, of these elements, it is the people (the "men"—the human species, not the gender—or the personnel) that are the most important factor in the assurance of quality. This is true of all levels within an organization, from company president and managing director to the most-junior employee. It may well be possible (if not altogether desirable) for high-quality, well-trained, dedicated personnel to compensate for a lack or deficiency in the other elements. Nothing, not even the finest premises, equipment, materials, or procedures can compensate for the quality hazard represented by low-standard, ill-trained, or poorly motivated staff.

The head of quality control unit and the head of production are identified as the key personnel in GMP guidelines to ensure that manufacturing facility is functioning in compliance with the cGMP requirements. Their responsibilities may vary as per the requirements of local regulatory authorities and organization structure. The heads of production and quality control must be independent from each other. There should be no gaps or

unexplained overlaps in the responsibilities of those personnel concerned with the application of GMP.

In most cases, the quality of production in one department has an important effect on the operations of another, and employees should be trained to understand such interrelationships. For example, if the tablet department is producing a product at the low end of the tablet hardness range, in turn causing an excessive tablet breakage during packaging, then one department is transferring problems to another department, thereby making the problem more difficult and expensive to correct.

## Training

It would be difficult to exaggerate the importance of personnel training. It has been said that “People who have been trained, by example that high production is more important than extreme care, are going to make mistakes.” (Edmond Fry, former senior official, US FDA).

Since GMP affects all areas of manufacturing, participation by engineering, quality control, materials handling, warehousing and distribution, purchasing, and other departments is a necessary step in ensuring compliance. Each person engaged in the manufacture, processing, packaging and holding of a drug product must receive GMP training (at least as applicable to their responsibilities) and specific training depending upon the complexity of their jobs. New employees should also receive introductory, background (induction or orientation) training. Employees who require training can be categorized into three main categories:

1. New employees
2. Those employees who are assigned a new job (nature or content of their job changes)
3. Those employees whose performance at a task falls below required standards.

The training of hourly personnel in the manufacturing area is particularly important, since new employees may find themselves in a relatively technical environment, dealing with potent or dangerous chemicals, and working with a system of weights and measures that may be unfamiliar to them. GMP clearly states that all materials going into a batch must be properly identified and checked, and that during all steps in the manufacturing process, all containers, drying racks, and other equipment should clearly state the product name and lot number as well as the process stage. In the packaging operations, a well-conceived line purgation system should ensure that all materials from a previous batch have been removed and that the new package order is properly identified with respect to the correct labeling as well as the batch number. Shipping and distribution departments should ship only approved materials and should deal effectively with problems related to recording lot numbers on invoices, which is extremly critical in the development of a product recall system. A proper stock rotation system and

all other controls exercised at the producing location must also be provided and monitored at every warehouse and branch location. Warehousing personnel should understand the importance of any quarantine system for supplies and should never deliver unapproved materials to a processing department. In large operations, it may be difficult to place approval labels on every bag, drum, or container, and any substitute system must be clearly understood, be foolproof, and be acceptable to FDA. For these and other reasons, each employee must clearly understand why accuracy, cleanliness of operations, and strict adherence to instructions are necessary.

Written operating procedures are required under the regulations, and they are considered helpful in the employee training program. A well-documented system and standardized procedures not only help in personnel training, but also provide a basis for the development of a self-inspection program that will permit auditing of operations by production management from time to time to ensure continuing compliance. A typical self-inspection program for a packaging operation can be found in the Eastern Regional Proceedings of PMA-FDA Seminar on “Quality Assurance in Drug Manufacturing.”

The development of a training program for the pilot plant staff is complicated by the inherent diversity between the operations and the personnel. Unlike the production facility, which is highly specialized and subject to infrequent operation changes, the pilot plant must be prepared to make a variety of dosage forms in response to a wide range of product development programs. The diversity of the pilot plant operations, equipment, and personnel requires a flexible entry, reinforcement, and remedial training program. An extensive tracking system and flexible scheduling system are also required. Most importantly, the training program must be developed to meet the changing priorities of the product development cycle while maintaining the cGMP compliance in the manufacture of clinical supplies.

## Review of the Formula

A thorough review of each aspect of the formulation is important and should be carried out early in the scale-up process. The purpose of each ingredient and its contribution to the final product manufactured on the small-scale laboratory equipment should be understood. As a result of this, the effects of scale-up using equipment that may subject the product to stresses of different types and degrees would be more readily predicted, or recognized when they actually occur. The need to modify the formulation during the scale-up is not unusual, and this should be done as early as possible in the phase III trials to allow sufficient time for the generation of meaningful long-term stability data in support of a proposed new drug application (NDA). If these studies are not completed until after the application is made, long and costly delays in the approval process may result.

## Raw Materials

One responsibility of the pilot plant function is the approval and validation of the active and excipient raw materials used in the pharmaceutical products. This is necessary because the raw materials used during the small-scale formulation trials may not be representative of the large-volume shipments of materials which would be used during the large-scale production, or also because the active ingredients, which may only have been prepared on a laboratory scale, are also being subjected to scale-up to meet the rising needs of the product. Even though all analytic specifications are met, these larger lots of active ingredient may change in particle size, shape, or morphology, resulting in different handling properties or differences in bulk density, static charges, rate of solubility, flow properties, color, etc. The quality of active ingredients needs to be verified because having alternate suppliers is usually desirable. This is an important consideration for companies who purchase their active ingredients, because a single supplier leaves the company vulnerable with respect to both supply and acquisition price. The evaluation of alternate suppliers requires that several batches of product be manufactured with these alternate materials and that their performance in the formulation and the stability of finished products be evaluated relative to the standard product.

## Relevant Processing Equipment

It is quite certain that most of the formulation development work is carried out on small, relatively simple laboratory equipment and during the subsequent scale-up, alternative manufacturing equipment should be considered. Based on the known processing characteristics of the product, the equipment that promises to be the most economical, the simplest, the most efficient, and the most capable of consistently producing a product within the proposed specifications should be evaluated. For feasibility studies, if a particular technology is not available in-house, small-scale trials can be carried out at the various equipment vendors' facilities. Then, when a decision has been made to use a particular process, the selected pilot plant equipment should be acquired. The size of this equipment should be such that experimental trials can be run that are meaningful and relevant to the production-sized batches that will eventually be made. If the pilot plant equipment is too small, the process developed will not scale up well. If the equipment is too large, excessive costs will be incurred, especially if the product involves the use of a large quantity of a new and expensive active ingredient.

When a reasonable process has been developed on the pilot plant equipment, intermediate-sized experimental batches should be run. Once again, these runs can be made in the facilities of the equipment vendor if the equipment is not available in-house, and thus provide some indication of the reliability of the extrapolation from the pilot plant to larger production equipment. When the decision involves equipment available from several vendors (e.g. fluid bed dryers), the next step is to determine from the various vendors the advantages of their particular equipment. Ease of cleaning should be considered, especially if multiple products are destined to be manufactured by the equipment. The time required to tear down the equipment for cleaning and to change from one product to another should also be determined. In some instances, this can be longer than the actual time required to manufacture a batch, and if frequent change-overs are anticipated, this becomes an important consideration. To evaluate these and other parameters accurately, trials should be run at the vendor's facilities. Such experiences help determine the true capability of the equipment and the quality of technical support available from the proposed vendors.

## Production Rates

The immediate and future market requirements must be considered when determining the production rates and the type and size of production equipment needed. The size of the equipment should be such that it is properly utilized. The equipment and process should be chosen so as to produce batches at a frequency that takes into consideration product loss in the equipment during manufacture, the time required to clean the equipment between batches, and the number of batches that will need to be tested before release. To accommodate future growth, increased production capacity may be realized more economically through more efficient utilization of smaller equipment than through the purchase of oversized equipment. For example, several smaller lots produced serially (without major clean-up) may be combined in a final blend to make a single large batch.

## Process Evaluation

The previous steps have developed the product scale-up program to the point that, the manufacturing process has been proposed and the equipment for production has been evaluated, selected, installed, and debugged. The next step is to evaluate the process critically and to optimize its performance based on that evaluation. Items that should be examined include the following:

- Order of addition of the components, including adjustment of their amounts
- Mixing speed
- Mixing-time
- Rate of addition of granulating agents, solvents, solutions of drugs, slurries, etc.
- Heating and cooling rates
- Filter sizes (liquids)
- Screen sizes (solids)
- Drying temperatures
- Drying time.

Knowledge of the effect of these important process parameters on the in-process and finished product quality is the basis for process optimization and validation. The purpose of process validation is to assure that the selected manufacturing procedure ensure the quality of the product at various critical stages of the process as well as in the finished form. This is accomplished by monitoring the within-batch variation of measurable parameters, such as content uniformity, moisture content, and compressibility. These data indicate where the process is performing as intended and where the problem areas may be found.

Parts of the process such as milling, mixing, heating, cooling, drying, sterilizing, compacting, and filling, which may cause some measurable change in the state of the material being processed, need to be evaluated. Such process data should be accumulated for a series of batches using a particular equipment configuration and a well-documented process. If the data show that the process performs consistently at the critical steps to produce a product that lie within the release specifications, then that process

has been validated. The process remains validated only if there are no changes in the formula, the quality of the ingredients, or the equipment configuration. Changes in any of these areas would have to be carefully evaluated and a determination would have to be made as to the need and extent of revalidation required. The personnel responsible for the process should be adequately trained so as to understand the directions and to carry out the process as intended.

The manufacturing process and quality control information should be reviewed on an annual basis, and if deemed necessary, some revalidation studies should be carried out to ensure that changes have not occurred. A validated process establishes a data base of cause and effect relationships between critical steps and in process or end product specifications. Therefore, documentation obtained during process validation can often be used predictively to shorten the time required to identify the factoids) in a process that has “drifted” from the normal.

## Preparation of Master Manufacturing Procedures

The manner in which the manufacturing directions, the chemical weigh sheet, the sampling directions, and the in-process and finished product specifications are presented is of utmost importance, as is the degree to which the processing technician understands and complies with them. The weight sheet should clearly identify the chemicals required in a batch. Further, the weigh sheet should present these in the quantities and the order in which they will be used. To prevent confusion and possible errors, both names and the identifying numbers for the ingredients should be used on batch records, and these should correspond with those on the bulk raw material containers.

The processing directions should be precise and explicit. They should be written in a style that uses language and terms with which the operators are familiar. While writing the manufacturing procedures, considerable input should come from the actual operators or from someone with current knowledge and expe-rience in the weighing and processing areas. In accord with GMPs, the batch records need to provide space to show the weighing and addition of each ingredient with appropriate countersignatures for each. A manufacturing procedure that becomes a multipage document and takes the place of good operator training and a library of standard operating procedures (SOPs) invites problems. Too detailed or wordy directions are cumbersome, and too many sign-offs are as bad as too few. There are numerous examples in the industry of problems arising because the operators will not read or follow highly detailed procedures. Experience has shown that when too many signatures are required in a batch record, they may not be signed off in the intended sequence during the manufacturing, but instead are completed all at one time at the completion of the process. When this occurs, the intention of verifing that each step was performed correctly is lost.

The batch record directions should include specifications for addition rates, mixing times, mixing speeds, heating and cooling rates, and temperature. It should also include the appropriate ranges for each of these parameters. The actual time, temperatures, and speed used should also be documented. These can best be monitored and recorded by the appropriate controller recorders, which free the operator to pay attention to the actual process and verify that all equipments are functioning properly and that the process is performing in a normal and acceptable manner. Strip charts from

the controllers are used to verify the compliance with the batch directions.

The time and manner in which the inprocess and finished product samples are to be taken from a batch, and the way in which they are handled and stored, should be clearly specified within the batch record. Samples that are improperly taken or handled give rise to unreliable data which can result in lost time or jeopardize the quality or acceptability of a batch. When this occurs, batch reassay, special stability studies, reprocessing, and extensive investigations may become necessary to assure the quality of the batch before its release.

In-process specifications included within the processing directions should be realistic but slightly narrower than the finished product specification. If a specification is too broad, it cannot provide the alert desired when something goes wrong with the process. Too tight a specification results in the needless rejection of batches of acceptable quality.

Finished product specifications set the standards by which a product is evaluated and help to ensure that each batch manufactured delivers the drug in the dose specified throughout the designated shelf-life of that product. Therefore, when finished product specifications and release specifications are set, they should take into consideration the capability of the process, the reliability of the test methods, and the stability kinetics of the product.

It is obvious but often forgotten that the quality of a batch cannot be determined by the analysis of the small number of samples usually required to meet compendial specifications (e.g. content uniformity). These results are only a valid estimate of the quality of a batch when process validation studies have been conducted to establish the statistical integrity of the sample. Periodic revalidation, good manufacturing procedures, and monitoring of the finished product test results via control charts are essential to maintaining consistent product quality.

## GMP Considerations

The term Good Manufacturing Practices is interpreted differently by different people. While there are FDA guidelines describing the GMP, these are also interpreted differently by different people in the FDA and in industry. Common sense exercised by people who have a good theoretic knowledge of pharmaceutical principles, and who are technically competent and have adequate relevant experience in manufacturing, guarantees compliance to the GMP.

A checklist of the GMP items that should be a part of the scale-up or new product or process introduction include the following:

- Equipment qualification
- Process validation
- Regularly scheduled preventative maintenance.
- Regular process review and revalidation
- Relevant written standard operating procedures
- The use of competent, technically qualified personnel
- Adequate provision for training of personnel
- A well-defined technology transfer system
- Validated cleaning procedures
- An orderly arrangement of equipment so as to ease material flow and prevent crosscontamination.

## Transfer of Analytic Methods to Quality Assurance

During the scale-up of a new product, the analytic test methods developed in research must be transferred to the quality assurance department. Early in the transfer process, the quality assurance staff should review the process to make sure that the proper analytic instrumentation is available and that personnel are trained to perform the tests. Recovery studies on the product and on spiked placebo samples should be carried out by different operators over a period of several weeks to verify that the tests perform reliably and yield results with precision and accuracy comparable to that obtained in research. If required, the assay method can be reformatted and rewritten using terminology and procedures consistent with current quality assurance laboratory practice. To complete the transfer process, the research personnel should review the assay procedure and the data obtained during the validation studies, to verify that the analytic methods have not been altered in a way which might affect the reliability, precision, or accuracy of the tests.

# DESIGN AND FACILITIES

## Pilot Plant Layout

In designing the facility and its layout, the impact on GMPs must be kept in mind. The layout should address the need for flexibility, restricted access, personnel flow, and material flow. The flexibility can be achieved in a number of ways. For example, fixed process equipment may be kept to a minimum and portable equipment be installed whenever possible. The movement of portable equipment is made easier by the construction of wide corridors and doors with height of 10 ft (3 m) or above. The process rooms can be designed for multipurpose use and provide desired flexibility for utilities, such as the power supply, dust collection, vacuum cleaning, water, static grounding, and compressed air.

The size of the restroom, lockers, gowning, and shower areas should be based on the anticipated nominal daily personnel level. The types of the garments worn by personnel in the pilot plant and gowning and degowning procedures should be specified and provisions be made for their cleaning, disposal, and storage. Material receipt, sampling, and quarantine processes should be well planned. The material and components should enter the facility through a controlled access area. Special precautions for the sampling of potent (highly active) compounds should be considered. This may require the use of specially designed rooms that have air locks, gowning areas, and isolation sampling methods. A receiving office should be located immediately adjacent to the actual receiving area. A restroom, telephone, and seating area for the transportation personnel are also other design options.

Receipt, sampling, and quarantine storage of flammable solvents require special consideration such as location, explosion-proof lighting, and electrical connections. Space is also required for the temporary holding of the waste materials. Temperature and relative humidity conditions to be maintained during storage should be specified and controlled through the use of air conditioners, refrigerators and freezers. The warehouse layout depends on the manufacturing scale, type of products, and the storage requirements. Storage within the facility is best, but off-site storage for medium-to long-term supplies or large quantity purchases may be a more cost-effective alternative. The method of inventory identification determines the design and capacity of the receiving, shipping, and warehousing areas. Figure 26.2 shows a pilot

plant layout design, which supports the development and clinical manufacture of solid dosage forms, liquids, semisolids, aerosols, and sterile products.

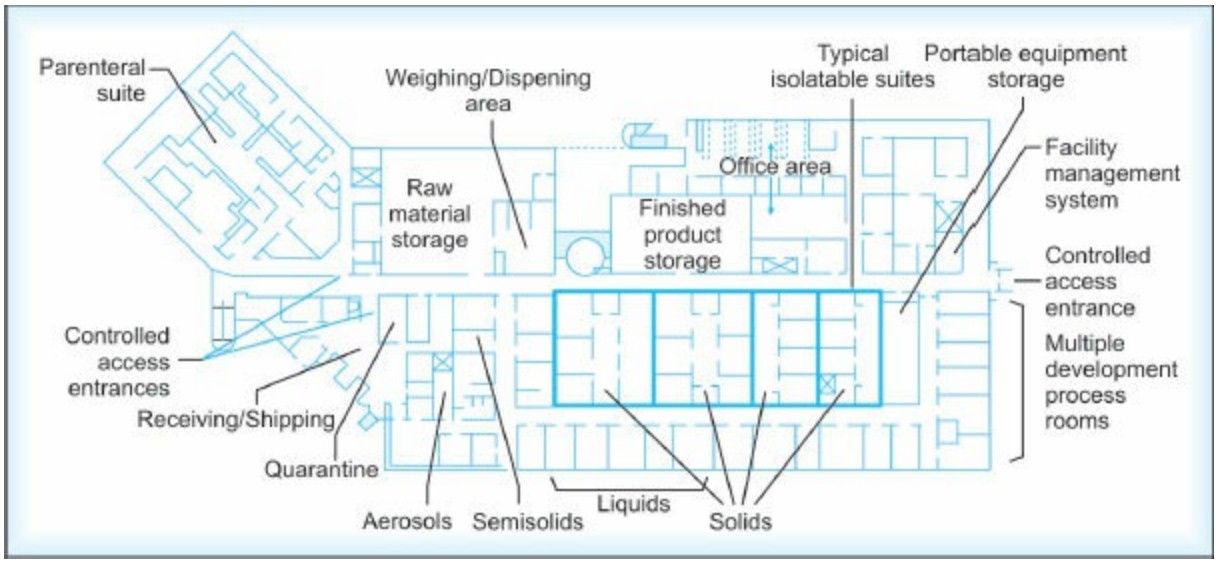



**Fig. 26.2:** A pilot plant layout design

## Environmental Factors

### *Materials, Lighting, and Air-conditioning Specifications*

Adequate cleaning, construction, lighting and ventilating are (among others) important factors in a production operation. Since the word “adequate” is used throughout the GMP, one cannot be specific with respect to the amount of lighting in each department, the type of ventilation, or the materials of construction for the floors and the walls.

Based on personal preferences and on the experiences of other companies, a choice can be made from an enormous array of available materials. These decisions are important when a new building is being designed, but become extremely difficult in the redesigning of an old facility, in which ceiling heights, wooden floors, and other outmoded features must be taken into consideration.

One of the most significant steps taken in this regard to provide industry-wide information on these areas, was a survey conducted by the Production and Engineering Section of the Pharmaceutical Manufacturing Association (PMA) in 1967 which was again updated in 1973. This publication was entitled as “Survey of Current Manufacturing Practices in the Drug Industry”, and it is divided into three chapters—Personnel, Buildings, and Equipment. In response to questionnaires sent out to both large and small members of PMA, a detailed summary of information has been compiled that represents a cross section of practices and materials utilized in the industry.

Table 26.1 provides a departmental checklist of a selection of materials that have proven durable over the years for walls, floors, and ceilings. Levels of lighting and general airconditioning requirements are also listed. In the selection of the construction materials, initial costs are only a partial consideration, since the maintenance cost enters into the total cost.

**Table 26.1:** Facilities guide for GMP

| | Chemical Weighing | Granulating | Tabletting | Coating | Liquid Manufacturing | Packaging | Warehousing |
|---|---|---|---|---|---|---|---|
| **Floors** | | | | | | | |
| High-density concrete | X | X | X | X | X | X | X |
| Troweled epoxy finish | X | X | X | | X | | |
| Terrazzo epoxy finish | X | X | X | | X | | |
| Concrete w/metal chips | | | | | | | X |
| Drains | X | X | X | X | X | X | X |
| Vinyl asbestos tile | X | | X | X | | X | |
| **Walls** | | | | | | | |
| Smooth | X | X | X | X | X | X | X |
| Epoxy paint | X | X | X | X | X | X | X |
| Washable | X | X | X | X | X | X | X |
| Glazed tile | X | X | X | X | X | X | |
| Wallpaper (washable) | | | X | | | X | |
| **Ceilings** | | | | | | | |
| Hung | X | X | X | X | X | X | |
| Acoustical | X | | X | X | X | X | |
| Smooth waterproof | | X | | | | | |
| Open | | | | | | | X |
| **Lighting** | | | | | | | |
| Footcandles | 75 | 100 | 100 | 75 | 75 | 75 | 50 |
| Flush mounted | X | X | X | X | X | X | |
| **Air-conditioning** | | | | | | | |
| Optional | | X | | | X | | |
| Comfort control | | X | | | X | X | |
| 45% RH/72°F | X | | X | X | | | |
| <20% RH/72°F | | | X | | | | |

The selection of decorating colors in the manufacturing and office areas is important, especially in plants with few or no windows. The sectional use of different colored paints in long corridors avoids a tunnel-like impression; and the hanging of washable wallpaper in areas such as the packaging room provides pleasant surrounding and low maintenance costs.

## *Dust Collection and Cross-contamination*

Cross-contamination constitutes an everpresent danger in pharmaceutical manufacturing. Identification and removal of its causes have had a profound effect on the plant layout as well as on the design and construction of the production areas. In many cases, a reassessment of the dust collection requirements and specifications has been necessary. For example, the development of increasingly sensitive test methods for penicillin, coupled with the awareness that some individuals are extremely sensitive to this antibiotic, necessitates a complete relocation of penicillin production unit into a separate facility away from all other operations to avoid possible cross-contamination.

## *Types of Dust-collecting Systems*

Although several types of dust collectors are available, the selection of equipment should be based primarily on its intended application. Cyclone collectors, for example, have limited application in pharmaceutical production except for their use in the spray-drying operation in which large volumes of powder are processed on a continuous basis. On the other hand, replaceable filters, cloth bags, wet scrubbers, and high-efficiency particulate air filters (HEPA), or a combination of some of these, are commonly used.

## *Wet Collectors*

A *wet scrubber* operates by mixing the dustladen air stream with water in an enclosed chamber, followed by discharging the effluent into a sewer or treatment plant. A Rotoclone operates by spraying the air in the collector with water and mixing it by means of highspeed paddles. This type of scrubber is efficient and uses a minimal amount of water. Both systems are particularly effective when a dye is used as a component of the dosage form, and it is therefore necessary to prevent trace quantities from contaminating the nearby operations. Either process can be used when water-soluble, volatile solvents are used in the pharmaceutical manufacturing, since the entrained vapors are diluted with water to a safe level of concentration.

## *Filters, Bags, and High-efficiency Particulate Air Filter*

The most commonly used dust collectors are the *replaceable filter*, *cloth bag*, and *HEPA*. In large operations, the collectors are placed on the roof or on a mezzanine. These air filters are rated on an efficiency basis as defined by the National Bureau of Standards (NBS) and the American Filter Institute (AFI), utilizing the Dioctyl-phthalate (DOP) test method. Figure 26.3 illustrates a DOP efficiency table for particles in the 0.3 micron range. Retention of Astrocel indicate HEPA efficiency of 99.97%, however, retention of Varicel 50 indicate 35% efficiency.

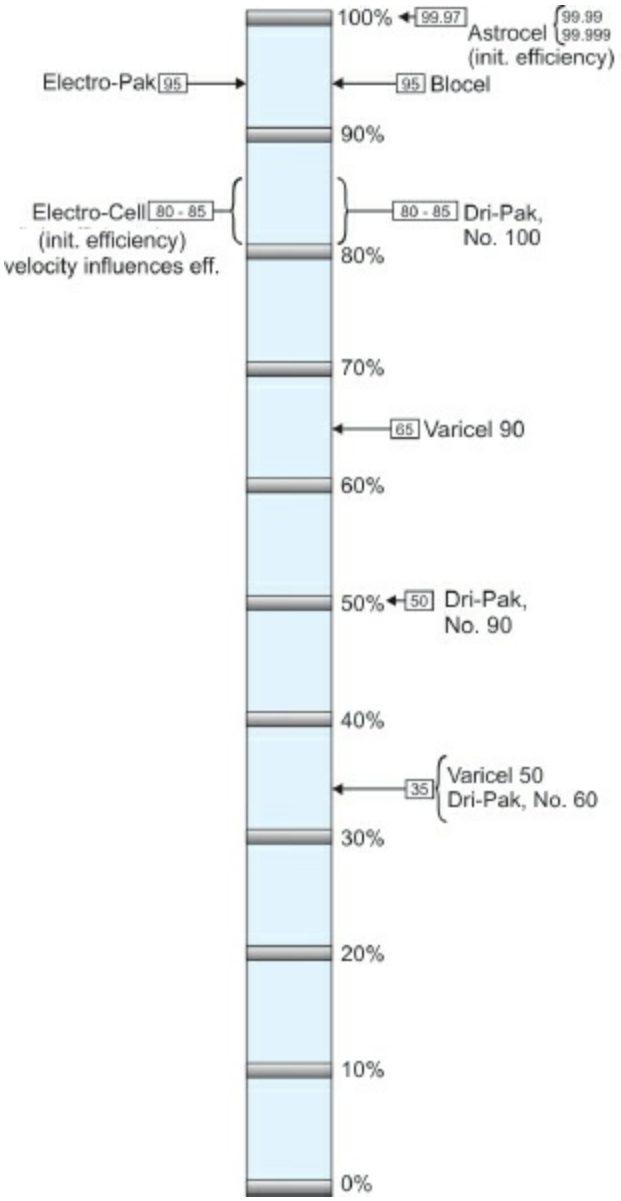



**Fig. 26.3:** DOP efficiency. Dioctyl-phthalate aerosol averaging 0.3 microns. Determinations made by lightscattering method. Average value over life of media unless otherwise noted

HEPA filters are 100% effective in particle size ranges of a few microns, and at one time, these were used exclusively in sterile operations to remove particulate matter. However, current emphasis on reducing cross-contamination, has resulted in their use in many dry-product operations, such as granulation, tableting, and capusle filling.

The handling of products containing attentuated viruses requires the use

of HEPA filters with no return air circulation as a minimum precaution and when viable pathogenic viruses or other dangerous organisms are processed, it is best to incinerate all exhaust air at a temperature above 800°F.

## *Departmental Specifications*

In addition to selecting the correct type of dust collector for each department, it is also desirable to provide the correct air volume and velocity required by the collector to do its job effectively. One of the most common failures in dust collection is an insufficient air velocity to carry all the particles from the pickup location to the collector, and the most sophisticated unit can become ineffective, especially if it is located some distance away from the pickup location. For example, the dust pickup stations on a tablet press should handle not only the fine and coarse granulation, but also the broken tablet pieces that may have been picked up on the turntable or placed there by the press operator.

## *Chemical Weighing*

Figure 26.4 shows a typical weighing operation units located inside booths and each booth equipped with a dust collection hood. Although a point-of-use collector may be used, a central unit gives better results when a large room is required or when several weighing booths are used. In such cases, HEPA filtration should be used. The hood for each of the 14 × 15 foot booth should have a capacity of 4500 cm with a face velocity in excess of 150 feet per minute.

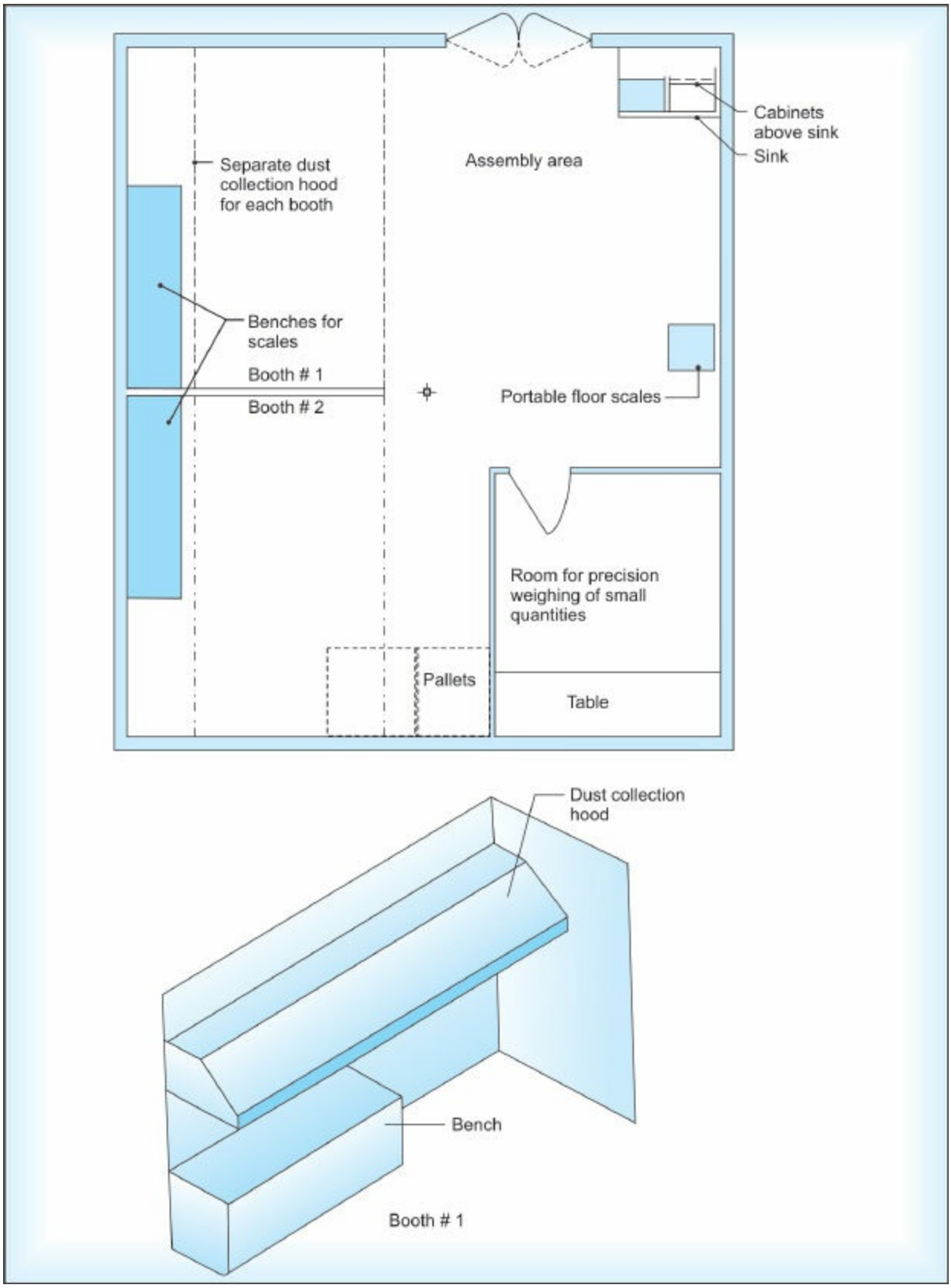


**Fig. 26.4:** Top, two weighing booths together with enclosed room for weighing steroids and the like. Bottom, workbench and hood, which should be connected to dust collector

## *Tablet Granulating and Compressing*

Since granulating operations use several different pieces of equipment, each with its own unique problems, dust collection can best be described in general terms. If the area is air-conditioned, it is possible to minimize cost by reusing 85% of the air, provided a HEPA filter is used. Flexible 3 or 4 inch hoses should have an air capacity of 200 to 300 cm and a linear velocity of more than 2500 feet per minute for best results. When applied to tablet presses, however, air capacity of at least 450 cm and a velocity of over 3000 feet per minute are needed.

## *Coating and Tablet Packaging*

A typical 42 inch coating pan should have a supply inlet of 200 cm and an exhaust of 300 cm when standard ducts are used. Absolute filters are preferable if 85% of the air is to be recirculated. If solvent-type film coating is performed in a conventional Accela Cota or Pellegrini pan, not only does the air volume have to be increased substantially, but also the discharged air must be treated to conform to the local and government environmental standards.

Flexible hoses used at tablet counters, powder fillers, and cottoning machines should handle about 200 to 300 cm at a minimum velocity of 2000 feet per minute.

## Space Requirements

A pilot plant has the following four types of space requirements.

### *Administration and Information Processing*

Documentation is extremely important and therefore adequate office and desk space must be provided for both the scientists and technicians. This should be adjacent to the work area but sufficiently isolated so as to permit the personnel to work without undue distractions. Since this group is the link between research, operations, and other disciplines, members of this group frequently meet with the personnel from other departments and should have an area available where three to four people can meet and discuss subjects of mutual concern. There should also be space for a computer terminal for the convenient entry and retrieval of the data as well as an archives for the stability data protocols and historical files.

### *Physical Testing Area*

The second area required is an adequate working area in which the samples can be laid out and examined and where physical tests on these samples can be performed. This area should provide permanent bench-top space for routinely used physical testing equipment (e.g. balance, pH meter, and viscometer).

### *Standard Pilot Plant Equipment Floor Space*

The third area is a discrete plant space where equipment needed for manufacturing all types of pharmaceutical dosage forms is located. The equipment should be available in a variety of sizes known to be representative of the production capability. This arrangement helps assure the quality of the scale-up data collected, while meeting the concern to be prudent with expensive materials. Intermediate-sized and full-scale production equipments are essential for evaluating the effects of scale-up on the research formulations and processes. Utilization of this area is most efficient when it is subdivided into separate areas for solid dosage forms, semisolid products, liquid preparations, and sterile products. Further subdivision of the areas should allow multiple operations to be conducted simultaneously without raising any GMP concerns.

Because the utilization of the pilot plant equipment is sporadic and dependent on the project assignments, equipment should be made portable, where ever possible. The equipment can then be stored in a relatively small area and brought out into the suitable work areas for use when required. This system helps relieve some of the congestion often found in pilot plant operations and provides more working space around the equipment that is in use. Such a system also provides more space if equipment is brought in for evaluation on a loan or on a rental basis. An essential requirement that is frequently neglected but that is important to all pilot plant operations is the provision of adequate space for cleaning of pilot plant equipment. While some equipment can be cleaned in place, most equipment is better handled in a dedicated cleaning area.

## *Storage Area*

The fourth area and the one most often described as inadequate is the storage space. Separate provision should be made for the storage of active ingredients and excipients. These should further be segregated into the approved and unapproved areas according to GMPs. There should be generous storage areas for in-process materials, finished bulk products from the pilot plant, and material from experimental scale-up batches made in production, which for GMP reasons cannot be stored in operations storage areas. Space should be provided for the storage of retained samples from pilot plant and experimental production batches. All of these space requirements are in addition to the controlled environment space allocated for the storage of the stability samples.

Finally, there should be space for the storage of packaging materials. These materials tend to be bulky, but since a common requirement of the scale-up function is to evaluate alternate suppliers of packaging materials, it is essential to provide the space required to store bottles, closures, tubes, vials, ampuls, etc. Here too, it would be preferable if the material could be segregated into approved and unapproved categories.

## Space Allocation

As important as the provision of adequate space for production is the proper space allocation for each operation. A departmental analysis of space allocation is beyond the scope of this section, but a comparison of some major areas is of interest.

The data in Table 26.2 represent the facilities of a major pharmaceutical company whose operations extend to more than 45 countries and over 140 plants. Included in the tabulation are domestic and international operations representing prescription products (Items 1, 5, 6, 7 and 8); proprietary products (Item 4); and for general comparative purposes, Items 2 and 3, which represent a combination of proprietaries and/or mints and chewing gum operations.

**Table 26.2:** Space allocation (thousand square feet)

| Plant | Total | Office | Production | (%) | Warehouse | (%) | Other |
|---|---|---|---|---|---|---|---|
| 1 | 1652.6 | 159.7 | 756.5 | (45.7) | 367.7 | (22.2) | 368.7 |
| 2 | 615.8 | 7.5 | 163.3 | (26.5) | 196.3 | (31.8) | 248.7 |
| 3 | 595.8 | 23.1 | 270.5 | (45.4) | 172.1 | (28.8) | 130.1 |
| 4 | 502.7 | 23.4 | 128.3 | (25.5) | 300.7 | (59.8) | 50.3 |
| 5 | 502.6 | 24.5 | 207.3 | (41.2) | 215.5 | (42.8) | 55.3 |
| 6 | 304.5 | 40.9 | 84.4 | (27.7) | 70.5 | (23.1) | 108.7 |
| 7 | 286.2 | 16.7 | 57.3 | (21.0) | 77.1 | (26.9) | 135.1 |
| 8 | 77.7 | 3.2 | 30.6 | (39.3) | 16.0 | (20.5) | 27.9 |

It is obvious that the production department and the warehouse occupy most of the space. One might also observe that plants 6, 7, and 8, whose total size is less than 500,000 square feet, utilize 21 to 39.3% of the total area for production. In plants occupying more than 500,000 square feet (1, 3 and 5), however, a significantly larger percentage of the area (i.e. over 40%) is devoted to production.

The same Table shows little percentage difference in the space occupied by the warehouse, regardless of plant size. The only exception is plant 5, which includes a significantly larger biologic operation, and plant 4, which is a proprietary operation and has the highest warehousing occupancy among plants. This operation is dominated by a few high-volume proprietaries that require large continuous manufacturing operations, thus minimizing space

needs for production while increasing the size of warehousing areas.

Plant 2 is a combination of high-volume proprietaries and gum operations, a substantial number of which require large-batch and continuous operations. The result is a production area that is slightly smaller than the warehouse. Plant 3, a pressed mints and gum operation, is largely automated, but the massive equipment occupies considerable space. On the other hand, although there is a high volume of case goods, the products are small and compact, and less space is needed for the warehouse.

In addition to space allocation by these broad categories, GMP Section 133.3 (Buildings) states that adequate space must also be provided for the following (1) orderly placement of equipment and materials to minimize mixups or cross-contamination, (2) receipt and storage of all material in a quarantine area prior to use, (3) holding of rejected materials and (4) manufacturing and processing operations and other such functions.

# PRODUCT CONSIDERATIONS AND MANUFACTURING FACILITIES

## SOLID DOSAGE FORMS

In scaling up the manufacture of tablets and capsules from experimental laboratory batch sizes to intermediate and large-scale production, each stage of the operation must be carefully considered. A process using the same type of equipment performs quite differently when the size of the equipment and the amount of material involved is increased significantly. Even a simple operation like loading a mixer can become a complicated operation utilizing sophisticated equipment when large volumes are involved. In some instances, scale-up may involve a major process change that utilizes techniques and equipment that were either unsuitable or unavailable on a laboratory scale.

The following are the typical unit operations involved and manufacturing facilities required for production of solid dosage forms.

## Material Handling

In the laboratory, materials are simply scooped, dumped, or poured by hand. This may also work well in some small or intermediate-sized production operations, but in other intermediate or large-scale operations, mechanical means of handling these materials often become necessary. These mechanisms range from simple post hoists or other mechanical devices for lifting, and tilting drums, to more sophisticated methods of handling materials, such as vacuum loading systems, screw feed systems, and metering pumps. The type of system selected also depends on the characteristics of the materials, e.g. density or static charge.

Any material handling system must deliver the accurate amount of the ingredient to the intended destination. Lengthy transfer lines may result in material loss, for which there must be accountability and compensation. If the system is used to transfer materials for more than one product, steps must be taken to prevent cross-contamination. This can be accomplished by the use of validated cleaning procedures for the equipment.

## Chemical Weighing

This is an important step in the manufacturing process and has been receiving an increasing amount of attention because of possibilities for cross-contamination and misbranded products due to incorrect ingredients or quantities. Many companies have adopted a central weighing department to service all of the processing areas. The advantages of this system are the (1) centralization of responsibility, (2) avoidance of duplicating weighing facilities, (3) lower labor costs. After an item is weighed and properly initialed on the batch sheet by the weigher, check-weighed and initialed by a checker, and properly packaged and identified, additional weighing at the time of usage is usually unnecessary. A chemical weighing department should be designed to provide supervision, checkers, proper weighing equipment, lighting, dust collection, and adequate sanitation.

High-potency drugs such as steroids and alkaloids should be weighed in a separate room equipped with absolute filters to avoid even minimal cross-contamination. This room could also be used for weighing dyes (*see* Fig. 26.4).

Sinks and drain boards should be conveniently located to facilitate frequent cleaning of measuring equipment. Cabinets should be provided for the storage of utensils.

Vacuum hoses should be available in the weighing area immediately adjacent to the weighing booths so that the tops of drums and other containers can be cleaned free of dust before they are opened for removal of contents.

Balances and scales having the proper capacity and sensitivity needed for weighing operations should be specified, and arrangements made for frequent calibration. Printing scales that record weights on formula sheets and container labels should also be provided.

Meters should be used when liquid materials are being transferred from storage tanks directly to manufacturing tanks. Each quantity should be recorded on batch sheets, either manually or by means of a printing system. The meters should be calibrated and checked periodically.

# Tablet Blending and Granulation

## *Blending and Granulation*

Powders which are to be used for encapsulation, or are to be granulated prior to tabletting or encapsulation, must be well blended to ensure good drug distribution. Inadequate blending at this stage could result in discrete portions of the batch being either high or low in potency. This could result in drug content uniformity variation, especially if the tablet or capsule is small and the drug concentration is relatively low in the blend. Ideally, the dry blend should take place in the vessel in which subsequent processing such as granulation occurs. Not every manufacturing facility, however, has equipment with sufficient capacity to accommodate such process. Consequently, a larger batch may be dry blended and then subdivided into multiple sections for the granulation operation. Steps should also be taken to ensure that all the ingredients (excipient and active) are free of lumps and agglomerates prior to the dry blend. Failure to remove or break up all agglomerates could cause flow problems through the equipment, creating nonreproducible compression and encapsulation processes with a detrimental effect on the content uniformity of the product. For these reasons, screening and/or milling of the ingredients prior to blending usually makes the process more reliable and reproducible.

The blending equipment used in production operations differs considerably in size from that used in the product development laboratories. Consequently, attention should be paid to the scale-up of this operation so that equipment of the right design is used and blender loads, mixing speeds, and mixing times are properly established. In any blending operation, both segregation and mixing occur simultaneously. Both processes are a function of particle size, shape, hardness, and density, and of the dynamics of the mixing action. Therefore, the characteristics of the different particles in the blend must be known, and the cause of segregation understood, so that the blending operation can be optimized and a uniform blend obtained.

## *Dry Blending and Direct Compression*

Processes that yield free-flowing granulations without the aid of granulating solutions are desirable in that they do not require the time and energy necessary to volatilize the solvent used in conventional wet granulation

procedures. With many of the common excipients and even active drugs now available in a form that makes them directly compressible, the possibility of dry blending and direct com pression should be considered. When exploring this option, a careful analysis of particle characteristics that influence mixing and segregation, such as size, size distribution, shape, and static charge should be evaluated. Scaling up a dry blending operation for a directly compressible formulation requires that special attention be paid to blender loads, optimum mixing speeds, and the blending time so that the drug distribution within a batch is acceptable and it is consistent from batch to batch.

The following are aspects of the dry blending operation of directly compressible materials that can be adjusted to optimize the process.

*The order of addition of components to the blender:* A low-dose active ingredient may be "sandwiched" between two portions of directly compressible excipient in the blender to improve dispersion and/or to avoid loss to the surface of the blender.

*The mixing speed:* Examples are blade rotation speed for a planetary type mixer, and mixer tumbling or rotational speed for a twin-shell, cone-type, or similar type of mixer.

*The mixing time:* The mixing time can be decreased if available data show the materials to be consistently and uniformly mixed in less time than originally directed. Alternatively, the time may need to be increased if the mixing time is shown to produce material with borderline uniformity. Mixing time also affects the compressibility of the finished blend. Excessive mixing time may fracture fragile excipients and ruin their compressibility.

*The use of auxiliary dispersion equipment within the mixer:* An example is an intensifier bar or chopper blade in a twin-shell mixer. These increase the efficiency of dispersion of solid and liquid ingredients added to the mixture. They also reduce agglomerates that may be present in a material, thereby aiding the efficiency of dispersion.

*The mixing action:* Mixing action is determined by the mechanics of the mixer and can only be changed by converting from one blender to another or by modifying the blender through addition of baffles or plates, which would alter the mixing characteristics.

*The blender load:* The amount of material volume to the total mixer

volume affects the efficiency of the blender. Each blender has an optimum working volume and a normal working range. Overloading a blender retards the free flow of the granulation and reduces the efficiency of the blender. Localized concentrations of the drug remain, causing content uniformity problems in the finished dosage form. Conversely, if the load is too small, the powders slide rather than roll in a blender, and proper mixing does not occur, or the time needed for uniform mixing of the powders increases.

## *Slugging (Dry Granulation)*

A dry powder blend that cannot be directly compressed because of poor flow or compression properties may in some instances be processed using a slugging operation. This is done on a tablet press designed for slugging, which operates at pressures of about 15 tons, compared with a normal tablet press, which operates at pressures of 4 tons or less. Usually, extra-large tablet punches are used to form compressed slugs of the powdered material. This procedure is slow because, the inherently poor compressibility of the powders requires slower press speeds to provide the extended compression dwell time under load needed to hold the compacted material together. Slugs range in diameter from 1 inch, for the more easily slugged material, to inch in diameter for materials that are more difficult to compress and require more pressure per unit area to yield satisfactory compacts. After compression, the slugs are broken down using either a hammer mill or an oscillating granulator to obtain a granulation with a suitable particle size distribution. If an excessive amount of fine powder is generated during the milling operation, the material must be screened and the “fines” recycled through the slugging operation. During scale-up of such an operation, the pilot plant scientist should pay particular attention to the forces used for the slugging operation, the diameter of the punches, and the subsequent sizing and screening operations. The optimization of these variables affects the particle size and particle size distribution.

Granulation by dry compaction can also be achieved by passing powders between two rollers that compact the material at pressures of up to 10 tons per linear inch. Because of the similarity in application, processes developed in the laboratory using a slugging operation might be adapted to the roller compaction process as shown in Fig. 26.5. Materials of very low density require compaction to achieve a bulk density sufficient to allow encapsulation

or compression. One of the best known examples of this process is the densifi-cation of aluminum hydroxide. Pilot plant personnel should determine whether the final drug blend or the active ingredient could be more efficiently processed in this manner than by conventional processing in order to produce a granulation with the required tabletting or encapsulation properties.


**Fig. 26.5:** Roller compactor for dry compaction operation *{Courtesy* of ERWEKA)

## *Wet Granulation*

To scale up a granulation process in the most efficient manner, the purposes for granulating must be clearly understood. The most common reasons given to justify granulating are (1) to impart good flow properties to the material so that the tablet presses and encapsulators can be properly fed and a uniform tablet or capsule weight maintained, (2) to increase the apparent density of the powders and (3) to change the particle size distribution so that the binding properties on compaction can be improved. A small amount of a potent active ingredient can be dispersed most effectively in a carrier granulation when the drug is dissolved in the granulating solution and incorporated into the batch during granulation. Use of the granulation process to disperse an active ingredient is also commonly cited as a reason to granulate.

Some pieces of equipment are more suitable than others for helping to develop the desired characteristics of a finished granulation. Traditionally, wet granulation has been carried out using sigma blade or heavy-duty

planetary mixers (Fig. 26.6). Production equipment of this type equipped with large motors of 7 to 10 horsepower can process 100 to 200 kg of material. The weight of material and the large shear forces generated by these powerful units affect not only the granulating time but also the amount of granulating fluid required relative to that used in experimental laboratory trials. Wet granulations can also be prepared using tumble blenders equipped with high-speed chopper blades (Fig. 26.7).

**Fig. 26.6:** Planetary mixer (*Courtesy* of ERWEKA)

**Fig. 26.7:** Twin shell blender. Material splits and refolds as the blender rotates, producing uniform blends. Various intensifier bars can handle specific applications

High-shear mixers are often more effective in densifying light powders, but require large amounts of energy and have limited load size. There are pieces of equipment available that combine the high-shear mixing action

often required for good densification with the advantages of high-speed choppers, which break up agglomerates and ensure uniform distribution of the granulating fluid and more controlled granule size (Fig. 26.8).

**Fig. 26.8:** High-shear pharmaceutical grade mixer/granulator

Binders are used in tablet formulations to make powders more compressible and to produce tablets that are more resistant to breakage during handling. Some of these are added in the dry state and impart their binding properties when exposed to the granulating fluid. Others are dissolved or dispersed in the granulating fluid. In some instances, the binding agent imparts considerable viscosity to the granulating solution, so that transfer of the fluid by either pumping or pouring becomes difficult. During the scale-up of such a process, problems could be encountered during the addition of the granulating agent to the powders being processed in enclosed equipment. If the problem is anticipated during the formulation stage, the viscosity of the granulating solution can be adjusted so that scale-up problems of this type can be avoided. One way of avoiding this problem is to disperse some or all of the binding agent in the dry powder prior to granulating. The granulating liquid containing any remaining binder can then be easily pumped and metered into the batch during granulation.

Occasionally, nonaqueous solvents or solutions that are composed of water and water-miscible solvents are used to improve the granulating properties of a formulation or to disperse poorly soluble drugs. While a reduced amount of energy is required to remove the more volatile solvent(s) from the granulation, proper ventilation and additional safety precautions

(against fire, toxicity, explosion) must be considered in the selection and design of the equipment and manufacturing area. In addition, solvent recovery systems may be necessary to comply with Environmental Protection Agency (EPA) or Occupational Safety and Health Administration (OSHA) regulations.

Some granulations, when prepared in production-sized equipment, take on a doughlike consistency and may have to be subdivided to a more granular and porous mass to facilitate drying. This can be accomplished by passing the wet mass through an oscillating type granulator with a suitably large screen, or a hammer mill with either a suitably large screen or no screen at all.

## *Drying*

The most common conventional method of drying a granulation continues to be the circulating hot air oven, which is heated by either steam or electricity, and in which the granulation is spread on paper-lined trays on a rack truck, which is then wheeled into the oven where evaporative drying occurs. The important factors to consider as part of scale-up of an oven drying operation are airflow, air temperature, and the depth of the granulation on the trays. If the granulation bed is too deep or too dense, the drying process will be inefficient, and if soluble dyes are involved, migration of the dye to the surface of the granule may occur. During scale-up of this operation, the granulation bed depth should be carefully controlled and the drying process monitored by the use of moisture and/or temperature probes in the granulation, or by frequent multipoint sampling of the granulation for moisture content throughout the drying phase. Drying times at specified temperatures and airflow rates must be established for each product, and for each particular oven load.

Fluidized-bed dryers (Fig. 26.9) are an attractive alternative to the circulating hot air ovens. Their main advantage is a reduction in drying time. Fluidized-bed drying times are usually less than 1 hour, compared with 8 hours or more in the conventional ovens. Many products can also be dry blended and granulated in a fluidized-bed granulator/dryer, further reducing the handling, and consequently the time, required to process a batch. Scale-up of a fluidized bed drying operation is more involved than scale-up of a circulating hot air oven process. First, optimum loads must be established. Then, rate of airflow and inlet air temperature as Well as the humidity of the

incoming air must be established, since these all affect the drying time. If the air is drawn from outside the plant without being conditioned, the large seasonal variations in temperature and humidity that may exist can alter the drying process. Scale-up is further complicated in that it has been shown that data from small-scale batches (1 to 5 kg) cannot be used to extrapolate processing conditions for intermediate-scale (100 kg) or larger batches. To obtain a granulation comparable to that obtained on a smaller scale, considerable experimentation and adjustments to the process are required.

**Fig. 26.9:** Fluidized-bed dryers of 1,5 and 50 kg capacity

## *Reduction of Particle Size*

Particle size and especially particle size distribution are important to the compression characteristics of a granulation. In the laboratory, hand screening or short-duration handling with small-scale milling equipment is used to obtain the desired particle size distribution prior to compression or encapsulation. When such a process must be increased in capacity to accommodate large, high-speed presses with more elaborate feed systems, it becomes important that the equipment chosen can yield the desired throughput while controlling the particle size and size distribution of a granulation.

Compression factors that may be affected by the particle size distribution are flowability, compressibility, uniformity of tablet weight, content uniformity, tablet hardness, and tablet color uniformity. A granulation with too large a particle size and insufficient fines is unable to fill the die cavities uniformly during compression, and the weight of the tablets fluctuates considerably. For colored granulations, the coarser the granulation, the more

mottled the final tablet appearance is. If too many fines are present, tablet weight variation occurs because of flow problems. Also, the tendency toward capping increases and is further exaggerated as the speed of the press is increased. Both oversized and undersized granulations can adversely affect tablet content uniformity. Particle size reduction of the dried granulation of production-size batches can be carried out by passing the material through an oscillating granulator (Fig. 26.10), a hammer mill (Fig. 26.11), a mechanical sieving device (Fig. 26.12), or in some cases, a screening device.

**Fig. 26.10:** Oscillating granulator. Oscillating motion of horizontal bars forces granulation through a screen, resulting in uniform granules that are more easily dried

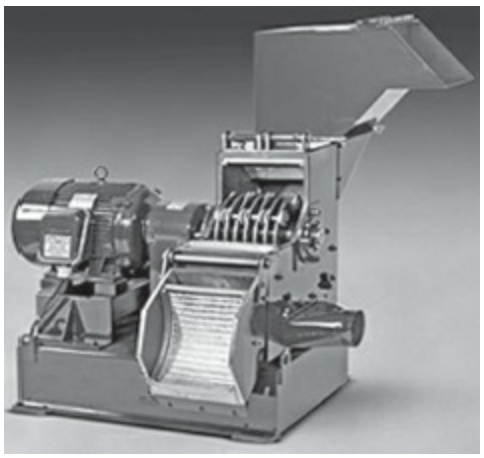


**Fig. 26.11:** Hammer mill (comminuting machine). The comminuting action occurs inside the mill head, where the high-speed centrifugal force of the blades hurls the granulation through the screen, which is held in place at the

discharge port

Many of the newer granulation operations produce material with particle size distributions already quite close to the desired range. Then, the sizing operation only requires that a small amount of agglomerates be broken down. For this type of operation, the material can either be subjected to a screening operation as described or be passed through a mechanical sieving operation, using equipment of the type shown in Fig. 26.12. The advantage of using this type of equipment is that there is no metal-to-metal contact during the screening process, so that the possibility of metal contamination is reduced. Since there is also little milling action, the initial particle size range is not significantly reduced. In addition, the throughput is rapid, and because of the enclosed nature of the equipment, little dust is created; consequently, material losses are low, and exposure of personnel to dust is minimized.

**Fig. 26.12:** Sieving equipment (*Courtesy* of ERWEKA)

As part of the scale-up of a milling or sieving operation, the lubricants and glidants, which in the laboratory are usually added directly to the final blend, are usually added to the dried granulation during the sizing operation. This is done because some of these additives, especially magnesium stearate, tend to agglomerate when added in large quantities to the granulation in a blender. To assure adequate distribution of these dry additives, a preliminary dispersion of these materials is often made during the sizing operation. This part of the process must be carefully optimized so that the lubricants are not over-mixed or undermixed during the screening and subsequent blending

operations.

## *Facilities*

In general, several different products are in production at any given time. The numerous steps in the granulating procedure increase the possibilities of cross-contamination, incorrect product identification, and/or mixups. To eliminate these possibilities, a separate room or booth is recommended for each step. Thus, more space is required and maintenance costs are higher because the equipment and each room must be thoroughly cleaned between operations. In many cases, the cleaning costs are buried in the indirect labor category, when truly it represents changeover costs as in a packaging operation.

Compartmentalizing the granulating process has, unfortunately, fragmented the operation and increased space, capital, and labor costs. Granulating should be considered a unit operation composed of closely integrated manufacturing steps, and process development work should be directed to this area for cost reduction and process improvement. Such effects help reduce granulating costs, which are invariably higher than tabletting costs when compared on a cost-per-thousand-tablet basis.

A washing facility (Fig. 26.13) should be available for cleaning portable equipment such as granulators and mills. To facilitate cleaning of non-portable equipment, such as fluid-bed driers, mixers, etc. each room should be provided with floor drains and a pitched floor (4 inch per foot) as well as hot and cold water and steam for special cleaning jobs. Particular attention should be devoted to the cleaning of drying racks and trays, which should be designed for easy cleaning and made of stainless steel or other non-rusting material. If the department is not air-conditioned, all windows should be screened against the entrance of insects. The aforementioned precautions are equally applicable to the manufacturing of powders and bulk materials for capsule filling.

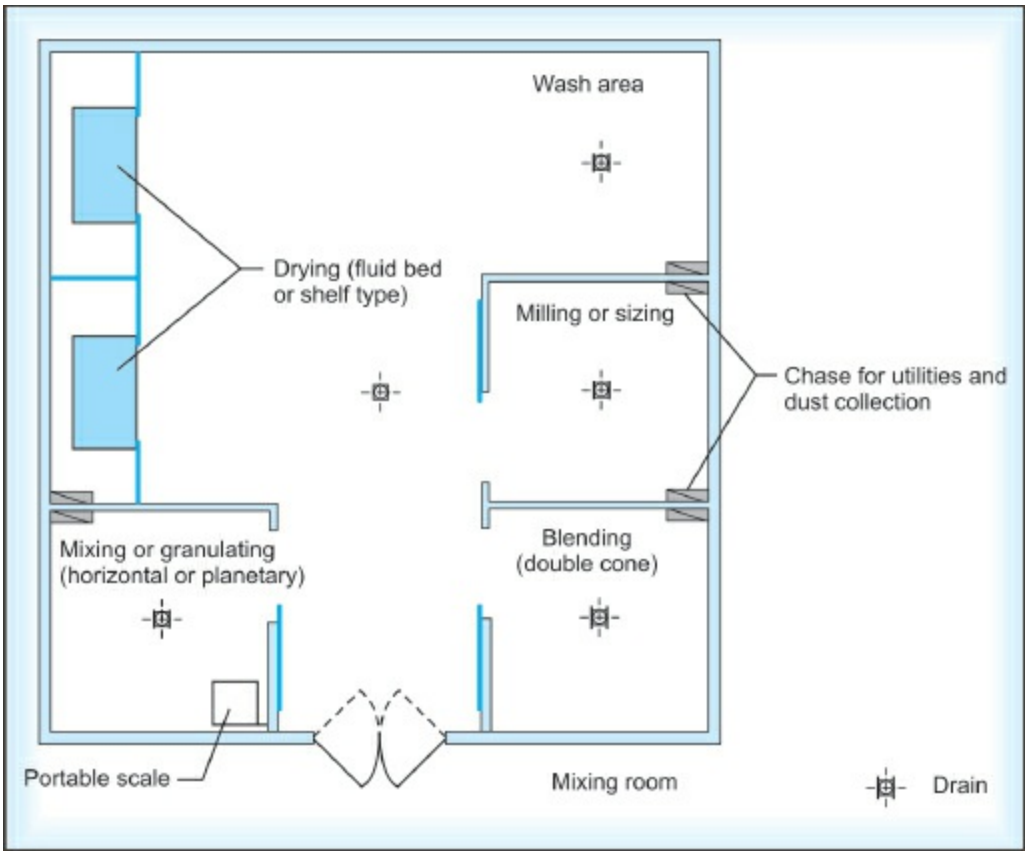



**Fig. 26.13:** Typical granulating room showing separation of milling, blending and granulating areas. Note drain in each area

More recently there has been trend towards the use of multifunctional processors (Fig. 26.14). These are units capable of performing all the functions required to prepare a finished granulation such as dry blending, wet granulation, drying, sizing and lubrication in a continuous process in a single piece of equipment. The advantages of using such equipment during scale-up of a product can be significant in terms of space and man power requirements. Closed continuous systems have the added advantage of less handling of materials, thereby reducing the danger of personnel exposure to potent materials.

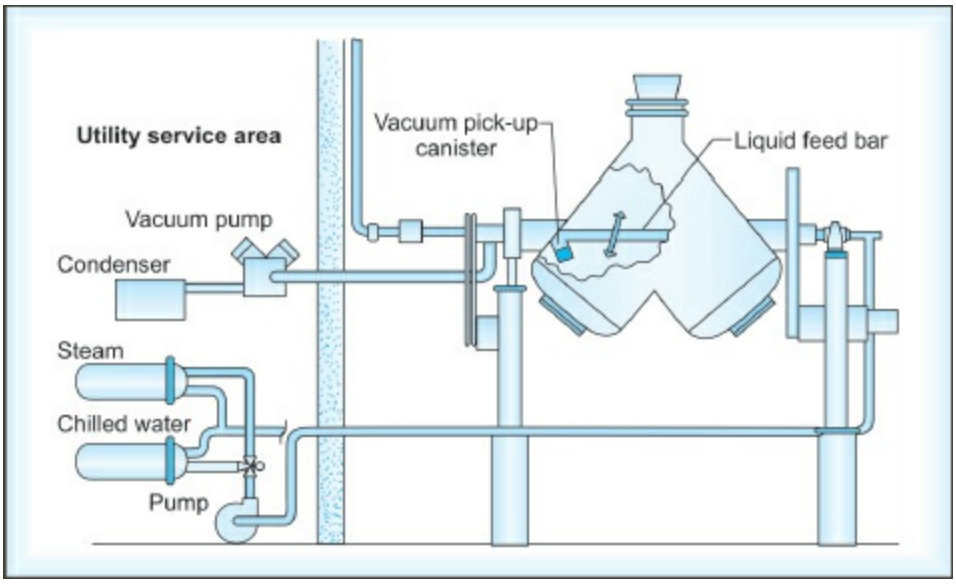



**Fig. 26.14:** Schematic of a multifunction pharmaceutical processor. The system contains a solids mixing shell, which includes a liquid dispersion bar. This unit is jacketed to allow heating and cooling. Material can be dried within the unit by use of the vacuum canister and condenser

## *Granulation Handling and Feed Systems*

The handling of the finished granulation in the compression area can be a simple operation such as hand scooping the material from a drum into the press hopper, or for larger operations, a sophisticated automated handling system using vacuum or mechanical systems to convey the granulation. For the latter system, studies should be undertaken to determine the effect that this additional handling has on the content uniformity of the drug and on the particle size distribution. Segregation due to static charges built up during vacuum and/or the mechanical handling of the granulation can lead to problems with material flow through tablet press hoppers and feed frames. This in turn makes it difficult to control tablet weight, thickness, and hardness. Poor content uniformity may be the final result.

More sophisticated material handling systems cause the added concern of cleanability. Long lengths of transfer tubes, valves, pneumatic pumps, vacuum canisters, cyclone traps, and other components of these systems must be engineered for efficient and total cleaning. Well-written, documented, and validated cleaning procedures are essential for such a system.

## Tablet Compression

The ultimate test of a tablet formulation and granulation process is whether the granulation can be compressed on a high-speed tablet press.

During compression, the tablet press performs the following functions:

1. Filling of empty the cavity with granulation.
2. Precompression of granulation (optional).
3. Compression of granulation.
4. Ejection of the tablet from the die cavity and take-off of compressed tablet.

The means by which these functions are accomplished varies, depending on the design of the press. Machine design also determines the usable range of compression forces at which the machine can safely operate, and the press speed at which output can be optimized without negative impact on tablet quality.

Sometimes, because of raw material characteristics or formulation constraints, a particular product cannot be successfully compressed at the upper speed range of a press. When this occurs, the press speed is reduced, or a slower press is used, to allow more time for the dies to fill and to extend the dwell time of compression, both of which help to facilitate the compaction process. With advances in tablet press technology, which provide better control of feeder mechanisms, precompression, and compression forces, many granulation problems or inadequacies can be overcome by making appropriate adjustments to the press. The handling and compression characteristics are important factors that must be considered in the selection of a tablet press. Output alone is not reason enough to select a press (Table 26.3) contains a partial list of presses available and their capabilities and features.

When evaluating the compression characteristics of a particular formulation, prolonged trial runs at press speeds equal to that to be used in normal production should be tried. Only then are potential problems such as sticking to the punch surfaces, tablet hardness, capping, and weight variation detected. Such preproduction trials in the pilot plant are important for identifying these problems early in the scale-up process, when changes are more easily made than during later production runs, when marketing

requirements may make it difficult to interrupt production schedules to modify the formulation.

**Table 26.3:** Compression rates of typical production presses

| Model | Max. tablet diameter (in.) | Number of stations | Tablets per min. |
|---|---|---|---|
| Colton 216 | ½ | 16 | 600–800 |
| Colton 233 | ⅝ | 33 | 2500–4000 |
| Colton 241 | 7/16 | 41 | 3000–5000 |
| Stokes B-2 | 7/16 | 22 | 450–900 |
| Stokes BB-2 | ⅝ | 27 | 800–1400 |
| Stokes BB-2 | 7/16 | 33 | 1000–1700 |
| Stokes 540 | ⅝ | 35 | 800–2400 |
| Stokes 541 | 7/16 | 41 | 1000–2700 |
| Stokes 551 | 7/16 | 51 | 1200–3700 |
| Stokes 552 Tripact | 7/16 | 51 | 3000–5000 |
| Stokes Ultrapress 565-1 | 7/16 | 65 | 3500–10,000 |
| Stokes Ultrapress 565-2 | ⅝ | 53 | 2900–8100 |
| Stokes Eagle-3 | 5/16 | 41 | 2150–6150 |
| Stokes Eagle-2 | ⅝ | 53 | 2800–8000 |
| Stokes Eagle-1 | 7/16 | 65 | 3500–10,000 |
| Manesty Betapress | ⅝ | 16 | 700–1500 |
| Manesty Betapress | 13/22 | 23 | 100–2000 |
| Manesty Express × 20 | 1 | 20 | 800–2000 |
| Manesty Express × 25 | ⅝ | 25 | 1000–2500 |
| Manesty Express × 30 | 7/16 | 30 | 1200–3000 |
| Manesty Rotapress | 7/16 | 55 | 1300–5000 |
| Manesty Rotapress | ⅝ | 45 | 1100–4300 |
| Manesty Rotapress mark III | 7/16 | 69 | 3300–10,000 |
| Manesty Unipress | 15/16 | 20 | 970–2420 |
| Manesty Unipress | ⅝ | 27 | 1300–3270 |
| Manesty Unipress | 7/16 | 34 | 1640–4120 |
| Fette Perfecta 3000 | 15/16 | 37 | 2200–4400 |
| Fette Perfecta 3000 | ⅝ | 45 | 2700–6750 |
| Fette Perfecta 3000 | ½ | 55 | 3300–8250 |
| Fette Perfecta 1000 | ⅝ | 28 | 560–2100 |
| Courtoy R100 | 15/16 | 24 | 288–2300 |
| Courtoy R100 | ⅝ | 30 | 360–2880 |
| Courtoy R100 | ½ | 36 | 1260–4400 |
| Kilian LX | 15/16 | 15 | 270–1350 |
| KilianLX | ⅝ | 21 | 400–2260 |
| Kilian TX | ⅝ | 30 | 330–3150 |
| Kilian TX | ½ | 40 | 440–4200 |
| Kilian RX | 15/16 | 41 | 660–5500 |
| Kilian RX | ⅝ | 51 | 830–7660 |
| Kilian RX | ½ | 67 | 1080–10,000 |
| Kikusui Gemini | ⅝ | 55 | 2200–7700 |
| Kikusui Gemini | 7/16 | 67 | 2680–9380 |
| Kikusui Hercules | 1 ½ | 18 | 450–1080 |
| Kikusui Hercules | 15/16 | 29 | 725–1740 |
| Korsch PH230 | ⅝ | 17 | up to 1850 |
| Korsch PH230 | ½ | 20 | up to 2200 |
| Korsch PH336 | ⅝ | 36 | up to 3600 |
| Korsch PH343 | ½ | 43 | up to 4300 |
| Korsch PH423 | 1 ½ | 23 | 368–2300 |
| Korsch PH431 | 15/16 | 31 | 500–3100 |
| Korsch PH447 | 9/16 | 47 | 750–4700 |

High-speed tablet compression depends on the ability of the press to interact with granulation so that a certain series of operations is successfully performed. The granulation must be delivered to the die feed system at an adequate rate. The granulation delivery must not be interrupted, nor should the flow rate vary. The delivery system must not change the particle size distribution. The system must not cause segregation of coarse and fine particles, nor should it induce static charges, which would retard the flow of the granulation and could cause the active ingredient to become segregated. The die feed system must be able to fill the die cavities adequately in the short period of time that the die is passing under the feed frame. The smaller the tablet, the more difficult it is to get a uniform fill at high press speeds. The granulation must have good flow properties, a good particle size distribution, and a relatively small mean particle size to facilitate rapid but uniform fill of the die cavities. For high-speed machines, induced the feed systems are necessary. These are available with a variety of feed paddles and with variable speed capabilities, so that the optimum feed for every type of granulation can be obtained.

After the die cavities have been filled with granulation, the excess is removed by the feed frame to the center of the die table. If the feed frame has been overfilled and not all of the excess granulation is moved to the center of the die table, this excess may be thrown from the table by the centrifugal force of the rotating the table. For this reason, the clearance between the scraper blade and the die table must be carefully set. Too large a gap results in large granulation losses, especially if the granulation contains a lot of fines. Too close a setting causes scoring of the die table and metal contamination of the product.

Compression of the granulation usually occurs as a single event as the heads of the punches pass over the lower and under the upper pressure rollers. This causes the punches to penetrate the die to a preset depth, compacting the granulation to the thickness of the gap set between the punch surfaces. The rapidity and dwell time in which this event occurs is determined by the speed at which the press is rotating and by the size of the compression rollers. The larger the compression roller, the more gradually the compression force is applied and released. The tendency toward capping in a formulation can often be reduced by slowing down the press speed or using presses with larger compression rollers. Granulations that are difficult to

compress and that have a tendency to cap can often be more effectively compressed on a press with a series of pressure rollers that impart a stepwise increase in pressure, thus allowing entrapped air to escape gradually rather than as -an abrupt event at the end of a single-step compression.

Courtoy has developed a tablet press that is capable of minimizing capping by a unique compression system whereby pressure is applied to the punch heads by pressure rollers that are pneumatically loaded. Instead of the pressure profile being the customary sine wave (  ) it becomes more of a square wave (  ). Thus, the dwell time at the point of maximum pressure is prolonged considerably. Subjecting the tablet to the maximum compression force for an extended time period reduces the tendency of the tablet to cap and permits running a problem granulation at higher than normal press speeds.

The final event in the compression process is the ejection of the compressed tablet from the die cavity. This involves the separation of the upper punch from the upper surface of the tablet and withdrawal from the die cavity. The lower punch face then moves up through the die cavity, breaking the tablet free from the die wall and forcing it out of the die cavity. As the die table rotates, a take-off bar positioned just above the table forces the tablet to separate from the lower punch face and sweeps the tablet off the press table into a collection chute.

During compression, the granulation is compacted, and in order for a tablet to form, bonds within the compressible material must be formed. The forces that give rise to strong cohesive bonds within the material also exist at the tablet interfaces and may result in adhesive bonds between the punch and the surfaces and the tablet. Embossing on the punches tends to accentuate sticking. A good internal lubricant system is necessary to prevent sticking of the tablet to the metal surface of the punches or die. If the granulation is adequately and properly lubricated, the tendency for sticking and binding of the tablet to the punch face can be eliminated. Magnesium stearate and calcium stearate are the most commonly used tablet lubricants. The level at which they should be employed and the degree to which they need to be blended with the granulation must be determined experimentally. Too high a level of lubricant or overblending can result in a soft tablet, a decrease in the wettability of the powder, and an extension of the dissolution time.

The design and condition of the punches can also be the cause for

sticking. Embossing on the punch faces often causes sticking if the embossed letters, numbers, or symbols are too high, if the angle of the embossing is too steep, or if the corners are too sharp. New punches often have to be “run in” over a 4 to 8-hour period before they can run cleanly. Microscopic nicks or pits in the punch faces can also cause sticking.

Binding at the die walls can sometimes be overcome by designing the die to be 0.001 to 0.005 inch wider at the upper portion than at the center in order to relieve pressure during ejection.

## *Facilities*

### *Booths or Rooms*

Separate rooms for tablet machines have now become a necessary design feature to avoid cross-contamination. When special low-humidity conditions are necessary to ensure product stability, a chemical unit employing lithium or silica gel is satisfactory for relative humidity levels below 20%. Such rooms should have special low vapor transmission treatment of walls and should be equipped with air locks.

Since it is now common practice to place each tablet press in its own separate location when in operation, the rooms can all be the same size or vary in size to accommodate the smallest and largest presses. When large volumes are being produced, for instance, it is practical to have a booth or room large enough to accommodate two or more tablet machines for the same batch (Fig. 26.15). The booth walls should extend from floor to ceiling and may be made of tile up to the four-or five-foot level, with a glass or transparent partition extending it to the ceiling. Tile or other hard surfaces in these booths should be used sparingly, since they contribute to the noise level. Space should also be provided for inprocess testing equipment such as balances and tablet hardness testers.

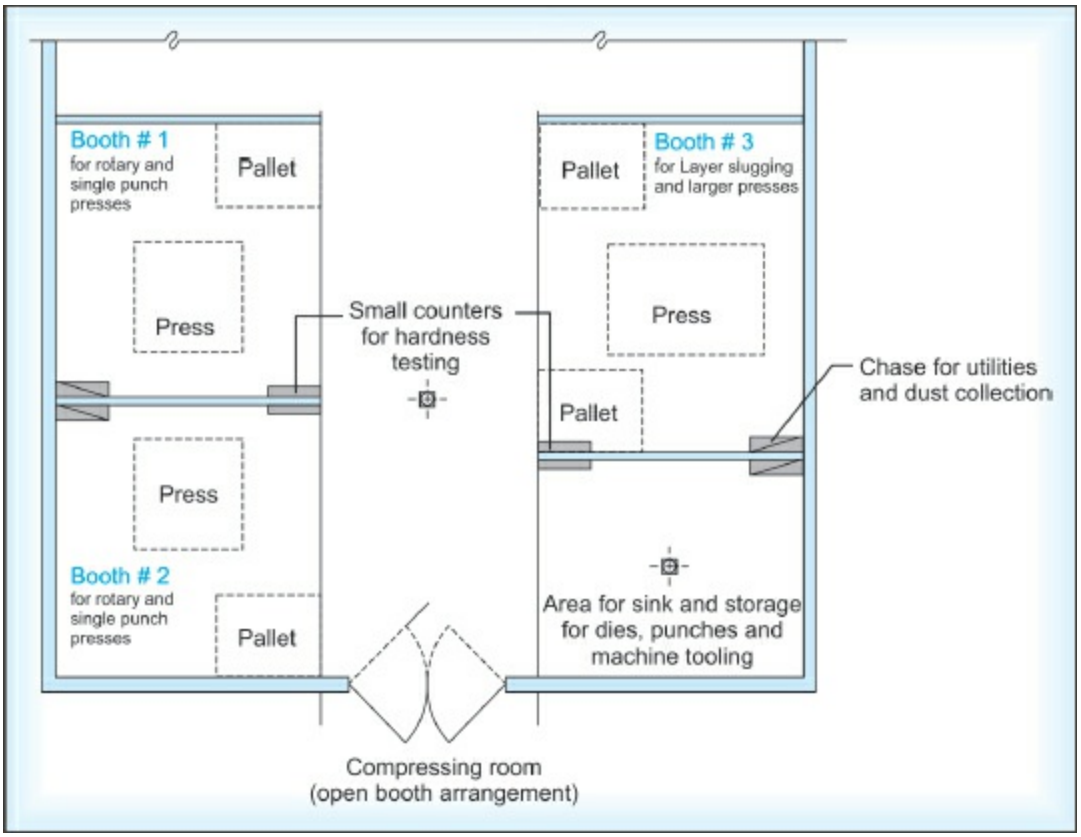



**Fig. 26.15:** Two booth sizes designed to accommodate single or small rotary presses or larger slugging or multilayer presses

## *Tablet Presses*

Each press should be mounted on metal frames so that it can be moved by lift trucks into a cleaning area. The number of booths or rooms needed in the compressing department usually does not equal the number of tablet presses on hand, since all presses are not likely to be in operation at the same time. Once a batch has been completed, the machine should be removed promptly from the booth and replaced with one that has been cleaned and prepared for the next product. A room should be made available nearby for the cleaning of presses and replacement of punches and dies for the next product.

The exact number of tablets produced is compared to the expected yield by a process called *reconciliation*. A major discrepancy between theoretic and actual yields signifies that an error may have been introduced at some stage of the procedure. To discover the discrepancy, rotary presses should be equipped with automatic counters, which can be set to place the same number of tablets in each bulk drum, thereby facilitating accountability calculations

and taking physical inventory. In this regard, commercial equipment is available that not only will count tablets but also will monitor and adjust presses to conform to weight standards (Fig. 26.16).

**Fig. 26.16:** Multicheck apparatus for tablet testing (*Courtesy* of ERWEKA)

## *Tablet Coating*

Tablet coating, which at one time consisted of sugar coating in conventional coating pans, has undergone many changes because of new developments in coating technology and changes in safety and environmental regulations. The conventional sugar coating pan has given way to perforated pans or fluidized-bed coating columns. The development of new polymeric materials has resulted in a change from aqueous sugar coating to solvent film coating, and more recently, to aqueous film coating.

Film coating is a specialized operation, and although film coating systems can be developed in a laboratory, the final coating process needs to be denned on production scale equipment. A properly designed core tablet greatly facilitates the success of scale-up. The tablets must be sufficiently hard to withstand the tumbling to which they are subjected in either the coating pan or the coating column. Tablet designs with sharp edges or flat surfaces should be avoided because they are difficult to coat. Engraved surfaces also make coating more difficult. This problem can be minimized, however, if the engravings are kept shallow and the cuts are angled to avoid sharp edges. Some tablet core materials are naturally hydrophobic, and in these cases, film

coating with an aqueous system may require special formulation of the tablet core and/or the coating solution.

A film coating solution may have been found to work well with a particular core tablet in a small laboratory coating pan or column, but may be totally unacceptable on a production scale. This difference in performance is due to the increased pressure and abrasion to which the tablets are subjected when the batch size is large, and to differences in the temperature and humidity environment to which the individual tablets are exposed during the coating and drying cycles. These differences are predictable from basic engineering equations that describe temperature and humidity gradients and transfer within the coating system. Operating conditions that must be established for either a pan or column operation are optimum tablet load, operating tablet bed temperature, drying airflow rate and temperature, and the solution application rate. The atomizing nozzles for typical pharmaceutical applications can be high-pressure airless or air-atomizing. For airless sprayers, the size and shape of the nozzle aperture is important. For air-atomized sprayers, the atomizing air pressure and the liquid flow rate are critical factors in establishing proper spraying characteristics. A high airflow yields a fine spray, but it also creates more turbulence and causes a spray drying effect. The coating solution can be sprayed continuously or intermittently. If it is sprayed intermittently, the cycles must be timed to prevent the rotation of dry uncoated or partially coated tablets, which would result in abrasion and edge chipping of the tablets and a poor quality coating.

Not only do chipping and abrasion affect those tablets damaged, but the debris formed adheres to the other tablets in the batch, thus ruining the whole pan or column load. In addition to increasing tablet hardness, the use of appropriate baffles in the coating pans can further reduce chipping and abrasion. These baffles prevent the tablet bed from sliding instead of rolling, and they also redistribute the weight of the tablet load, and spread the weight more uniformly over the entire tablet bed. In a column operation, the abrasion can be minimized by controlling the length and diameter of the center spouting column and controlling the force of the air fluidizing the tablets.

With the advent of computers and microprocessors, automated processing systems are available to make tablet coating a more controlled and reproducible process than it was several decades ago, when tablet coating was considered an art rather than a science.

## *Facility*

Traditionally, tablet coating has been considered an art and has required a rather lengthy apprenticeship. The process is noisy and dusty, and since one person operates about five pans, it is a labor-consuming and somewhat inefficient procedure. In recent years, a number of technologic developments have removed some of the “art” and substituted automated techniques that have increased pan sizes and improved the drying cycles. Automated spray coating is now available, and new products as well as old ones are being given a film coating or an abbreviated thin sugar coating applied by mechanical rather than manual methods.

These technologic changes have necessitated a new approach to the design and layout of a coating department, but some of the fundamental considerations still apply. For example, regardless of coating pan size or of whether coatings are applied manually or by spraying, the pans are placed in line and may be freestanding or enclosed (Fig. 26.17). Dust collection considerations as previously described are still important even though some new designs in pans vent the dust through the back of the pan.

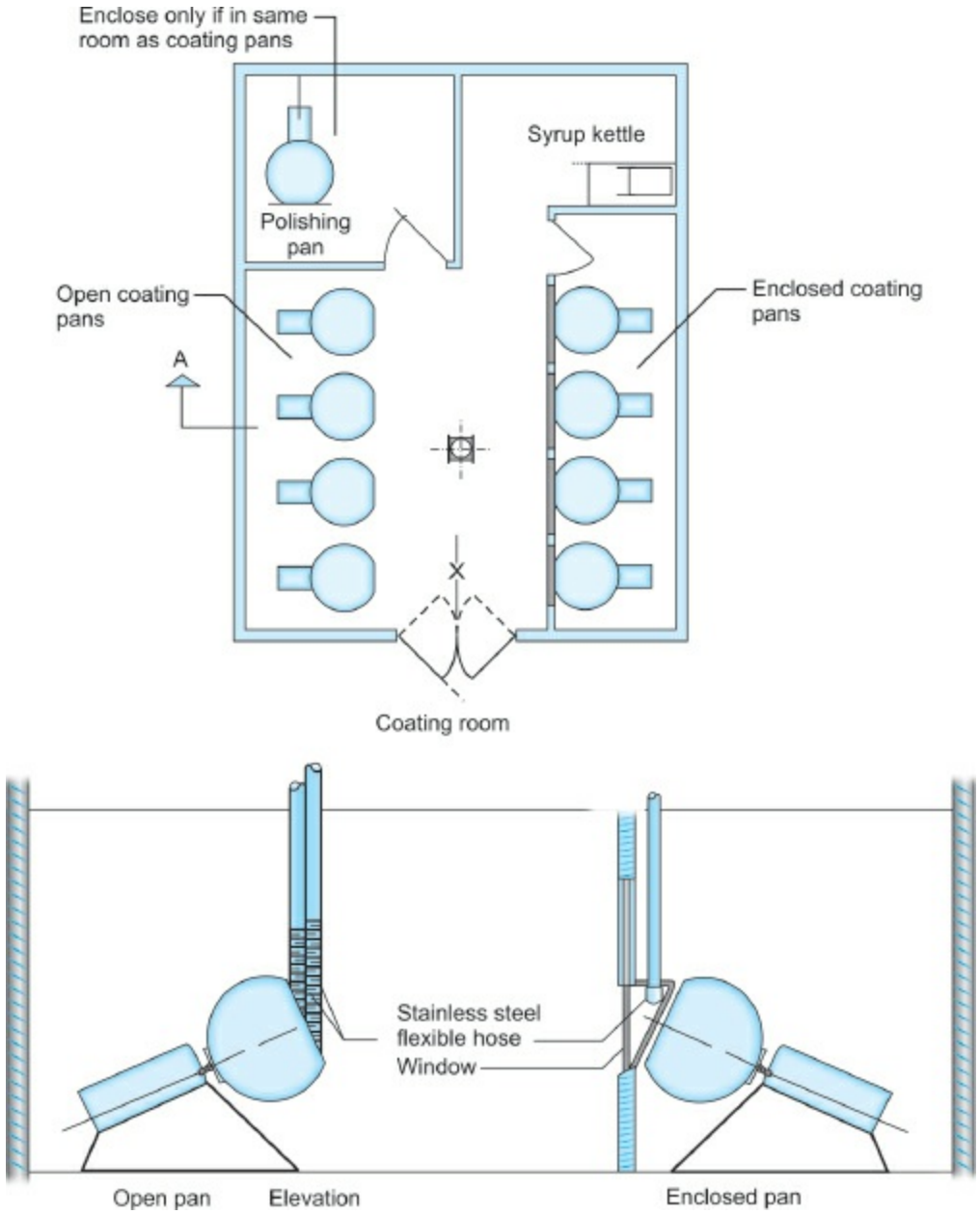



**Fig. 26.17:** Top, An enclosed bank of coating pans. Polishing pans should be in a different location from coating pans. Bottom, Differences between open and enclosed pans.

Enclosing pans in groups of five or more offers some advantages. The enclosure muffles the noise level to acceptable limits. This is particularly helpful in a large coating operation when the noise level approaches the maximum permitted under the OSHA (Occupational Safety and Health Administration) maximum average (80 decibels). The noise level can also be reduced in open pans by the use of insulating material or foam around the ******ebook converter DEMO Watermarks*******

outside of the coating pan, but product temperature control is thus rendered more difficult.

As indicated in Fig. 26.17, each pan can be equipped with a window that can be closed during dusty operations, thereby improving dust collection and reducing cross-contamination hazards.

Polishing cans of either the metal or cloth type should be isolated from the general coating operation, and any solvent-laden exhaust should be sufficiently diluted with air to meet fire and environmental standards of safety.

Adequate cleaning of floors and equipment can be a problem in coating operations because of dust and the frequent use of dyes. Sufficient floor drains should be provided for this purpose, and pumps should be used for the transfer of wash water from coating pans to either floor drains or nearby sinks. For large operations in which coating solutions made with dyes are formulated, it is desirable to have a small adjacent room equipped with a sink and mixing equipment for this purpose.

If coated tablets are imprinted with a monogram or a product identification number, each printing machine should be in a separate booth to avoid cross-contamination. If an inline, one-at-a-time printing machine is used, each machine should be equipped with an electric eye or other counting device to count tablets as they move down the discharge chute. Such devices give the official yield and can be used for product reconciliation. In addition, if it is necessary to inspect coated tablets, the inspection equipment should be placed in separate booths.

## Encapsulation for Hard Gelatin Capsules

The manufacturing process for encapsulated products often parallels that for tablets. Both tablets and capsules are produced from ingredients that may be either dry blended or wet granulated to produce a dry powder or granule mix with a uniformly dispersed active ingredient. To produce capsules on today's high-speed equipment, the processed powder blend must have the particle size distribution, bulk density, and compressibility required to promote good flow characteristics and to result in the formation of compacts of the right size and of sufficient cohesiveness to be filled into capsule shells.

Equipment used in capsule filling operations involves one of two types of filling systems. Encapsulated manufactured by Zanasi or Martelli form slugs in a *dosator* (a hollow tube with a plunger to eject the capsule plug), while the operating system of the Hofliger-Karg machines is based on formation of compacts in a die plate using tamping pins to form a compact. With both systems, the most common problems encountered in scale-up involve bulk density, powder flow, compressibility, and lubricant distribution. The common problem of weight variation can be caused by the poor flow characteristics of the granules, which may in turn be due to the bulk density and particle size distribution of the granules. Weight variation may also be caused by plugs sticking to the dosator plunger surfaces or die walls because of inadequate lubrication. Overlubrication of the granules may result in weight variation problems because the softer plugs that are formed may not be completely transferred to the capsule body.

Overly lubricated capsule granules are also responsible for delaying capsule disintegration and dissolution, which may result in reduced bioavailability. As in tablet operations, prolonged trials of many hours using multiple batches are required before a process can be judged as acceptable for routine production. The size and type of equipment used in blending, granulating, drying, sizing, and lubrication of capsule granulations or mixtures can greatly influence the characteristics of the granulation and the finished product.

Because of the differences previously noted, the type of encapsulating equipment chosen for routine production should dictate the properties required of a powder blend. During testing designed to determine the optimum process conditions, the need for controlled environmental

conditions must be considered. Many encapsulation processes are less reliable than anticipated because the humidity in the processing and encapsulation rooms is not adequately controlled. Left as an unknown variable, humidity often has a significant effect on the moisture content of the granulation and on the empty gelatin capsules. Granulation moisture content can be important to chemical or physical stability of the finished product, and uncontrolled moisture leads to machine problems of flow and sticking during material transfer and filling. Empty gelatin capsules have a recommended storage condition of 15 to 25°C and a relative humidity of between 35 and 65%. This condition is designed to minimize moisture absorption or loss, and the resultant changes in physical dimensions, during the encapsulation operation. During encapsulation, the processing room humidity should be controlled to within 45 and 55%. At higher humidities, the capsules may swell because of the moisture absorbed. This may make separation of the capsule parts more difficult and interfere with the transport of the capsule throughout the encapsulation process. Low humidity conditions make the capsules brittle and increase their static charge, thereby seriously interfering with the encapsulation operation.

## LIQUID DOSAGE FORMS

In the discussion that follows, liquid pharmaceuticals encountered in the pilot plant are defined as nonsterile solutions, suspensions, or emulsions. Scale-up of each of these presents a different set of processing concerns that must be considered. Sterile liquids represent an additional, unique class of pharmaceutical products; these are covered elsewhere within this book.

Simple solutions are the most straightforward to scale up, but then require tanks of adequate size and suitable mixing capability. Most equipment has heating/cooling capabilities to effect rapid dissolution of components of the system. Adequate transfer systems and filtration equipment are required, but they must be monitored to assure that they can clarify the product without selectively removing active or adjuvant ingredients. All equipment must be made of suitable, nonreactive, sanitary materials and be designed and constructed to facilitate easy cleaning. Liquid pharmaceutical processing tanks, kettles, pipes, mills, filter housing, and so forth are most frequently fabricated from stainless steel. Of the two types of stainless steel used in the industry (type 308 and 316), type 316 is most often used because of its less reactive nature. Stainless steel is not nonreactive, however, it does react with some acidic pharmaceutical liquids. When this situation is a concern, the problem can be minimized by prereacting the stainless steel with an acetic acid or nitric acid solution to remove the surface alkalinity of the stainless steel. This procedure, known as passivation, may need to be repeated at periodic intervals. For example, if an alkaline cleaning agent is being used between batches of a reactive product, passivation may need to be repeated before the subsequent batch can be prepared. Interaction with metallic surfaces can be minimized by use of glass or polytetrafluoroethylene (Teflon) liners. Although these are highly inert surface materials, they have the obvious disadvantages of cracking, breaking, flaking, and peeling, with resultant product contamination.

## Facilities

In locations where oral, external, or cosmetic preparations are made, it is necessary to have separate facilities for each group. If this is not possible, a separating wall should be constructed to isolate one group from another, thereby preventing cross-contamination and the transference of odors.

Special attention should be given to the design and installation of equipment and washing facilities, especially those used for products that are susceptible to microbiologic contamination. Sanitary pumps and fittings should be used, together with stainless steel tubing with snap-on connections, to facilitate easy removal and cleaning. Troughs should also be available to permit the cleaning and soaking of piping and transfer lines on an overnight basis. They should be made of materials that withstand commercially available detergents and germicidal solutions, e.g. stainless steel.

Although the use of potable water is necessary in all operations, it is particularly important in liquid manufacturing. If deionizers and other water treatment equipment are used, special attention must be given to routine microbiologic and chemical testing. Storage tanks and piping used for bulk storage of such liquids as glycerin and mineral oil should be constructed to facilitate examination as well as cleaning.

Good Manufacturing Practices requires accurate yields for liquid preparations and reconciliation of yields as in solid dosage forms. If the same tanks are used to manufacture more than one product, liquid meters and tank calibrations are important to product reconciliation. In many cases, it is practical to install on each tank load cells that provide readout of its contents.

Manufacturing tanks located on either side or around a work platform or gantry should be sufficiently far away from each other to avoid cross-contamination, especially when dry powders are an ingredient. The elevated platforms should be sufficiently large to permit the positioning of pallets or raw materials around the tanks and still provide sufficient room for the separation of adjoining tanks.

## Suspensions

Suspensions require more attention during scale-up than do simple solutions because of additional processing needs. The addition and dispersion of suspending agents, which on a laboratory scale may merely involve sprinkling the material into the liquid vortex, may require use of a vibrating feed system or other novel approach when production scale batches are involved. A powder eductor may facilitate the addition of a material that tends to clump during the process or that is difficult to disperse. In some instance, suspending agents that are difficult to disperse can be successfully incorporated by making a slurry with a portion of the vehicle. The suspending agent in a concentrated slurry is easier to wet and can be more completely dispersed using a high-shear mixer in a smaller volume of the vehicle. Such a slurry facilitates rapid and complete hydration of the suspending agent when added to the larger portion of the vehicle. The time and temperature required to hydrate suspending agents is often critical, and unless the hydration process is complete before other ingredients are added, the quality of the suspension is adversely affected.

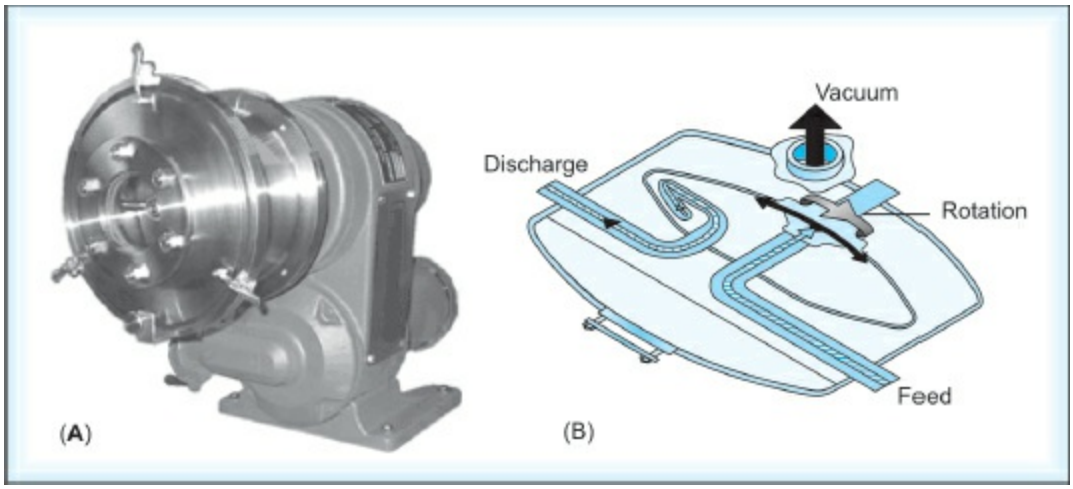
Active ingredients in a suspension must be uniformly dispersed throughout the batch. The best dispersion procedure to use in the production process depends on the physical characteristics of the active ingredients. If they wet easily, disperse readily, and tend not to agglomerate, a simple addition of the chemicals at a convenient stage in the manufacturing process is appropriate. If the active ingredients are difficult to wet or tend to agglomerate, however, other methods for adding these ingredients must be sought. One is to prepare a slurry with a wetting agent and with the aid of high-shear mixing equipment. Another method is to pretreat the hard-to-wet material by blending it in a high-shear powder blender with one or some of the liquid ingredients, possibly with a surfactant included. This converts a bulky material, which is difficult to handle because of static charges, to a dense, readily wettable powder, which is much easier to handle. Such approaches minimize wetting difficulties and eliminate the formation of dry agglomerates in the finished product. If these agglomerates should occur, the air trapped in this dry material may cause the product to “cream” (separate), thereby causing physical instability or poor content uniformity.

In preparing pharmaceutical suspensions, the type of mixers, pumps, and

mills, and the horsepower of the motors, should be carefully selected based on scale-up performance. The equipment must be selected according to the size of the batch and the maximum viscosity of the product during the manufacturing process. As an example, the use of an appropriate type of mixer is important because if the mixer is undersized, the obvious problems of inadequate distribution or excessive production time result. Mixing at too high a speed can result in the incorporation of an excessive amount of air into the product. Air that is entrapped in the product as very small bubbles is difficult and time-consuming to remove, and if not removed, can affect the physical and chemical stability of the product and/or the reproducibility of the filling operation. If air entrapment is a problem that cannot be rectified with process or equipment modifications, the air can be removed using a vacuum unit such as the Versator (Fig. 26.18). During operation of a Versator, product is drawn into a vacuum chamber through an inlet line, where it is spread onto the center of a highspeed rotating disc. The centrifugal force produced by the rotation of the disc causes the product to form a thin film on the disc surface. As the film thins and moves toward the outer edge of the vacuum chamber, the entrapped air is drawn off, and the deaerated product is collected from the outer edge of the vacuum chamber.

Unwanted and discolored particulate material in a batch can come from the raw materials or be introduced from the bags, cases, and drums in which the raw materials were supplied. Even when all precautions are taken, some unwanted material may find its way into the product during manufacturing, so that filtration of the finished suspension through an appropriate size screen is a normal batch processing step. The mesh size chosen must be capable of removing the unwanted foreign particulates but should not filter out any of the active ingredients. Such a sieve can only be selected based on production batch size trials. Most active ingredients have particle sizes less than 10 microns with almost none over 25 microns. Therefore, when dealing with particulates, screens of 150 mesh, having openings of around 100 microns, remove unwanted suspended materials that are below the easily visible range without retaining the suspended active ingredient(s).

At the completion of the batch, the transfer and filling of a finished suspension should be carefully monitored. If suspensions are not constantly mixed or recirculated during transfer processes, they may “settle out” and thereby adversely affect the uniform distribution of the active ingredient.




**Figs 26.18A and B:** (A) The Versator consists of a vacuum chamber and a high-speed revolving disc; (B) During operation, material is spread into a thin film by the centrifugal force of the disc, and deaeration is achieved under vacuum (*Courtesy* of the Cornell Machine Co)

## Emulsions

Emulsions are disperse systems similar to suspensions except that the dispersed phase is a finely divided immiscible liquid instead of a solid. The dispersed phase is usually made up of oils or waxes that may be in either a liquid or solid state. Manufacturing of liquid emulsion products entails specialized procedures, and as a result, scale-up into production equipment involves extensive process development and validation. Processing parameters and procedures that must be adjusted and controlled for the various types of emulsions include temperature, mixing equipment, homogenizing equipment, in-process or final product filters, screens, pumps and filling equipment. The degree to which the emulsion is refined by the reduction of the globule size of the internal phase affects the physical properties of the emulsion, such as appearance and viscosity, as well as the physical stability of the product. Manufacturing systems that utilize high-shear mixers are more likely to lead to air entrapment and may adversely affect the physical and chemical stability. Conversely, the use of vessels that can be operated with the contents under a controlled vacuum avoids the problem of unwanted aeration. The filtration of an emulsion to remove particulates originating from the raw materials or introduced during processing can affect the quality of the emulsion. The unwanted particulates are most efficiently removed by filtering the separate oil and water phases before emulsification.

## Semisolid Products

Pastes, gels, ointments, and creams are closely related to suspensions, liquids, and emulsions except that they are products with higher viscosities. The scale-up of these products involves many of the same factors that must be considered in the scale-up of the comparable lower viscosity products already discussed, but the high viscosity renders certain aspects of the scale-up of semisolid products more critical. As an example, the natural turbulence created by the mixers used to make liquid suspensions or emulsions is not adequate to produce a homogeneous ointment or cream. For these products, the mixing equipment must be capable of effectively and continuously moving the semisolid mass from the outside walls of the mixing kettle to the center and from the bottom to the top of the kettle. This action is required both to distribute the ingredients and to bring about a rapid and efficient heat transfer to and from the product during the heating and cooling steps.

The power required to carry out the mixing operation varies greatly during the manufacturing sequence and is directly related to changes in the viscosity of the product. Motors used to drive the mixing system of semisolid manufacturing equipment must be sized to handle the product at its most viscous stage. Motors that drive the mixers used to disperse or dissolve components that have been added early in the manufacturing sequence, when product viscosity is low, may be required to operate at a slower speed to prevent splashing of the intermediate phases. For this reason, most semisolid equipment is designed to provide variable speed mixing. An automatic manufacturing plant for the production of semisolid products such as creams and ointments is depicted in Fig. 26.19.

Many processing steps such as the mixing of oil and water phases during emulsification processing, component homogenization, addition of active ingredient, and product transfer are usually carried out at carefully predetermined temperatures. The working temperature range at which these operations are carried out are usually critical to the quality of the final product. In the formation of a cream, the aqueous phase and oil phase must be heated to a temperature above the solidification point of the oil phase, and then emulsified. Failure to have both phases at the correct temperature results in a poor-quality product with improperly dispersed wax. Consequently, an accurate understanding of the heat transfer characteristics of the system and

the temperature gradient throughout the batch is important.

Reliance on the temperature recorded by a single sensor at a fixed point in the mixing vessel is often misleading. Unacceptably wide ranges in product viscosity are frequently the result of inadequate temperature control during the critical emulsification steps. Improper temperature control can have an adverse effect on the particle size of poorly soluble active ingredients. If these are added at too high a temperature, the solubility may be artificially increased, creating a metastable product. On subsequent cooling, crystal growth or recrystallization may occur from the saturated solution. This recrystallized material may be a different polymorphic form or a different crystal type or size. The result may be a change of particle size distribution, yielding a gritty, less elegant product, or one with altered stability or biologic activity.

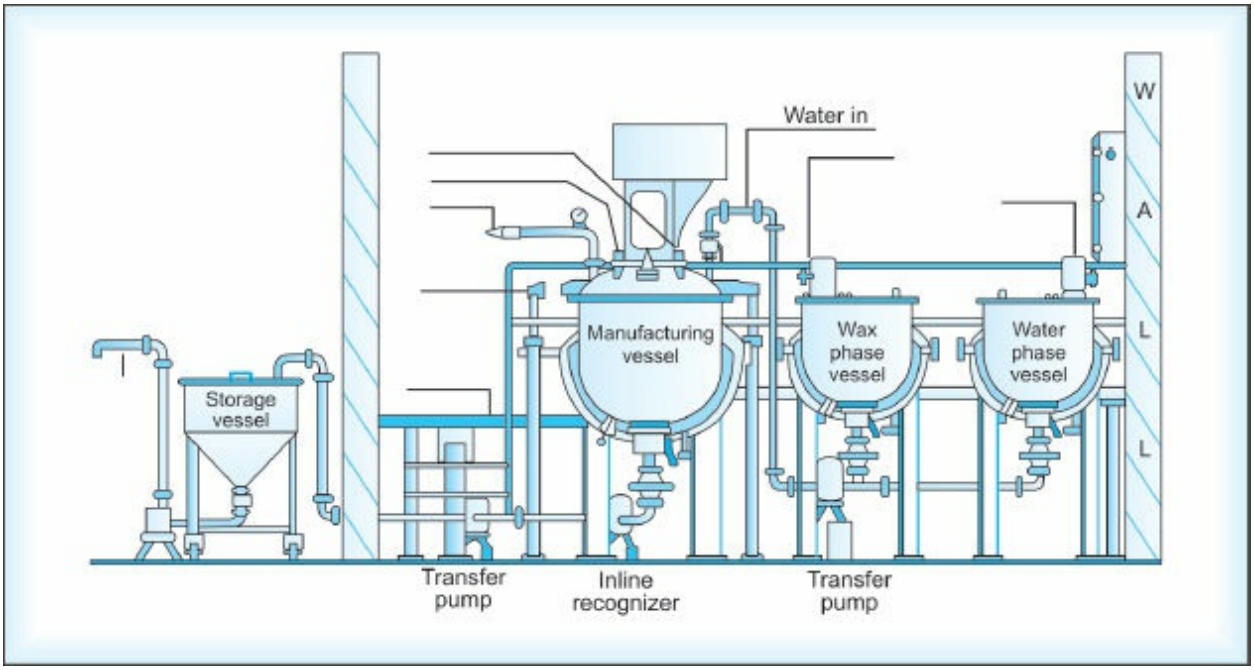



**Fig. 26.19:** Representation of an automatic manufacturing plant for the production of semisolid products

Many cream formulations and some gel products are shear-sensitive. Handling such products during transfer from the manufacturing kettle to holding tanks or to the filling lines requires that attention be given to the amount of shear that such products will encounter. Changes in measured viscosity are frequently seen when viscous products are pumped through long

transfer lines or are filtered to remove unwanted particulates. Because of this, the relationship between shear stress and the measured viscosity values of the product must be understood. When carrying out such evaluations, the pilot plant scientist needs to remember that most viscometers determine relative viscosity rather than an absolute viscosity. Therefore, the accurate evaluation of the effect of a process change on viscosity must recognize the effect of sample conditions such as temperature, processing history of the sample, and age. The more of these variables that can be controlled, the more accurate the interpretation will be of the effects of processing conditions on the viscosity of the product.

The most critical processing steps that need to be carefully evaluated and controlled during the manufacture of a cream are the emulsification of the two phases and the dispersion of any suspended active ingredients. Pharmaceutical equipment used in the homogenization of the emulsion and dispersion of suspended active ingredients includes various types of high-shear mixers, homogenizers, and colloid mills, supplied by a number of different manufacturers.

Transfer pumps for semisolid products must be able to move viscous material without applying excessive shear and without incorporating air. Pumps designed to meet these criteria are known as positive displacement pumps. They are available from many sources with subtle differences in design, but they all operate using a rotating member inside of a close fitting stationary housing. They are selfpriming and can create adequate head pressure to force product through long transfer lines and filtration equipment. In choosing the size and type of pump for a particular operation, product viscosity, desired pumping rate, product compatibility with the pump surfaces, and the pumping pressure required should be considered.

## Suppositories

The manufacture of suppositories on a laboratory scale usually involves the preparation of a molten mass, the dispersion of drug in the molten base, and the casting of the suppositories in a suitable mold. When the mold has been adequately cooled, it is opened, and the suppositories are removed. Such an operation provides little experience that is relevant to new processes used in large-scale production of a suppository product. However, many commercial suppositories are still produced by a fusion method in which (1) a molten mass is prepared, (2) the suppository is molded, cooled in a refrigeration tunnel, and removed from the mold and (3) the product is packaged in an off-line wrapping or blistering operation. A detailed discussion of this operation and processing equipment can be found in Chapter 20, “Suppositories”. The manufacturing and packaging processes for suppositories have recently been simplified to a one-stage operation. This new technology eliminates many of the troublesome molding, cooling, and unmolding steps of the older technology. The basic improvement of the newer processing equipment is that the molten suppository mass is filled into formed PVC or foil shells, which serve both as the mold and finished package. Such a process eliminates many of the problems encountered during the removal of the suppository from the two-piece molds in which they were formed on the older equipment. The extra work and equipment required to complete the off-line packaging operation of wrapping or blistering are also eliminated.

The manufacture of suppositories using modern equipment can be divided into several operations involving first the manufacture of the molten suppository mass and then the molding and packaging of the suppository.

### *Preparation of the Molten Suppository Mass*

The preparation of the suppository mass on a production scale involves heating various wax-like components of the suppository base to a temperature at which they become molten. The higher melting components should be placed in the manufacturing kettle first and the lower melting ingredients added when the first components are almost completely in the molten state. To avoid overheating the waxes and altering their melting points, this operation should be carried out in jacketed vessels in which the jacket temperature can be controlled. The best systems allow monitoring of

both the temperature of the vessel jacket and the molten contents of the vessel. Systems in which only live steam or cold water are available to regulate the temperature of the molten suppository mass are almost impossible to control within the fairly narrow temperature ranges required to produce a continuous source of well-controlled heating and cooling water. The normal operating range is 35 to 65°C, with temperatures of 45 to 65°C used for melting and mixing of the suppository base material and temperatures of 35 to 45°C used during addition of active ingredients and during molding (filling). Since the viscosity of suppository masses is normally temperature-dependent, the filling operation should be conducted at a temperature just above the solidification point. Settling of suspended material during cooling is prevented through both the increased viscosity and the reduced time required to solidify the mass. Because this is the most critical step, the working range of temperature is often no larger than ± 3°C around the set point.

The precise temperature control required affects the design and type of mixing system used in the melting and holding tank. The vessel in which the mass is prepared may require high-shear mixing capability to break up agglomerates of the active ingredient and to disperse the active ingredient effectively throughout the molten base. The equipment should also be capable of providing gentle stirring action that will effectively keep the active ingredients well dispersed.

Transfer of the molten mass to the filling heads, of the suppository machine must be done through heated lines. When a separate holding tank is used, this must be jacketed and should be part of a recirculating loop designed to prevent settling of the active ingredient and congealing of the mass.

Any extraneous particulate material that may be present in the molten mass can be removed using an in-line filter of an appropriate mesh size. Most active ingredients in suppositories are less than 10 microns in size, so 100-mesh filters with openings of approximately 70 microns provide filtration without retention of the active ingredients. The filter’s ability to remove extraneous material but not hold back any of the drug should be validated by a series of carefully monitored pilot plant trials. Extensive sampling and analysis of suppositories over an extended time period designed to mimic production conditions of filtration, holding, recirculation, and filling are

required to show that the active ingredient is not retained. In addition, extensive visual inspection of filtered suppository mass is required to show that the filter is efficient in removing unwanted material.

## *Molding and Packaging*

Current suppository manufacturing technology utilizes equipment that forms, fills, and seals the suppositories in a continuous process. The mold for the suppositories is formed from special thermoplastics sheets of PVC, polyethylene, and aluminum foil laminates. A variety of problems can occur during the forming of the shells. These can be minimized by careful attention to the temperature and dwell times used during the formation of the molded cavities. Inappropriate temperatures or dwell times that are either too short or too long produce inadequate mold shells that may leak because of improper seals. The temperature of the mass during filling is also important. If the molten mass is filled at a temperature more than a few degrees above the congealing point, a hole forms in the center of the suppository upon cooling, owing to excessive contraction. Filling at too low a temperature causes clogging of the transfer lines and filling nozzles and results in erratic fill weights. Therefore, the temperature of the molten mass in the hopper and lines must be carefully controlled, and the mass constantly stirred or recirculated, to maintain the mass within a narrow temperature range.

After filling, the shells pass through a cooling tower where the suppositories are allowed to solidify before the ends of the PVC or aluminum shell are sealed. Then the strips of suppositories are trimmed and cut into the specified length. The temperature in the cooling tower determines the cooling rate of the suppositories. If the temperature is too low and the cooling rate too fast, the suppositories can become brittle. Sufficient scale-up trials should be run to optimize the process so that the temperature of the product throughout the process is maintained at the level required for the particular formulation.

## PACKAGING

Packaging lines should be far enough apart to prevent cross-contamination, product mixup, or other serious problems. Normally, a separation of 15 to 20 feet is adequate, and in some operations, a wall or movable partition between packaging lines has been used.

The choice of straight lines or U-shaped lines can be made only on the basis of department layout or line speeds. If U-shaped lines are selected, essential materials-handling activities that are not related to actual filling or labeling, for both the input and output sides of the U, should be performed outside the department (Fig. 26.20). This arrangement not only eliminates noise, heavy traffic, and paper dust, but also helps to reduce air-conditioning loads in the packaging department.

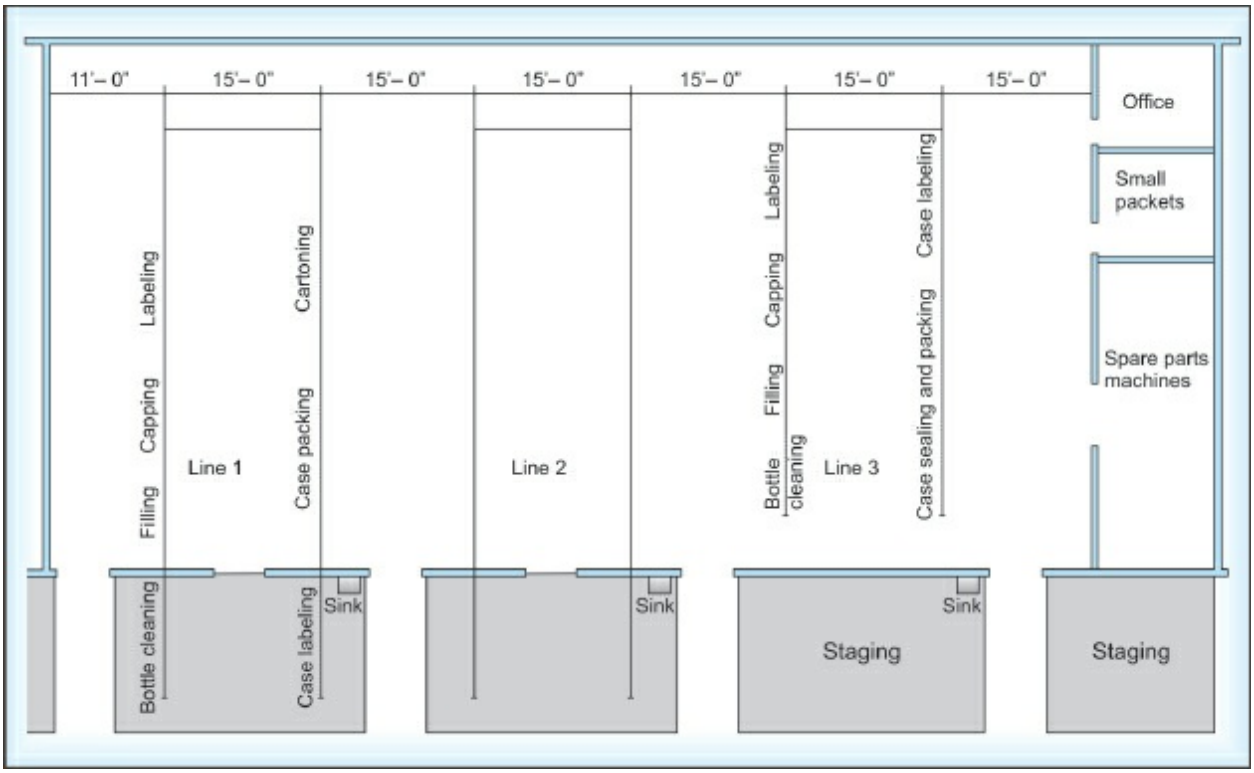



**Fig. 26.20:** U-shaped lines with bottle filling and case sealing external to department. Distance between lines is 15 ft, and a divider can be used for separation when needed

Straight packaging lines, on the other hand, do not lend themselves to such an arrangement, since the output end of the line must be provided with

shipping cases and case-sealing machinery, and all finished products must be carried out of the room with lift trucks.

For operations in which there are considerable numbers of labels of the same size or color, the concept of roll labeling equipment versus cut labels should be explored carefully. Roll labels have many advantages over cut labels in avoiding mislabeling or label mixups in either the label-printing operations or the storage and handling of labels in packaging. As the name implies, when the continuously printed sheet of labels comes off the printing press, it is slit, and each row of labels is wound into a roll around a center spool. Each label also has an identifying mark on it that not only provides correct label identification, but also allows electronic counting, thereby permitting good label reconciliation.

If cut labels must be used, label-scanning equipment should be provided for label identification either before the labeling operation, at the labeling machine, or at both times.

Labels should be stored in an air-conditioned room, and in winter, the air should be humidified to maintain a relative humidity of about 50% to avoid overdrying of labels. The room should have sufficient space for storage of inserts and be subdivided to separate approved from unapproved labels and inserts in accordance with GMP.

The department should be equipped with an adequate number of sinks to permit washing of measuring utensils such as graduates and packaging line equipment parts. Usually, one sink is sufficient for two lines.

The cleaning of dry products filling equipment by vacuum or compressed air should be conducted during shutdown hours when the danger of cross-contamination is minimized. Whenever compressed air is used for cleaning, the equipment should be removed from the line if possible or completely enshrouded during the cleaning operation.

Space should be available for use by departmental and/or quality control line inspectors for the storage of counting boards, graduated cylinders, torque testers, and other in-process testing equipment; also, cabinets should be provided for clean utensils and parts.

A staging area should be available for the storage of packaging equipment not in use and machine change parts, and facilities for cleaning and dismantling packaging equipment should also be provided.

## Warehousing

Since warehousing is normally the largest operation in the plant in terms of area, special attention should be focused on maintaining cleanliness, freedom from infestation, and orderliness. The entire warehousing area should be cleaned as often as necessary to maintain sanitary conditions. Mechanical floor washers may be used in large facilities. From time to time, wooden pallets are subject to spillage of materials, and an area in which to clean pallets should be provided. Occasional pesticide treatment of pallets is also advisable to minimize insect infestation.

A quarantine area for incoming raw materials and packaging components is necessary, and an enclosed quarantine area must be provided for raw materials, packaging components, bulk products, and finished goods that have been rejected for failure to meet various standards.

## Shipping and Receiving

Constant movement of materials in and out of the building subjects this area to the greatest possibility of insect and rodent infestation. This is particularly troublesome in tropical areas and when night operations are in progress. Air curtains have been used to prevent flying insects from coming into the building, but their effectiveness is somewhat limited.

When an inside dock is provided, it should be large enough to permit both the trailer and the tractor to park inside the building. Overhead doors can be used to close off the dock area. Each opening at the loading platform where trucks back into an outside wall should be equipped with compressible receptacles that effectively seal the truck's opening with the entry port into the building.

Only approved finished goods should be kept in the shipping area. Items awaiting quality control approval should be kept in the quarantine area. If this is not possible, a system that clearly identifies approved finished goods in the warehouse must be used.

Alcohol and other combustible solvents should be stored in explosion-proof rooms equipped with special fire protection facilities. If narcotics and other dangerous drugs are handled, vaults approved by the Drug Enforcement Administration (DEA) must be used for storage of finished goods as well as for in-process materials. It is advisable to have the alarm signal connected directly with the nearest police station.

An inspection center immediately adjacent to the receiving dock should be provided where facilities for the examination of incoming materials are made available. This can be used by the quality control department for statistical inspection and the sampling of raw materials and finishing supplies. The area should have lighting of not less than 150 footcandles. The inspection center should also be equipped with sinks and other facilities for washing test equipment, and space should be provided for storage of retained samples for quality control as well as permanent production records.

Dust collection hoses should be provided to clean the tops of containers prior to placement in the general warehouse.

## CONTRACT MANUFACTURE

On occasion, scale-up or manufacture of a product may need to be done at an outside contract manufacturer. The reasons for considering contract manufacture include the needs for additional manufacturing capacity, highly specialized technology, or specialized equipment. When choosing a contractor, consideration should be given to the experience, capability, and reputation of the contractor within the industry. In addition, the technical competence of the contractor's personnel and compliance to good manufacturing practices are important factors in the choice of a vendor. Before pilot-size batches or larger are prepared at an outside facility, a team of personnel who are knowledgeable and experienced in production, QA, QC, and GMP should visit the potential contract manufacturer. The site visit should include an in-depth assessment of facilities, equipment, personnel, and policy. The following checklist of items to be reviewed originates from the Good Manufacturing Practices regulations as set forth in the Code of Federal Regulations (21 CFR211):

- Review of procedures, facilities, and personnel (training and documentation) involved in acquisition, storage, testing, and handling of raw materials (excipients and actives).
- Equipment availability, utilization, and maintenance; documentation of equipment operator training, equipment cleaning procedures, and equipment calibration.
- Adequacy of manufacturing facilities (size, spatial separation, cleanliness, design, cleanability; dust, temperature, and humidity control).
- Adequacy of storage space for in-process, quarantine, and approved bulk finished goods storage.
- Controls for preventing product mixup and cross-contamination.
- Label storage, approval, dispensing, and accountability procedures and controls.
- Packaging equipment utilization, maintenance, cleaning, and use controls.
- Packaging component control procedures.
- Structure and function of the quality control unit, including personnel, training, SOPs, and responsibilities.

- Warehouse space, separation of products, control procedures, and environmental controls.
- Record control policies.

After all of these checklist items have been discussed to everyone's mutual satisfaction, the visiting team can complete discussions of division of responsibilities between the originating company and the contract vendor. These discussions should be finalized in a comprehensive document that outlines all areas of responsibility and that can serve as the basis for follow up investigations by the pilot plant scientist and/or the quality control department.

Technical fact-finding missions of this type can underscore the unique characteristics of the pilot plant scientist, whose scientific expertise, knowledge of production equipment and processes, awareness of GMPs and SOPs, and ability to critically evaluate facilities, personnel, and procedures are key components to the direction of subsequent negotiations.

# MATERIALS MANAGEMENT

The role of materials management is to convert the sales forecast into a production forecast and then into raw materials, finishing supplies, intermediates, equipment loading, and labor hours. It is then necessary to see that all of these are available at the right time and place to maximize the use of company assets and provide the best customer service with the lowest inventory investment. To do this effectively, there must be control over purchasing, receiving, shipping, warehousing, and distribution as well as production planning and scheduling, and continuous liaison is needed with production, marketing, sales, and quality control departments.

Materials management has the dual responsibility of determining the amount of inventory as well as its accountability. Inventories are an important part of a company's working capital and are reported to stockholders in the annual report.

Activities most identified with materials management are production planning and inventory control. Production and inventory control principles were developed from statistical or mathematical approaches. More recently, the evolution of operations research has permitted many production and inventory control problems to be expressed mathematically, and statistical probability theories have provided new methods for solving complex business problems by means of a computer. In this regard, the selection of computer hardware and the development or purchase of software systems are important responsibilities of the materials management department to ensure that any management information system provides timely and accurate data on thousands of daily transactions occurring in manufacturing and related operations.

## Inventories

Basically, inventories are needed to satisfy future demands, and the pharmaceutical industry has relatively long process cycle times and procurement lead times. The inventories may be described as a combination of fluctuations in anticipation, lot size inventories, and inventories to cover movement of materials from one location to another. All of these are affected by fluctuations in demand and manufacturing lead times, which are covered by reserve stock or safety stock (*see* "Concepts of the Order Point System," later in this chapter). Inventories are classified as follows:

1. *Materials:* These are such chemicals as active ingredients, diluents, and excipients needed to manufacture intermediates or components of the finished product. Included in this category and best shown separately are finishing supplies such as containers, labels, caps, and shippers needed in the packaging operation.
2. *Components:* These are parts or subassemblies needed for the final assembly of the end product (e.g. bulk tablets awaiting packaging).
3. *Work-in-process:* These are materials and components on which work is being done.
4. *Finished goods:* These are the salable items, samples, or other promotional items held in inventory awaiting customer orders or made for specific customers.

### *Inventory Management*

It is customary in any production operation to consider return on investment in buying capital equipment, and many appropriation requests are turned down if the rate of return is too low. Commitments for inventories must be considered in the same way, and obviously, the purchase and holding of a one-month supply of an item gives a better return on investment and inventory turnover than a two-months supply. This is an oversimplification since there are many costs associated with inventory decisions, for example, ordering costs, out-of-stock costs, clerical costs, computer costs, and quality control costs; others are too numerous to list here. Examination of the annual reports of several top pharmaceutical companies that have the greatest return on equity shows that inventories can represent anywhere from 35 to 80% of working capital, and some of these have worldwide inventories approaching

$700 million! Not only does it cost money to acquire inventory, it also costs money to hold it until used.

A generally accepted method of quantifying inventory costs is as follows:

| Inventory carrying cost breakdown (%) | |
|---|---|
| Cost of money | 11.0* |
| Storage and handling | 8.5 |
| Obsolescence | 4.4 |
| Taxes | 1.0 |
| Insurance | 0.1 |
| | 25.0 |

* Currently about 15.0%

A practical illustration of the importance of this quantification is as follows:

Average worldwide inventories = $ 280,000,000
Inventory carrying cost at 25% = $70,000,000/yr
A 10% reduction in investment would save = $ 7,000,000.
Inventory reduction would release in cash = $ 28,000,000.

Obviously, a well-managed inventory can exert considerable financial leverage, and inventory reduction can release much needed cash which the corporation can invest in more profitable ventures and reduce borrowing. The inventory investment, however, must always take into consideration its effect on out-of-stock situations as well as the desired customer service level.

## *The ABC Concept*

One of the most important and simplest tools used for inventory management is the ABC classification of inventories. This classification is based on a principle first outlined in the late 1800s by V Pareto, an Italian engineer and mathematician. In its simplest terms, it states that in a large population in which many items are involved, relatively few items account for the major part of activity. For example, 15% of the highest-value items in inventory amounts to 70% or more of the total inventory value. Another 25% of medium-value items accounts for an additional 20% of inventory value. 

Therefore, the combined A and B values, which amount to only 40% of the total items, nevertheless represent a combined value of 90% of the total inventory. The remaining 60% of items represents a small value. These are classified as A, B and C, respectively. This means that when there are a substantial number of items to be controlled, emphasis should be given to A and B items, since on the average, they constitute the major portion of total inventory value. By concentrating attention on the management of A and B items, one is in effect covering about 90% of the inventory value. Inventory levels of C items should be given little attention and can even be kept at a high level since they contribute only a small percentage to the raising or lowering of inventories.

Figure 26.21 shows that 40% of all inventory items fall on the curve at 90% of the total inventory value, and other relationships can easily be calculated from the curve. Each item, such as raw materials, can be expressed as the total annual procurement value in descending order for easy classification.

| Items | Combined Annual Value | % of Total Annual Value | Cumulative (%) |
|---|---|---|---|
| 1–4 | $1,800,000 | 36 | 36 |
| 5–6 | 600,000 | 12 | 48 |
| 7–30 | 500,000 | 10 | 58 |
| 31–40 | 1,600,000 | 32 | 90 |
| 41–200 | 500,000 | 10 | 100 |
| | $5,000,000 | | |

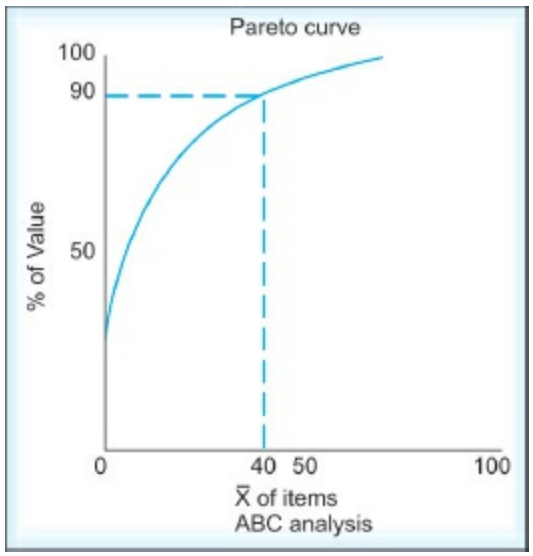


**Fig. 26.21:** Pareto curve. Graph shows that 40% of items can account for 90% of total value.

This table shows that the first 30 items (15% of 200) represent 58% of the total annual value. The next group of 31 through 40 items represents an additional 32%. This means that in this particular illustration, by closely following 30 raw materials that could be classified as A items, 58% of the total annual value would be covered. By examining 10 items that can be given a B classification, an additional 32% would be covered. Therefore, by closely watching 40 items out of 200, there would be a coverage of 90% of the total annual value. Obviously, little attention needs to be given to items 41 through 200, since they represent only $500,000 and 10% of total annual purchases.

## *Inventory Reporting and Analysis*

Since inventories are reported in dollars, the figures are of little value unless they are related to something, such as the ratio to cost of goods. A commonly used expression is "Turnover" (TO), or "Stock Turn Rate" (STR), which shows the relationship of inventories to the amount of goods that could be produced from those figures when related to factory door cost on an annual basis:

$$\text{TO or STR} = \frac{\text{COST OF GOODS IN PERIOD}}{\text{COST OF INVENTORY ON HAND}}$$

If the annual cost of goods is $20,000,000 and the inventory is $10,000,000, the STR is 2.0, meaning that the inventory "turns over" twice a year. Such information is generally given to the top people in management to provide a general inventory view, but at best, it shows an incomplete picture.

In addition, a STR of 2.0 for one division of a company or one group of products may be completely unacceptable to another division or group, which may be producing large-volume consumer products and may expect inventories to be turned over more often. What the STR of 2.0 is really showing is that there is a 6 months total inventory on hand.

Inasmuch as inventories are on hand to take care of future sales demands, it is not accurate to relate today's inventories to a previous cost of goods. An expression of inventory dollars as they relate to forecasted sales would be

more realistic.

To control and analyze inventories properly, it is best to develop a relationship to inventory dollars as well as to the length of time such inventory will last, expressed in number of months' supply. If it is established that a 10 million dollar inventory is too high, this figure by itself does not permit proper inventory analysis to determine in which categories the high inventory occurs. This can be done as follows.

## *Conversion of General Inventory Dollars to Months of Supply*

The cumulative monthly forecasted cost of goods sold, which equals the inventory dollars, gives the equivalency in months of supply:

*Example:* If the total inventory is $10 million and finished goods total $6 million, what are the equivalents in a month's supply?

| Forecasted Cost of Goods Sold (millions) | | Cumulative |
|---|---|---|
| January | 2 | 2 |
| February | 2 | 4 |
| March | 3 | 7 |
| April | 2 | 9 |
| May | 2.5 | 11.5 |
| June | 3.1 | 14.6 |

By subtracting finished goods from the total inventory, the remaining balance must be supplies and work-in-process. Since the $10 million inventory falls between April and May, by interpolation, this works out to be 4.4 months.

| Total Inventory | Finished goods | Supplies and work-in-process |
|---|---|---|
| $10 million | $6 million | $4 million |
| 4.4 mo | 2.64 mo | 1.76 mo |
| | 6/10 4.4 | 4/10 4.4 |

Carried to its logical conclusion, such a report should be broken down

further into individual products or product groups representing A or B items to ensure that inventories that are not in line or are above inventory policy limits will clearly show, so that investigation and further action can be taken (Table 26.4). Such a report issued monthly provides meaningful information to all managerial personnel in materials management, production, sales, and accounting, since it covers A and B items and therefore most of the inventory investment.

**Table 26.4:** Inventory and Months of Supply by Product Classification

| Description by classification | Raw materials | | Finished supplies | | In Process | | Bulk | | Finished goods | |
|---|---|---|---|---|---|---|---|---|---|---|
| | M$* | Mos | M$ | Mos | M$ | Mos | M$ | Mos | M$ | Mos |
| Product-A | 73.0 | 2.2 | 81.4 | 2.7 | 11.5 | 0.2 | 48.7 | 1.1 | 239.8 | 2.3 |
| Product-B | 10.8 | 1.6 | 6.2 | 1.8 | 14.2 | 1.3 | 24.0 | 2.1 | 40.0 | 2.9 |
| Product-B | 12.4 | 0.9 | 4.8 | 3.2 | 25.6 | 2.4 | 19.3 | 1.5 | 31.0 | 2.2 |

*$ in thousands

## Sales Forecasting

The importance of a well-planned, well-executed sales forecast deserves emphasis, because therein lies the basis for many business decisions. A sales forecast dictates future personnel, equipment, and warehousing requirements. It also generates the inventory investment plan and determines the amount of cash needed to operate the business. At the same time, purchasing personnel are apprised of the amounts needed, so that they can arrive at the best prices and delivery dates.

### *Responsibility for Sales Forecasts*

With few exceptions, sales forecasts are made by the marketing staff, and this perhaps is as it should be. They should know the conditions in the marketplace and be able to assess the effects of competition, advertising and promotion, changes in prices, and the size of the sales force in view of fluctuating demands. Poor forecasting can have serious ill effects on production operations. The production staff, on the other hand, continually calls attention to violent changes in sales demands as they affect operations. Sometimes forecast inaccuracies are offered as an excuse for problems in production management. When inventories are too high, the forecast often is blamed for not meeting the plan, and people are either laid off, or worse, kept on and underutilized, thereby increasing costs and reducing labor efficiency. Sometimes the problems are due to forecast inaccuracies and sometimes to errors in production planning. Both groups should get together to resolve significant differences.

Since sales forecasts have an effect on so many other operations, those operations most affected should be represented in the forecasting procedure. The materials management staff holds a pivotal position in this regard, inasmuch as it must be alert to significant variations between forecasted demands and actual sales.

### *Techniques*

Basically it is more accurate to forecast for a large group of items than for any one item. It is easier, for example, to forecast the total sales of all the thyroid products in the line than to predict how many bottles of 65 mg 100s will be sold. Obviously, a forecast of tomorrow’s sales can be more accurate

than a prediction of sales two years hence. Additionally, forecasting should not be thought of in absolute terms, and every forecast should include an estimate of error. Before the forecast system is used, the method should be tested, especially in statistical forecasting, in which trends, seasonality, and randomness must be taken into consideration.

## *Averaging*

**Simple averaging:** The simplest approach is to assume that demand is steady and that sales for any new period will be the same as for the current period. Simple averaging, then, is not realistic; it treats all data equally, since there is no way of emphasizing or "weighting" some portion of the data.

**Weighted averaging:** Arithmetic averages "weight" all data in terms of previous sales, which can be called the old forecast, and since inventories are for future sales, no means is provided for "weighting" future requirements. If the average weekly sales for last year were 100 pieces and the first week's sales for the new year were 70, it would be unreasonable to "weight" these numbers equally in forecasting future sales. If a "weighting factor" of 1 was used (same as 100%), it would be logical to recognize more "weight" in the 52 weeks of last year's data than in one week of the new data in computing the new forecast. For instance, consider the following:

**The weighted average**

**First week**

| | | Simple average | Weighted average |
|---|---|---|---|
| Old forecast (Avg/50 wks) | = | 100 × 0.5 = 50 | × 0.9 = 90 |
| Sales (Latest wk) | = | 70 × 0.5 = 35<br>170 Avg 85 | × 0.1 = 7<br>Avg 97 |

**First week**
**(Using weighted average)**

Old forecast = 97 × 0.9 = 87
Sales (2nd wk) = 105 × 0.1 = 11
New forecast = 98

In the foregoing example, the average weekly sales for the past 52 weeks (old forecast) plus the latest sales total 170, the average being 85. When an equal weight of 0.5 (50%) is given to each number, the result is an average of 85. Why should the average of 52 weeks of data be given the same weighting as data for one week? This is not a good method.

In the same illustration, one may wish to give the 52 weeks (old forecast) a value of 0.9 (90%) and the new one a value of 0.1 (10%): the results would produce a new forecast of 97 instead of 85. This in turn can be developed into forecasts for the ensuing weeks, a system well suited to computer application.

## *Exponential Smoothing-first-order Equation*

From the foregoing, it can be seen that it is possible to assign different weighting values to averages, and this is particularly useful on moving averages when dealing with averages of 3 months, 6 months, 39 months, and so on. In exponential smoothing, such values range from 0 to 1 using a, which is called a *smoothing constant*. Thus, to produce a new forecast, the old forecast, the current demand, and a selected value of are needed. This can be expressed as follows and is commonly referred to as the first-order equation.

New forecast = (1 – α) old forecast + α new demand. Accordingly, if the old forecast is 100, current demand 120, and an assigned α 0.2, then:

(1) New forecast = (1 – 0.2) 100 + 0.2 (120)
= 80 + 24
= 104

(2) If α is changed to 0.5

New forecast = (1 – 0.5) 100 + 0.5 (120)
= 50 + 60
= 110

In (1), the α of the old forecast has a value of 80, and in (2), when is changed to 0.5, the value drops to 50. Thus, it can be seen that as the value increases, the old average of 100 is progressively discounted and has less weighting, whereas the new data give more weighting since it moves from 24 to 60.

Forecast accuracy is best when data of 2 or 3 years are available; however, exponential smoothing can be used to advantage in forecasting for new products for which little prior sales history exists. As was seen before, the higher the α value, the less weighting there is on the old forecast, so that on new products high α values can be used to limited advantage. The following table shows the equivalency of assigned α values to months of data.

| Alpha values vs. Equivalent moving average | |
|---|---|
| **Alpha** | **Equivalent months** |
| 0.500 | 3 |
| 0.400 | 4 |
| 0.333 | 5 |
| 0.250 | 7 |
| 0.200 | 9 |
| 0.100 | 19 |
| 0.050 | 39 |
| 0.010 | 199 |

As stated before, in any forecasting system the forecast error should always be measured by plus or minus differences between actual versus forecasted sales.

## *Mean Absolute Deviation (MAD)*

As a measure of forecast error, the monthly differences between actual and forecasted sales should be noted and expressed as a plus or minus value from the actual. These monthly deviations are then added without regard to the sign and divided by the time period being measured. This is called mean absolute deviation (MAD). For example:

| (Week) | Time period Sales | Deviation |
|---|---|---|
| 1 | 1200 | –300 |
| 2 | 1500 | — |
| 3 | 1900 | + 400 |
| 4 | 1700 | + 200 |
| 5 | 1600 | + 100 |

$$\text{MAD} = \frac{1000}{5} = 200 \text{ Units}$$

The error can also be expressed statistically by calculating the standard deviation, but this is time-consuming and requires much computation. There 

is an approximate relationship that is expressed, however:

$$\text{Sigma} = 1.25 \times \text{MAD}$$

which would give 250 units in the example given above.

If the MAD is watched, the forecast error is monitored frequently. This monitoring prevents overproduction or underproduction and can be used as a "tracking signal" to report significant variations, which in turn would result in production or forecast changes.

## *Trends and Seasonality in Forecasts*

In the event of a definite sales trend, the use of the first-order equation produces forecast error differences, which on a cumulative basis gets larger and larger. In such cases, the method of "least squares" can be used; however, it is sometimes easier to use second-order smoothing.

## Economic Lot Size or Order Quantity (EOQ)

It is not uncommon to have tens of thousands of items in inventory, so that materials management personnel must decide how much to buy and when each item should be delivered, as well as when intermediates and finished products should be made. With excessively high inventories, too much cash is tied up, investment opportunities are lost, and obsolescence may be increased. Production and marketing staffs tend to favor high inventories to get longer production runs, avoid back orders, and provide the best customer service levels.

Fixed costs such as set-up and clean-up times are the same no matter how large or small the batch size or the packaging run is. Obviously, the more frequently an item is bought, the lower the carrying cost is, but such costs as quality control increases if something is purchased every month, as opposed to every third month.

Using simple figures for ease in calculation, the following trial and error method shows the EOQ.

*Total annual cost for various order quantities assume*

1. Annual usage of 12,000 units
2. Unit cost of $3.00
3. Cost of carrying inventory 10% per year
4. Ordering costs per order $50.

| **I. Assume a delivery of 500 kg** | **Cost** |
|---|---|
| Annual fixed costs<br>24 orders/yr $50) | $1,200 |
| Annual cost of carrying inventory<br>$\frac{500^*}{2} \times 3 \times 10\%$ | 75 |
| Total annual costs | **1,275** |
| **II. Assume a delivery quantity of 1,000 kg** | |
| Annual fixed costs | 600 |
| Annual cost of carrying inventory | 150 |
| Total annual costs | **750** |

| | |
|---|---|
| **III. Assume a delivery quantity of 2,000 kg** | |
| Annual fixed costs | 300 |
| Annual cost of carrying inventory | 300 |
| Total annual costs | **600** |
| **IV. Assume a delivery quantity of 3,000 kg** | |
| Annual fixed costs | 200 |
| Annual cost of carrying inventory | 450 |
| Total annual costs | **650** |
| **V. Assume a delivery quantity of 4,000 kg** | |
| Annual fixed costs | 150 |
| Annual cost of carrying inventory | 600 |
| Total annual costs | **750** |

*Average inventory is half the lot size.

Therefore, when a 2,000 kg delivery quantity is ordered, the annual fixed costs and the annual carrying costs are equal, resulting in the lowest total annual costs. The interrelationship of the costs are illustrated in Fig. 26.22, which shows that as order quantities go up, so do carrying costs, but ordering costs go down. The total cost curve shows that total costs decrease with increasing ordering quantities to a minimum of a $200 ordering quantity, after which total costs increase with larger lot sizes because of the increasing effect of inventory carrying costs. In this case, a $200 order quantity has a total cost of $20, but if a $50 order quantity were used, the total cost would be twice as much or $40.

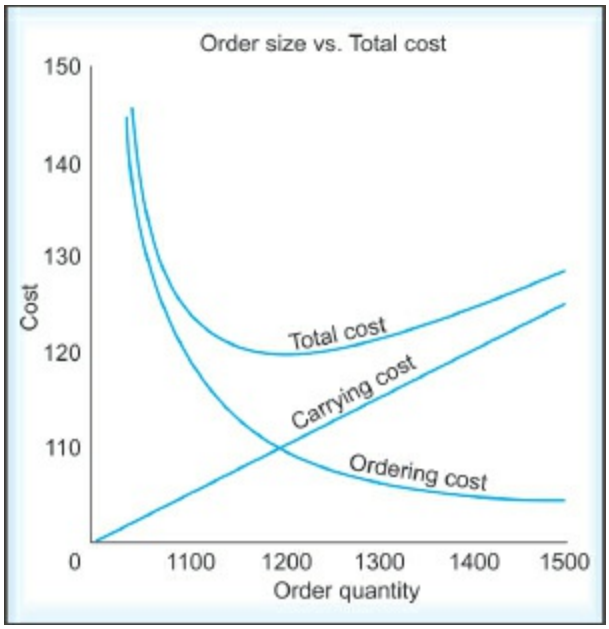



**Fig. 26.22:** Graph shows relationship of changes in ordering costs and inventory carrying cost to total costs

## *The EOQ Formula*

Considering the thousands of inventory items, a trial and error method would be impractical; in its place, the EOQ equation can be used.

$$EOQ = \sqrt{\frac{2\,AS}{I}}$$

where A is the annual usage in dollars, S is the ordering costs in dollars, and I is the inventory carrying costs expressed as a decimal.

Since the foregoing equation shows the least cost expressed in dollars, the least cost can be calculated directly as follows by substituting the figures used in the trial and error illustration.

$$EOQ = \sqrt{\frac{2\,(12000 \times 3) \times 50}{0.10}} = \$600$$

However, using the same example, if the number of units that give the lowest cost is required, the equation is expressed as follows:

$$EOQ = \sqrt{\frac{2\,AS}{IU}}$$

where A is the annual usage in units, S is the order cost in dollars, I is the inventory carrying costs expressed as a decimal, and U is the cost of one unit.

$$EOQ = \sqrt{\frac{2 \times 12000 \times 50}{0.1 \times 3}}$$
$$= \sqrt{4{,}000{,}000} = 2000 \text{ Units}$$

The EOQ equation shows that the most economical lot size is a function of the square root of the annual usages of items expressed in dollars.

The square root relationship also shows in a general way how much inventories should increase if the sales of a product increases by 20%. Obviously, the inventory should not be increased by 20%, and the maximum would be the square root of 20.

Generally speaking, the application of the EOQ formula for ordering and lot sizing result in good materials management. At times, however, its inappropriate use may lead to increased inventories and costs, since EOQ quantities may not balance with quantities needed for lot sizes. Set-up costs are also an important factor, and a technique called LIMIT (Lot-size Inventory Management Interpolation Techniques) can be used.

## Inventory Management Systems

In managing inventories, two sets of techniques can be used in a manufacturing operation (1) statistical inventory control or order point and (2) material requirements plan-ning (MRP).

**Statistical inventory control vs. material requirements planning:** As indicated earlier in this chapter, the concepts of mathematical probability, operations research, the Pareto curve, exponential smoothing, and EOQ all indicate that mathematical expressions can be applied in the field of inventory and production control. When properly used, all constitute statistical inventory control, or the order point system which most companies are using today. Before examining its three basic concepts, a closer inspection of the system is necessary.

The order point system represents an attempt to predict on the basis of past data. It is best used when there is an independent demand. For example, one only produces finished goods in response to depletion of inventories resulting from sales demands, and therefore a sales forecast is required. However, the components in finished goods, such as raw materials and finishing supplies, need to be purchased only when supplies needed to produce the finished goods need to be replaced. These requirements can be calculated. Even so, the purchasing department should be given adequate notice. The order point system calculates the quantity and delivery date for each item separately and independently of the other items that must also be used as part of the production.

In addition to the fact that the square root approach does not balance with lot size or the other items needed simultaneously, it is not time-phased and assumes that some inventory should be replenished as soon as it is depleted. If 1,000 kg of a raw material is consumed in February and no further material is needed until June, the order point system dictates immediate replenishment, and a large inventory may result.

### *Concepts of the Order Point System*

The following are three basic concepts of the order point system.

### *1. Safety Stock*

The function of safety stock is to ensure the best level of customer service

and to keep back orders to a minimum. Such stock is necessary because all forecasting techniques have errors, which are generally expressed as MAD or in standard deviations, and safety stock compensates for such errors. The calculation of safety stock is based on the normal distribution curve and its relationship to the desired customer service level. If the MAD or standard deviation of forecast error is known, the order quantity is multiplied by either of these figures, and the result equals the reserve or safety stock. For example, if the weekly forecast is 500, the MAD is 200 units, and a 98% service is desired, the reserve stock should be 512 units (Table 26.5). This is derived by multiplying the 200 units of MAD by 2.56. The use of standard deviation results in a smaller number. Either of these figures added to the 500 units of weekly forecast results in 1,012 units or 910 units, which in turn defines the order point. Obviously, this should be applied only when there is a random demand.

Additionally, the MAD column in the table shows the enormous increase in safety stocks needed when moving from a 90% to a 99.9% service level.

**Table 26.5:** Safety factors for normal distribution

| **Service level** | **Safety factor using** | |
|---|---|---|
| **(% Order cycles w/o stockout)** | **Standard deviation** | **Mean absolute deviation** |
| 50.00 | 0.00 | 0.00 |
| 75.00 | 0.67 | 0.84 |
| 80.00 | 0.84 | 1.05 |
| 84.13 | 1.00 | 1.25 |
| 85.00 | 1.04 | 1.30 |
| 89.44 | 1.25 | 1.56 |
| 90.00 | 1.28 | 1.60 |
| 93.32 | 1.50 | 1.88 |
| 94.00 | 1.56 | 1.95 |
| 94.52 | 1.60 | 2.00 |
| 95.00 | 1.65 | 2.06 |
| 96.00 | 1.75 | 2.19 |

| | | |
|---|---|---|
| 97.00 | 1.88 | 2,35 |
| 97.72 | 2.00 | 2.50 |
| 98.00 | 2.05 | 2.56 |
| 98.61 | 2.20 | 2.75 |
| 99.00 | 2.33 | 2.91 |
| 99.18 | 2.40 | 3.00 |
| 99.38 | 2.50 | 3.13 |
| 99.50 | 2.57 | 3.20 |
| 99.60 | 2.65 | 3.31 |
| 99.70 | 2.75 | 3.44 |
| 99.80 | 2.88 | 3.60 |
| 99.86 | 3.00 | 3.75 |
| 99.90 | 3.09 | 3.85 |
| 99.93 | 3.20 | 4.00 |
| 99.99 | 4.00 | 5.00 |

Adapted from Plossl and Wight.

## *2. Order Point*

Figure 26.23 shows that at some point in time, 25,000 items are in inventory, and as the weeks pass, the inventory is consumed as shown on the downward sloping line. This consumption continues until the inventory reaches a predetermined level or ordering point, at which time a replenishment quantity is ordered. This system assumes that usage is uniform; thus, the average inventory is equal to one-half the order quantity. The lead time, of course, is that time lapse between placement of an order and the time it is received and approved by quality control. Since both demand and lead time vary, the reserve stock is really established to take care of both. Accordingly, the order point is equal to the sum of the anticipated demand during lead time plus the reserve stock. In Fig. 25.10, during the five weeks of lead time, about 5,000 units will be depleted from stock, and this amount plus the reserve of 10,000 units indicate an order point of 15,000. It must be remembered that randomness of demand during lead time and gradual inventory depletion are 

the underlying assumptions.

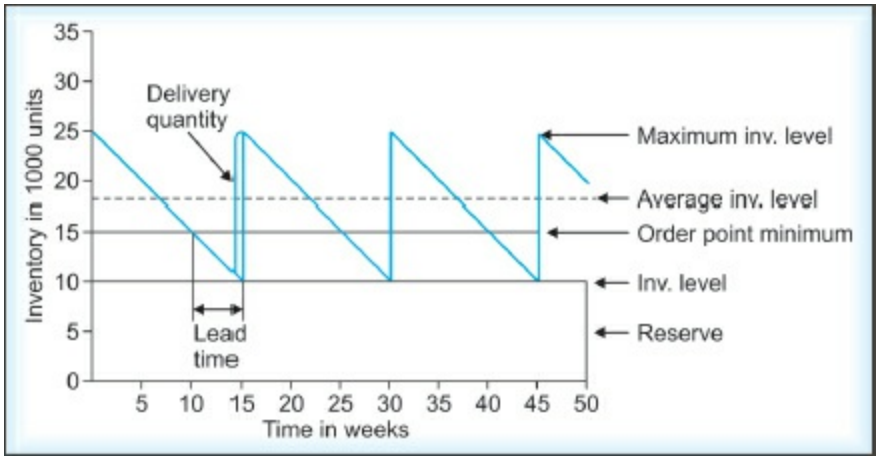



**Fig. 26.23:** Graphic representation of order point

### *3. Order Quantity*

The EOQ formula previously described assumes gradual inventory depletion, so that the lot size inventory is equal to half the order quantity.

## *Material Requirements Planning (MRP)*

As stated previously, the order point system is best applied when the demand is independent, as for finished goods. A more sophisticated and mature approach to production and inventory control is the recognition of the principles of independent versus dependent demand, first expressed by Dr Orliky of IBM in 1965. When the demand is for all components that "go into" an item, material requirements planning is a far more satisfactory technique.

If an item such as a bulk tablet is made infrequently, or at least not on a continuous basis, and is made in a large batch, the demand is occasional, and immediate replenishment of the raw materials may not be necessary, or in fact, may be undesirable. An order point system would immediately replenish any or all of the raw materials, but in requirements planning, ordering would be postponed until the next scheduled production day, in consideration of the lead time. In other words, in addition to holding orders until needed, this method also establishes the next requirement date arid is said to be time-phased. Under these conditions, if there is a sudden change in quantity or date, the frequent updating in an MRP system permits changing in sufficient

time to avoid problems.

There are three general principles to remember for an effective MRP:

1. The materials plan must be revised frequently to react to changes in requirements.
2. The smaller the time period used, the more effective the materials plan will be.
3. The materials plan must extend over a long enough period to cover the longest lead time of any component.

The following example illustrates the difference in order point versus MRP:

| | |
|---|---|
| On hand | 50 |
| Plus on order | 150 |
| Less gross requirement | 300 |
| Net requirement | –100 |

In an order point system, 100 units are needed immediately, and no one knows when the 150 on order are needed, but presumably, they are needed immediately also.

In an MRP system, the gross requirements are broken down by quantity and required date as follows:

| Required | Date needed |
|---|---|
| 50 | 1/3 |
| 30 | 2/14 |
| 70 | 3/4 |
| 100 | 3/27 |
| 50 | 4/25 |
| 300 | |

It is assumed that it is now January 2, and that the lead time is four weeks. Actually, the 100 ordered for immediate delivery in the order point system do not need to be ordered until the last week in February. The 50 needed on January 3 are already on hand, whereas the 150 on order will cover 

2/14, 3/4, and half of 3/27. By time-phasing the order, the material was brought in only when needed, thereby contributing to reduction of inventory and inventory carrying costs.

## Cost Controls

From the sales forecast, the materials management group generates a production forecast that takes into account seasonality, deals and promotions, introduction of new products or product sizes, and other factors that create a demand on inventory, equipment, and personnel. This information in turn permits the development of the operating or expense budgets for each department. The information thus provided is applicable to a standard cost accounting system, which in its simplest terms, means that individual item costs, as well as labor and burden rates, have been fixed or standardized for a period of time, usually a calendar or fiscal year.

### *The Three Elements of Cost*

The basic elements of cost are materials, labor, and burden; the last item is sometimes referred to as “overhead”.

### *1. Materials*

Based on the needs expressed in the production forecast, the purchasing department negotiates contracts and price arrangements, and ascertains the standard cost for each item. These standard costs form the basis of a monthly price variance report, which not only measures the effectiveness of purchasing, but also monitors changes in cost. This monitoring is particularly important in the pharmaceutical industry, in which materials represent the largest part of the cost of goods.

The purchase price variance report shown in Table 26.6 is for the month of April and shows a favorable ($74,800) cumulative total variance through the four-month period for three operating divisions of the company as well as for advertising. For Division B, for example, there is a favorable variance of $10,200 for the month of April and $33,200 total of all divisions. Since each division shows a favorable purchase price variance, their purchases are generally within the standard cost predicted. However, to permit better analysis of price variances, the figures are broken down into major categories, such as raw materials, certain classes of finishing supplies, and others, to help pick out favorable and unfavorable price variances.

By categories, for example, the report shows that raw materials were purchased below standard costs, showing a savings of $30,100 for the month

of April and a total savings of $43,300 over a four-month period. This would indicate that the purchasing division is doing a good job, and if the savings turn out to be substantial at year's end, perhaps selected raw materials cost standards could be revised downward next year. In general, it is desirable to have a favorable price variance, but on the other hand, a variance that is too high might indicate poor estimation on the part of the purchasing division; it would overstate or understate the cost of goods.

**Table 26.6:** Purchase price variance report (In M Dollars)

| | Month of april | | | Year to date | | | |
|---|---|---|---|---|---|---|---|
| By division | Actual Purchase Price | Standard Cost | Variance | Actual Purchase Price | Standard Cost | Variance | Budget |
| | ($) | ($) | ($) | ($) | ($) | ($) | ($) |
| Division A | 2.1 | 2.8 | (–7) | 27.6 | 30.8 | (3.2) | |
| Division B | 1,867.6 | 1,877.8 | (10.2) | 6,923.4 | 6,956.6 | (33.2) | |
| Division C | 671.0 | 684.9 | (13.9) | 1,526.1 | 1,549.2 | (23.1) | 20.0 |
| Advertising and promotion | 28.6 | 37.0 | (8.4) | 78.8 | 94.1 | (15.3) | |
| **Total** | **2,569.3** | **2,602.5** | **(33.2)** | **8,555.9** | **8,630.7** | **(74.8)** | **20.0** |
| *By category* | | | | | | | |
| Raw material | 579.8 | 609.9 | (30.1) | 2,251.8 | 2,295.1 | (43.3) | |
| Bulk | 213.7 | 184.1 | 29.6 | 501.5 | 463.3 | 38.2 | 20.0 |
| Set-up box | 23.6 | 23.8 | (.2) | 56.2 | 58.3 | (2.1) | |
| Folding carton | 69.7 | 64.6 | 5.1 | 328.0 | 330.9 | (2.9) | |
| Labels and inserts | 74.1 | 73.3 | .8 | 221.2 | 231.8 | (10.6) | |
| Bottles | 264.0 | 287.4 | (23.4) | 755.8 | 780.5 | (24.7) | |
| Caps | 44.2 | 47.7 | (3.5) | 116.9 | 126.0 | (9.1) | |
| Corrugated | 24.0 | 24.1 | (.1) | 59.8 | 56.8 | 3.0 | |
| Cans, drums | 32.6 | 34.2 | (1.6) | 73.5 | 75.7 | (2.2) | |
| Cotton, cello | 60.5 | 64.2 | (3.7) | 172.1 | 183.1 | (11.0) | |
| Advertising and promotion | 28.6 | 37.0 | (8.4) | 78.8 | 94.1 | (15.3) | |
| *Vendor charges* | 14.9 | 18.9 | (4.0) | 125.2 | 126.9 | (1.7) | |
| Finished goods, Division A | 613.5 | 613.4 | .1 | 1,437.5 | 1,438.9 | (1.4) | |
| Finished goods, Division B | 493.7 | 487.9 | 5.8 | 1,608.9 | 1,603.1 | 5.8 | |
| Finished goods, Other | 32.4 | 32.0 | .4 | 768.7 | 766.2 | 2.5 | |
| **Total** | **2,569.3** | **2,602.5** | **(33.2)** | **8,555.9** | **8,630.7** | **(74.8)** | **20.0** |

Denotes favorable variance.

Another control occurs in the usage of materials during production. From the

production forecast, production planning determines the number of batches of each product needed to meet sales forecasts, and batch sheets for each product give the exact formula and quantity of all raw materials.

In addition to this, and in conformance with GMP, the batch sheet shows actual bulk yields, which can be compared with the theoretical yields, adjusted to take care of normal losses during manufacturing. A materials usage variance report is issued for each department and Table 26.7 explains the cumulative results for tabletting through the month of May. This report expresses a favorable or unfavorable variance for each product, and it can be seen that this department shows a net favorable variance of $7,018. The value of this report, however, is that it shows the variances for each product, two of which show losses in excess of $2,000—a situation that should be further examined.

A similar report covers variances in materials usage in the packaging department (Table 26.7). In this case, these variances are expressed by packaging line, with respect to both variations in bulk such as tablets and liquid products, and variations in finishing supplies. For example, the May report shows that there was a total unfavorable bulk variance of $15,399 in the packaging department and an unfavorable variance of $27,796 in finishing supplies.

**Table 26.7:** Summary tabletting, granulating, and packaging cost performance (Year to date)

**Tabletting and granulating cost performance**

| Product name | Material Usage variance | Std. Hrs. Delivered to invty. | Labor time | Lost | Total cost |
|---|---|---|---|---|---|
| | $ | | $ | $ | $ |
| "A" | 116 | 615 | 912 | | 1028 |
| "B" | (1493) | 948 | 147 | | (1346) |
| "C" | 1392 | 648 | (79) | | 1313 |
| "D" | 866 | 877 | 579 | | 1445 |
| etc. | 2211 | 546 | 1025 | | 3236 |
| | 2568 | 375 | 1574 | | 4142 |
| | (896) | 401 | 166 | | (730) |
| | (4869) | 187 | 436 | | (4433) |
| | | | | 1992 | 1992 |
| **Total** | **$(7018)** | **9042** | **10130** | **1992** | **5104** |

**Packaging cost performance**

| Line No. | Packaging line descrip. | Material usage variance | | Std. Hrs. delivered to invty. | Labor perf. | Lost time | Total cost over std. |
|---|---|---|---|---|---|---|---|
| | | Bulk | Fin. supply | | | | |
| | | $ | $ | | $ | $ | $ |
| 1 | Tablet | 1 345 | 3406 | 3767 | 468 | 1185 | 5404 |
| 2 | | 2 4446 | 1574 | 3257 | 1234 | 1404 | 8658 |
| 3 | | 3 (56) | 1668 | 3559 | 2350 | 1774 | 5736 |
| 4 | | 4 2553 | (2498) | 15190 | (3561) | 1277 | (2229) |
| 5 | | 5 (976) | 927 | 1668 | 3827 | 1468 | 6249 |
| 6 | | 6 (380) | (1864) | 4569 | 1152 | 1370 | (725) |
| 7 | | 7 (432) | 732 | 5713 | 3497 | 2459 | 6256 |
| 24 | | (141) | 42 | 832 | 133 | 191 | 225 |
| 25 | | 8531 | 1738 | 14649 | 12877 | 6612 | 29758 |
| 26 | | 233 | 727 | 3567 | 3092 | 2300 | 6402 |
| | **Total** | **15399** | **27796** | | **31512** | **36272** | **110979** |

In all cases, the accounting, production, and quality control departments have agreed on a standard yield as compared with theory, and it probably has been derived on the basis of statistical analysis with losses built into the cost system. Variance in finishing supplies is derived by comparing the finished goods produced in packaging with the amount that should have been consumed on the packaging order and with losses on the packaging line. Both reports show that product losses on bulk can be quite expensive, and differences in yields between processing and packaging point to the need for precision in the yielding of bulk tablets, liquids, and ointments by using counters on tablet presses or meters and load cells in other areas, a practice that will help in GMP reconciliation of bulk.

As stated previously, the biggest item in pharmaceutical cost-of-goods is 

materials. Reports such as those illustrated are furnished to each operating department for the department head to analyze quickly and decide which products should be studied further so that problems can be corrected.

In the control of materials costs, the variances are budgeted as a hedge against unforeseeable price changes. The report described in this section is helpful to both purchasing and production departments, as it points to problems and permits their correction. Of course, the information is important to the accounting department and to management as well.

## *2. Labor*

**Direct labor:** In any manufacturing operation, the management and careful use of the labor force are important for cost control as well as for the stability and happiness of the work force. In most companies, the number of hours required in each department for production of most items is known either by virtue of historical information or as a result of time studies by industrial engineers. The latter standard is more practical, of course, since the historical data does not necessarily reflect the most efficient method of production. In either case, the production of any bulk or finished goods item is expressed as the number of direct labor hours needed to produce 1,000 bulk tablets or 1,000 bottles of 100's for every product made, including promotional samples. This information is important to several groups, e.g. the accounting department, production planning, and, of course, the operating managers.

**Indirect labor:** It is easy to identify the direct labor needed for an item, but to arrive at the true cost, indirect labor hours are needed. They are difficult to measure, however. In some plants, the indirect labor has not been adequately identified, and in some departments, it represents a high percentage of the total requirements. Without indirect labor hours, it would be difficult to analyze and control available hours of work and to use the labor force efficiently. If a serviceman is bringing caps and bottles to three packaging lines, his function represents some indirect labor to each of the lines. Whenever possible, it would be best to allocate time to the products on each line rather than completely lose its identity and cost by putting it into the indirect labor class, thereby spreading the cost over the rest of the products.

The indirect labor costs that cannot be put into a work standard should be attributed to a specific purpose, so that they can be identified and controlled. Some examples are the costs of mechanics, servicemen, equipment cleaning,

line changeover, quality control inspectors, and operation of the label room in a packaging department. In a department such as packaging, which may employ 200 to 300 people, the indirect labor could amount to several hundred thousand dollars a year; therefore, each category should be budgeted and compared against actual costs on a monthly basis.

Another cost to watch is the use of direct labor for an indirect labor job. The use of direct labor for such work should be minimized, since they have a labor rate higher than indirect labor people and would contri-bute to increasing the labor cost.

Every effort should be made to lower the percentage of indirect labor to total labor for good cost control. Each labor category should be examined periodically to see if it can be placed into a time standard.

It is in the interest of efficiency to break the work standard down into job skills, e.g. machine operators, table workers, and mechanics, since in the scheduling of production these services must be available when required.

## *3. Burden*

There are two types of burden, direct and fixed.

**Direct burden:** Categorized as direct burden are such expenditures as supervision and clerical help; lost time; premium on overtime; vacation, holiday, and sick pay; and such employee benefits as hospitalization, insurance, and retirement benefits.

Additional items falling into direct burden are controllable expenses incurred in the operation of each department and commonly called operating expenses.

These expenses would include items such as laundry for uniforms, maintenance department charges, travel, membership dues and seminars. Another important item in each departmental budget is operating supplies. For example, in the parenteral department, the cost of ethylene oxide used in sterilization and the replacement of HEPA filters should be included. Repairs and maintenance, as well as supplies used for production equipment such as punches and dies, must be estimated since these can add substantially to the budget.

**Fixed burden:** In addition to the operating departments, others such as engineering, quality control, materials management, and all departments

reporting to each of them also prepare operations budgets, and the expenses and labor costs form the overhead to be charged in the manufacturing costs. In addition, the costs of fuel, electricity, land and real estate taxes, and depreciation on the building and equipment. In the case of quality control, efforts have been made by some companies to allocate as much cost as possible directly into production departments to get as close as possible to true costs. With a little effort, it is practical in some quality control services to assign reasonably accurate labor cost and expenses to some production areas. As a simplified illustration of this, if someone is keeping a facility and requires rabbits to test for the presence of pyrogens, the cost of this can be allocated to the parenteral department, which requires this service. This principle is not unlike identifying indirect labor into the direct category, and in the example used, it avoids spreading the cost of pyrogen testing over the other cost centers, which do not use this service.

# 27:
# Packaging Material Science

In the pharmaceutical industry, it is vital that the package selected must preserve the integrity of the product. The selection of a package therefore begins with determination of the product's physical and chemical characteristics, its protective needs, and its marketing requirements. Desirable attributes which a packaging material must have include:

- Protection of product from environmental conditions:
  - Ligfa
  - Gases
  - Moisture
  - Solvent loss
  - Sterility
- Safety:
  - Nontoxic
  - Must not impart taste and odor
- Compatibility:
  - Must not be reactive with product.
- Performance:
  - Adaptable with common high speed packaging machines
  - Must meet tamper-resistance and tamper-evident requirements
  - Transit worthy
  - Economical
  - Eco-friendly

Today, packaging serves a wide range of functions. The FDA mandates that drug packages must "maintain the standards of identity, strength, quality, and purity of a drug for its intended shelf life". Not only is the container-closure system in which a drug is distributed, but also the packaging system 

itself is expected to ensure the properties of the pharmaceutical preparation. A container closure system refers to the sum of packaging components that together contain and protect the dosage form, limit the unauthorized tampering, support new delivery systems, and be both child-resistant and senior-friendly. In general, there are several recognized layers of packaging systems. *Primary packaging component* are the components which are in direct contact with the pharmaceutical dosage form, e.g. liners, bottles, desiccant in bottles and blister films. These components typically protect the drug from the environment (e.g. moisture, gases, and light). *Secondary packaging component*, means packaging components that are not, and will not be, in direct contact with the dosage form, e.g. cartons, and overwraps for blisters. These components generally provide protection and labelling for the primary container.

Packaging systems are also classified according to the dosage capacity of the container. A *single-unit container* is one that contains a only a unit quantity of medication which can be used one time. A single-dose container is a single-unit container designed for parenteral administration only, while a *unit-dose container* is a single-unit container intended for solid oral dosage forms. The unit-of-use container is one that may contain any number of doses, but is designed to be dispensed as packaged without further repackaging and is intended for a specific therapeutic dosing regimen or use. A *multiple-unit container* is one that encloses multiple doses and permits multiple withdrawals of oral dosage forms, while a multiple-dose container is a multiple-unit container designed for parenteral administration only.

In addition to the types of containers that constitute a package, there are also different types of containers that may be specified based on the stability and usage characteristics of a given drug. Over the counter, ophthalmic and otic products require tamper-resistant packaging. Drugs that undergo photooxidation need to be packaged in light-resistant containers. For moisture-sensitive drugs, a tight container (e.g. one that protects the contents from contamination or loss) or a hermetic container (e.g. one that is “impervious to... gas”) is necessary. Well-closed containers protect their contents from solid-type contaminants and other handling processes.

Owing to the broad scope of the subject, a detailed treatment of the science of packaging as related to pharmaceuticals cannot be adequately covered in this chapter. However, basic topics have been selected for

discussion to provide a general understanding of this subject. These are limited to the protective function of commonly used packaging materials, their limitations, and their possible interaction with various drugs. The commonly employed pharmaceutical packaging materials are glass, plastics, metals and paper.

# TYPES OF PACKAGING MATERIALS

Faulty packaging of pharmaceutical dosage forms can invalidate the most stable formulation. Consequently, it is essential that the choice of container materials for any particular product be made only after a thorough evaluation has been made of the influence of these materials on the stability of the product and of the effectiveness of the container in protecting the product during extended storage under varying environmental conditions of temperature, humidity and light.

The materials most commonly employed as container components for pharmaceutical preparations include glass, plastic, metal, and rubber.

## GLASS

Glass is commonly used in pharmaceutical packaging because it possesses superior protective qualities; elegancy and containers are readily available in a variety of sizes and shapes. It is essentially chemically inert, impermeable, strong, and rigid, and has FDA clearance. Glass does not deteriorate with age, and with a proper closure system, it provides an excellent barrier against practically every element except light. Colored glass, especially amber, can give protection against light when it is required. The major disadvantages of glass as a packaging material are its fragility, its weight and expensive machinability.

## Composition of Glass

Glass is composed principally of sand ($SiO_2$), soda ash ($Na_2CO_3$), limestone ($CaCO_3$), and cullet. Cullet is broken glass that is mixed with the sand, soda ash and limestone to acts as a fusion agent for the entire mixture. The composition of glass varies and is usually adjusted for specific purposes. The most common cations found in pharmaceutical glassware are silicon, aluminum, boron, sodium, potassium, calcium, magnesium, zinc, and barium. The only anion of consequence is oxygen. Many useful properties of glass are affected by the kind of elements it contains. Reduction in the proportion of sodium ions makes glass chemically resistant; however, without sodium or other alkalies, glass is difficult and expensive to melt. Boron oxide is incorporated mainly to aid in the melting process through reduction of the temperature required. Lead in small traces gives clarity and brilliance, but produces a relatively soft grade of glass. Alumina (aluminum oxide), however, is often used to increase the hardness and durability and to increase resistance to chemical action.

## Manufacture of Glass Containers

Four basic processes used in the production of glass containers are: blowing, drawing, pressing, and casting. *Blowing* uses compressed air to convert molten glass to the cavity of a metal mold. Most commercial bottles and jars are produced on automatic equipment by this method. In *drawing*, molten glass is pulled through dies or rollers that shape the glass. Rods, tubes, sheet glass, and other items of uniform diameter are usually produced commercially by drawing. Ampoules, cartridges, and vials drawn from tubing have a thinner, more uniform wall thickness, with less distortion than blow-molded containers. In *pressing*, mechanical force is used to press the molten glass against the side of a mold. *Casting* uses gravity or centrifugal force to cause molten glass to form in the cavity of the mold.

## Colored Glass-light Protection

Glass containers for drugs are generally available in clear flint or amber color. For decorative purposes, special colors such as blue, emerald green, and opal may be obtained from the glass manufacturer. Only amber glass and red glass are effective in protecting the contents of a bottle from the effects of sunlight by screening out harmful ultraviolet rays. The USP specifications for light-resistant containers require the glass to provide protection against 290 to 450 nm of light. Amber glass meets these specifications, but the iron oxide added to produce this color could leach into the product. Therefore, if the product contains ingredients subject to iron-catalyzed chemical reactions, amber glass should not be used.

## Glass for Drugs

The USP and NF describe the various types of glass (type I, type II, type III and type NP) and provide the powdered glass and water attack tests for evaluating the chemical resistance of given glass. These test are the measures of the amount of alkalinity leached from the glass by purified water under controlled elevated temperature conditions. The powdered glass test is performed on crushed grains of a specific size and is meant for type I, type III, and type NP glass. The water attack test is conducted on whole containers and is used only with type II glass.

### *Type I—Borosilicate Glass*

In this highly resistant glass, a substantial part of the alkali ($Na^+$ and $K^+$) and earth cations ($Be^{2+}$, $Mg^{2+}$, $Ca^{2+}$, $Sr^{2+}$, $Ba^{2+}$ and $Ra^{2+}$) are replaced by boron and/or aluminum and zinc. It is more chemically inert than the soda-lime glass, which contains either none or an insignificant amount of these cations. Although glass is considered to be an inert material and is used to contain strong acids and alkalies as well as all types of solvents, it has a definite and measurable chemical reaction with some substances, notably water. The sodium is loosely combined with the silicon and is leached from the surface of the glass by water. Distilled water stored for one year in flint type III glass (to be described) picks up 10 to 15 parts per million (ppm) of sodium hydroxide along with traces of other ingredients of the glass. The addition of approximately 6% boron to form type I borosilicate glass reduces the leaching action, so that only 0.5 ppm is dissolved in a year. Type I glass is generally used for preparation intended for parenteral administration.

### *Type II—Treated Soda-lime Glass*

When glassware is stored for several months, especially in a damp atmosphere or with extreme temperature variations, the wetting of the surface by condensed moisture (condensation) results in salts being dissolved out of the glass. This is called “blooming” or “weathering,” and in its early stages, it gives the appearance of fine crystals on the glass. At this stage, these salts can be washed off with water or acid. Type II containers are made of commercial soda-lime glass that is already de-alkalized, or treated to remove surface alkali. The de-alkalizing process is known as “sulfur treatment” and

virtually prevents “weathering” of empty bottles. The treatment offered by several glass manufacturers exposes the glass to an atmosphere containing water vapor and acidic gases, particularly sulfur dioxide at an elevated temperature. This results in a reaction between the gases and some of the surface alkali, rendering the surface fairly resistant, for a period of time, to attack by water. The alkali removed from the glass appears on the surface as a sulfate bloom, which is removed when the containers are washed before filling. Sulfur treatment neutralizes the alkaline oxides on the surface, thereby rendering the glass more chemically resistant. Type II glass is suitable for neutral and acidic parenteral preparations.

### *Type III—Regular Soda-lime Glass*

Containers are untreated and made of commercial soda-lime glass. Type III glass is generally not used for parenteral preparations

### *Type NP—General-purpose Soda-lime Glass*

Containers made of soda-lime glass are supplied for nonparenteral products.

## Drug-glass Considerations

Although glass exhibits many advantages over other packaging materials, it has two principal faults, namely, the release of alkali and the release of insoluble flakes to liquids stored in the container. The USP XX acknowledges that alkali ions can be released for each of four types of glass listed in the USP compendium, as shown by the data in Table 27.1.

By decreasing the soda content in the glass or replacing the sodium oxide with other oxides, it has been possible to overcome the property of glass to release alkali cations into solution. This is exemplified by the borosilicate glass in Table 27.1 as compared with the general purpose soda-lime glass. Several approaches have been used to increase the resistance of glass to alkali release by surface treatment. One method consists of treating the surface of soda-lime glass to produce a fire-polished skin of silica, which is more resistant than the inner layers of glass. This surface-treated glass is represented by Type II glass in Table 27.1.

**Table 27.1:** Glass types and USP test limits

| Type | General description | Type of test | Limits | |
|---|---|---|---|---|
| | | | Size (ml) | ml of 0.02 N Acid |
| I | Highly resistant borosilicate glass | Powdered glass | All | 1.0 |
| II | Treated soda-lime glass | Water attack | 100 or less | 0.7 |
| | | | Over 100 | 0.2 |
| III | Soda-lime glass | Powdered glass | All | 8.5 |
| IV | General purpose soda-lime glass | Powdered glass | All | 15.0 |

Another approach is to treat the surface of the glass with sulfur dioxide in the presence of water vapor and heat. This causes the surface alkali to react with the sulfur dioxide, and the glass becomes more resistant, as shown by the data in Table 27.2, for 5 hour extractions performed with boiling water for treated and untreated glass. The suggested reason for the greater resistivity of this glass is that a layer of positively charged hydrogen ion glass forms, causing the development of a compacted layer on drying and the expulsion of water. The diffusion of sodium through such a layer is greatly retarded; hence, the extraction by water is considerably reduced.

**Table 27.2:** Extraction of alkali

| Glass treatment | Extraction as mg $Na_2O$ |
|---|---|
| Annealed in absence of $SO_2$ | 2.9 |
| Annealed in presence of $SO_2$ | 0.9 |

In recent years, a number of drugs have been developed that are of high potency, and consequently, of low dosage. The stability of these drugs can be readily affected by the release of soluble alkali from glass containers. As a safety factor, whenever the dosage form is liquid, the solution is buffered to eliminate any effect due to possible change in pH if some alkali were released from the glass.

Sometimes, insoluble flakes have been found to appear in solutions stored in glass containers. The type of glass employed plays a major role in whether flake formation takes place. For example, flake formation may occur in nonborosilicate glass immediately after autoclaving, whereas in borosilicate glass, it occurs at temperatures much higher than those normally used for autoclaving.

Glass containers may possess various additives, such as oxides of boron, sodium, potassium, calcium, iron, and magnesium which alter physical and chemical properties of the glass. For example, when formulating sulfate salts, (drug substances or antioxidants) the glass container should have minimal amounts of calcium and barium to prevent the formation of insoluble inorganic salts. Table 27.3 shows six parenteral solutions containing barium sulfate crystals, which were identified by polarizing microscopy, scanning electron microscopy with energy-dispersive X-ray analysis, micro-X-ray powder diffraction, and micro Raman spectroscopy. On the basis that barium sulfate forms in situ, the sulfate ions appear to originate from the union of the drug and/or the bisulfite ion of the antioxidant. The barium ions originate from the borosilicate glass, which can contain up to 5% Barium oxide. Flakes are also likely to occur with phosphate, citrate, tartrate, and alkaline solutions. Pretreatment of the containers with dilute acid solution was found to delay flake formation.

**Table 27.3:** Parenteral solutions containing barium sulfate crystals

| Product | Type of glass | Source of sulfate or |
|---|---|---|

| | container | bisulfite ions |
|---|---|---|
| Procaine hydrochloride, 2% | 30 ml vial | Antioxidant |
| Kanamycin sulfate, 0.5 g/2 ml | 2 ml vial | Drug |
| Meperidine hydrochloride, 100 mg/ml | 1 ml vial | Antioxidant |
| Magnesium sulfate, 5 g/10 ml | 10 ml vial (plunger) and rubber closure | Drug |
| Atropine sulfate, 0.5 mg/ml | 1 ml ampule | Drug |
| Promethazine hydrochloride, 25 mg/ml | 1 ml ampule | Antioxidant |

Many pharmaceutical preparations exhibit physical or chemical changes due to the radiant energy of light. Light radiations can cause color development or color fading and initiate an oxidation-reduction reaction, resulting in drug degradation, rancidity of oil formulations, and flavor and odor loss. Products exhibiting physical and chemical changes resulting from the effects of radiant energy can, in most instances, be adequately protected by the use of special glass containers. Flint glass, which is the most widely used multipurpose container material, has the disadvantage of being transparent to light rays above 300 mp. As a result, amber glass, which has the property of shutting out certain portions of the light spectrum, has been used extensively by the pharmaceutical industry. The transmission curves for flint and amber glass in Fig. 27.1 show that although flint glass transmits significantly from 300 mp, amber glass does not begin to transmit to any appreciable extent until 470 mp. Since it is known that the photochemical activity of the light radiations drops off with increasing wave length, it would be expected that the amber container would afford a product better protection against light than the flint glass. This is illustrated by the following examples.

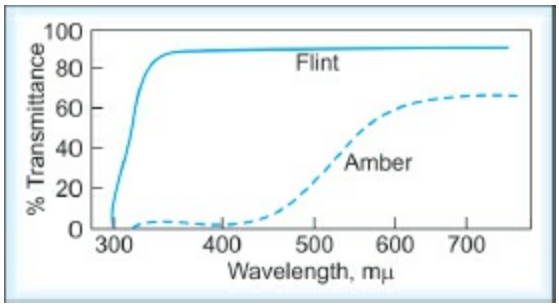



**Fig. 27.1:** Transmission curves for typical flint and amber glass

The data in Table 27.4 show the protective affect of amber glass on the fading of dyes from the surface of tablets stored in flint and amber glass bottles.

**Table 27.4:** The amount of dye (%) faded per day when irradiated under exaggerated intensity

| Glass sample | FD and C Blue #1 | D and C Yellow #10 |
|---|---|---|
| Flint | 12.9 | 38.1 |
| Amber | 4.6 | 6.2 |

The results in Table 27.5 summarize the stabilizing properties of amber glass on the photochemical degradation of an antihistamine injectable solution exposed to exaggerated illumination. The solutions stored in clear glass ampules show a 36% loss of antihistamine concentration, while those in amber glass show no loss in drug concentration.

**Table 27.5:** Chemical assay for residual antihistamine from solutions stored in clear and amber glass ampules

| | Assay (%) | |
|---|---|---|
| **Time (weeks)** | **Clear glass** | **Amber glass** |
| 0 | 100 | 100 |
| 4 | 83 | 98 |
| 8 | 64 | 98 |

The protective effect of amber glass on the darkening of sulfonamide tablets exposed to exaggerated lighting is shown in Table 27.6, where reflectance data is given for the surface of the sulfonamide tablets. It is evident that the tablets darken with storage, since the reflectance values decrease substantially when the tablets are exposed to light in an open dish. Visual observation of these tablets indicated that they had taken on a yellow-tan color; however, for the tablets stored in amber bottles, no significant change in reflectance took place, and visual observation showed no apparent darkening.

**Table 27.6:** Influence of light on the photosensitivity of a sulfonamide tablet formulation

| | % Reflectance at 450 mμ | |
|---|---|---|
| **Time (weeks)** | **Open dish** | **Amber bottle** |
| 0 | 52.6 | 52.6 |
| 2 | 41.3 | 52.3 |
| 4 | 35.6 | 53.6 |
| 6 | 31.7 | 53.8 |
| 8 | 28.9 | 52.9 |
| 10 | 25.4 | — |
| 12 | 23.8 | 51.5 |

The amber color of the glass is imparted by the addition of iron and manganese oxides, cations that are known to catalyze oxidative reactions. Studies have shown that these ions are released from the glass, and that the decomposition rates of several drugs-thiome-rosol, amitriptylene, and L-ascorbic acid, are enhanced in amber glass containers.

The parenteral drug association has published guidelines on the processing and selection of glass containers. Various surface treatments are used to improve chemical resistance and decrease alkalinity. For example, exposing hot containers to sulfur dioxide reduces sodium content at the surface, and a brief treatment with ammonium bifluoride effectively cleans the surface by dissolving a portion of it.

## PLASTIC

Plastics in packaging have proved useful for a number of reasons, including the ease with which they can be formed, their high quality, and the freedom of design to which they can be changed. Plastic containers are extremely resistant to breakage and thus offer safety to consumers along with reduction of breakage losses at all levels of distribution and use.

Plastic containers for pharmaceutical products are primarily made from the following polymers: polyethylene, polypropylene, polyvinyl chloride, polystyrene, and to a lesser extent, polymethyl methacrylate, polyethylene terephthalate, polytrifluoroethylene, the amino formaldehydes, and polyamides.

Plastic containers consist of one or more polymers together with certain additives. Containers manufactured for pharmaceutical purposes must be free of substances that can be extracted in significant quantities by the product contained therein. Thus, the hazards of toxicity or physical and chemical instability are avoided. The amount and nature of the additives are determined by the nature of the polymer, the process used to convert the plastic into the containers, and the service expected from the container. For plastic containers in general, additives may consist of antioxidants, antistatic agents, colors, impact modifiers, lubricants, plasticizers, and stabilizers. Mold release agents are not usually used unless they are required for a specific purpose. At present, a great number of plastic resins are available for the packaging of drug products. A general description of the more popular ones are presented here.

## Polyethylene

High-density polyethylene (HDPE) is the material most widely used for containers by the pharmaceutical industry and will probably continue to be for the next several years. It is a good barrier against moisture, but a relatively poor one against oxygen and other gases. Most solvents do not attack polyethylene, and it remains unaffected by strong acids and alkalies.

Lack of clarity and a relatively high rate of permeation of essential odors, flavors, and oxygen militate against the use of polyethylene as a container material for certain pharmaceutical preparations. Despite these problems, polyethylene in all its variations offers the best all-around protection to the greatest number of products at the lowest cost.

The density of polyethylene, which ranges from 0.91 to 0.96, directly determines the basic physical characteristics of the blow-molded container (1) stiffness, (2) moisture-vapor transmission, (3) stress cracking and (4) clarity or translucency. As the density increases, the material becomes stiffer, has a higher distortion and melting temperature, becomes less permeable to gases and vapors, and becomes less resistant to stress cracking. The molecular structure of high-density material is essentially the same as that of low-density material, the main difference being fewer side branches.

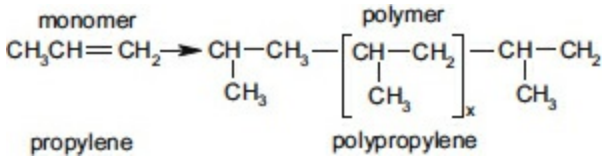
Since these polymers are generally susceptible to oxidative degradation during processing and subsequent exposure, the addition of some antioxidant is necessary. Usually levels of hundreds of parts per million are used. Antioxidants generally used are butylated hydroxy toluene or dilauryl thiodipropionate.

Antistatic additives are often used in bottle grade polyethylenes. Their purpose is to minimize airborne dust accumulation at the surface bottle during handling, filling, and storage. These antistatic additives are usually polyethylene glycols or long chain fatty amides and are often used at 0.1 to 0.2% concentrations in HDPE.

## Polypropylene (PP)

Polypropylene (PP) has recently become popular because it has many of the good features of polyethylene, with one major disadvantage either eliminated or minimized. Polypropylene does not stress-crack under any conditions. Except for hot aromatic or halogenated solvents, which soften it, this polymer has good resistance to almost all types of chemicals, including strong acids, alkalies, and most organic materials. Its high melting point makes it suitable for boilable packages and for sterilizable products. Lack of clarity is still a drawback, but improvement is possible with the construction of thinner walls.

Polypropylene is an excellent gas and vapor barrier. Its resistance to permeation is slightly better than that of HDPE, and it is superior to low-density or branched polyethylene (LDPE). One of the biggest disadvantages of polypropylene is its brittleness at low temperatures. In its purest form, it is quite fragile at 0°F and must be blended with polyethylene or other materials to give it the impact resistance required for packaging.



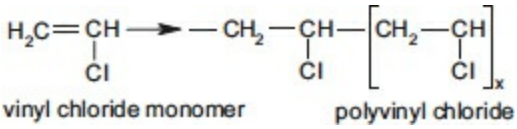
## Polyvinyl Chloride (PVC)

Clear rigid polyvinyl chloride bottles overcome some of the deficiencies of polyethylene. They can be produced with crystal clarity, provide a fairly good oxygen barrier, and have greater stiffness. In its natural state, polyvinyl chloride has poor impact resistance. It can be softened with plasticizers. Various stabilizers, antioxidants, lubricants, or colorants may be incorporated. PVC is seldom used in its purest form. It is an inexpensive, tough, clear material that is relatively easily processed. It must not be overheated because it starts to degrade at 280°F, and the degradation products are extremely corrosive. PVC yellows when exposed to heat or ultraviolet light, unless a stabilizer is included by the resin supplier. It is virtually impossible to process vinyls at elevated temperatures without a stabilizing agent. From the standpoint of clarity, the best stabilizers are the tin compounds, but the majority cannot be used for food or drug products. Dioctyl-tin mercaptoacetate and maleate compounds have been approved by the FDA, but these have a slight odor, which is noticeable in freshly blown bottles. Other acceptable stabilizers (sulfur, calcium, and zinc salts) have a yellowish cast, which makes the plastic hazy and undesirable. In the formulation of PVC compounds with calcium-zinc stabilization materials, all ingredients are used in concentrations below their maximum extractable concentrations.

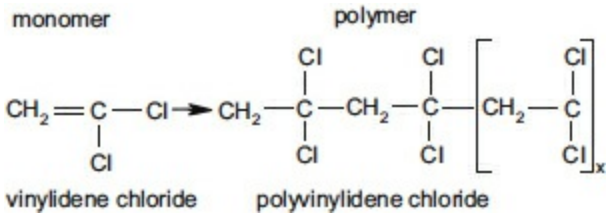
PVC is an excellent barrier for oil, both volatile and fixed alcohols, and petroleum solvents. It retains odour and is a good barrier for oxygen. Rigid PVC is a fairly good barrier for moisture and gases in general, but plasticizers reduce these properties. PVC is not affected by acids or alkalies except for some oxidizing acids.

Much of the concern about the safety of plastic products has focused on PVC plastics and on compounds that are derived from vinyl chloride monomer. Its possible incrimination in the development of cancer of the liver (angiosarcoma) in some persons exposed to vinyl chloride monomer and polyvinyl chloride during manufacture has received widespread interest. Thus far, it seems that the only suspect is vinyl chloride, and the disease is not associated with polyvinyl chloride itself. Major reductions in the amount of residual vinyl chloride monomer have been achieved by PVC producers and its use for the fabrication of plastic bottles now represents the fastest growing

application in the United States. Polyvinyl chloride may also be used as a skin coating on glass bottles. This is accomplished by dipping the bottle in a polyvinyl chloride plastisol and curing the coating, which produces a shatter-resistant coating over the glass bottle.




The relative moisture and gas impermeability in addition to the inexpensive cost and versatility has led to the widespread use of PVC in blister packs, although more impermeable materials are required for more highly sensitive preparations. These materials are often formed by coating PVC with polyvinylidene chloride (PVdC, Saran) or Polymonochlorotrifluoroethylene (PCTFE, Aclar).



## Polymonochlorotrifluoroethylene (PCTFE)

PCTFE comes under the trade name Aclar, is one of the most inert plastics, with the lowest permeability to moisture. It is the most expensive plastic and has so far only been used in packaging as a thin layer (laminated to PVC) for blister packing. Three copolymers are in use: 22A, 33C and 88A. 33 C is the cheapest and has a similar permeability at approximately half the thickness. New homopolymers Rx 160, ultRx 2000 and 3000, SupRx 900 offer certain advantages over the copolymers, including lower costs. The basic monomer, chlorofluoroethylene $C_2ClF_3$, is associated with toxicity. Although, PCTFE is a good barrier against moisture, but a poor one against $O_2$, $N_2$ and $CO_2$ relatively to PVdC.

## Polystyrene

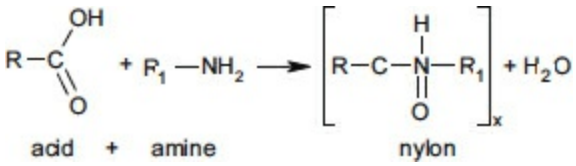
General-purpose polystyrene is a rigid, crystal clear plastic. It has been used by dispensing pharmacists for years for containers for solid dosage forms because it is relatively low in cost. At present, polystyrene is not useful for liquid products. The plastic has a high water vapor transmission (in comparison to high-density polyethylene) as well as high oxygen permeability. Depending on the methods of manufacture and other factors, polystyrene containers are easily scratched and often crack when dropped. Polystyrene also easily builds up a static charge; it has a low melting point (190°F) and therefore cannot be used for hot items or other high-temperature applications. Polystyrene is resistant to acids, except strong oxidizing acids, and to alkalies. It is attacked by many chemicals, which cause it to craze and crack, and so it is generally used for packaging dry products only.

To improve impact strength and brittleness (both of which are sometimes referred to as practical toughness), general purpose polystyrene may be combined with various concentrations of rubber and acrylic compounds. Certain desired properties diminish with impact polystyrene, e.g. clarity and hardness. The shock resistance or toughness of impact polystyrene may be varied by increasing the content of rubber in the material, and often these materials are further classified as intermediate-impact, high-impact, and super-impact polystyrene.

## Nylon (Polyamide)

Nylon is made from a dibasic acid combined with a diamine. Since there are many dibasic acids and many different amines, there is a great variety of nylons. The type of acid and amine that is used is indicated by an identifying number; thus, nylon 6/10 has 6 carbon atoms in the diamine and 10 in the acid. Nylon and similar polyamide materials can be fabricated into thin-wall containers. Nylon can be autoclaved. It is extremely strong and quite difficult to destroy by mechanical means. Important to the widespread acceptance of nylon is its resistance to a wide range of organic and inorganic chemicals. As a barrier material, nylon is highly impermeable to oxygen. It is not a good barrier to water vapor, but when this characteristic is required; nylon film can be laminated to polyethylene or to various other materials.

Its relative high-water transmission rate and the possibility of drug-plastic interaction have reduced the potential of nylon for longterm storage of drugs. Nylon 6, Nylon 6/6, Nylon 6/10, Nylon 11, and certain copolymers are cleared by FDA, subject to limitations or extractables.



## Polycarbonate

Polycarbonate can be made into a clear transparent container. This relatively expensive material has many advantages, one being its ability to be sterilized repeatedly. The container is rigid, like glass, and thus has been considered a possible replacement for glass vials and syringes. It is FDA-approved, although its drug-plastic problems have not been investigated adequately. It is only moderately chemically resistant and only a fair moisture barrier. The plastic is known for its dimensional stability, high impact strength, resistance to strain, low water absorption, transparency, and resistance to heat and flame.

Polycarbonate is resistant to dilute acids, oxidizing or reducing agents, salts, oils (fixed and volatile), greases, and aliphatic hydrocarbons. It is attacked by alkalies, amines, ketones, esters, aromatic hydrocarbons, and some alcohols.

Polycarbonate resins are expensive and consequently are used in speciality containers. Since the impact strength of polycarbonate is almost five times greater than other common packaging plastics, components can be designed with thinner walls to help reduce cost. Polycarbonate articles can be subjected to repeated sterilization in steam or water without undergoing significant degradation.

## Acrylic Multipolymers (Nitrile Polymers)

These polymers represent the acrylonitrile or methacrylonitrile monomer. Their unique properties of being gas barriers, good chemical resistance, excellent strength properties, and safe disposability by incineration make them effective containers for products that are difficult to package in other polymers. Their oil and grease resistance and minimal taste transfer effects are particularly advantageous in food packaging. This medium cost material produces a fairly clear container (not as brilliant as styrene). The use of nitrile polymers for food and pharmaceutical packaging is regulated to standards set by the Food and Drug Administration. The present safety standard is less than 11 ppm residual acrylonitrile monomer, with allowable migration at less than 0.3 ppm for all food products.

## Polyethylene Terephthalate (PET)

Polyethylene terephthalate, generally called PET, is a condensation polymer typically formed by the reaction of terephthalic acid or dimethyl terephthalate with ethylene glycol in the presence of a catalyst. Although used as a packaging film since the late 1950s, its growth has recently escalated with its use in the fabrication of plastic bottles for the carbonated beverage industry. The development of the biaxially oriented PET bottle has had a major impact on the bottling of carbonated beverages, accounting for an estimated annual resin usage of approximately 350 million pounds. Its excellent impact strength and barrier properties make it attractive for use in cosmetics and mouth washes as well as in other products in which strength, toughness, and barrier are important considerations. Furthermore, the resin has been sanctioned for over 25 years by the FDA for food contact applications and has been the recipient of a favorable environmental impact statement.

## Acetal Polyoxymethylene (POM)

POM is available under trade name Delrin and Kemetal. It is obtained from the polymerisation of formaldehyde, and is basically an engineering plastic with high tensile strength, stiffness, and good fatigue endurance. Usage usually lies with devices, aerosol valves and similar engineering components. The density of POM is approximately 1.41. POM is less hygroscopic than nylons, and if correctly stored it does not require pre-drying. Being a polymerisation product of formaldehyde, residues may have to be checked.

## *Regenerated Cellulose*

Regenerated cellulose is included under plastics although it is a derivative of wood pulp. Uncoated regenerated cellulose film is very hygroscopic, highly permeable to water and has poor dimensional stability. Coated film is strong, flexible, transparent with good grease resistance. Coated films are mainly used as transparent overwraps, strip packs and as an outer ply in laminations. It is widely available under trade names such as cellophane and rayophane.

## *Other Plastics*

Coextruded resins are being used to fabricate bottles and thermoformed blisters with barrier characteristics not previously attainable with single resins, resin blends, or copolymers. Coextrusion technology permits the use of high-barrier resins, such as ethylene vinyl alcohol, which could not be used alone because of either high cost or physical/dimensional instability. The resins used in the coextrusion can be selected to provide optimum performance characteristics for the particular product needs. A coextrusion such as polypropylene/ethylene-vinyl-alcohol/polypropylene provides the moisture barrier of polypropylene coupled with the enhanced gas barrier of ethylene vinyl alcohol. Coextruded resins are providing packaging alternatives for products that previously were packaged only in glass.

High-barrier plastics that might compete with glass and metal containers may be available through a new processing technology developed by Du Pont Co. This technology involves dispersing nylon in a polyolefin resin so that the final polymer matrix contains a unique laminar arrangement of nylon platelets, which provide a series of overlapping barrier walls. Reportedly, this technique produces a plastic, which, when compared with the polyolefins,

demonstrates a 140-fold increase as a barrier against certain hydrocarbons and an eightfold increase as a barrier for oxygen.

## Drug-plastic Considerations

A chief disadvantage of plastic containers, as compared with glass, is the problem of permeation in two directions, namely, from the solution into the container and from the ambient environment through the plastic into the preparation. In addition, materials can be leached from the plastic container into the liquid preparation, materials from the preparation can be adsorbed or absorbed onto and into the plastic container, and in certain instances, the contents of the container can chemically or physically react with the plastic components of the container, causing container deformation. The degree of permeation, leaching, sorption, diffusion, and chemical reactivity varies considerably from one plastic material to another. Drug-plastic considerations, therefore, have been divided into five separate categories (1) permeation, (2) leaching, (3) sorption, (4) chemical reaction and (5) alteration in the physical properties of plastics or products.

### *Permeation*

The transmission of gases, vapors, or liquids through plastic packaging materials can have adverse effects on the shelf-life of a drug. Permeation of water vapor and oxygen through the plastic wall into the drug can present a problem if the dosage form is sensitive to hydrolysis and oxidation. Penicillin tablets were found to degrade in polystyrene containers, owing to permeation of water vapor. A change in color and taste of tetracycline suspension was observed owing to permeation of air through the walls of a polyethylene container. Temperature and humidity are important factors influencing the permeability of oxygen and water through plastic. An increase in temperature reflects an increase in the permeability of the gas.

Great differences in permeability are possible, depending on the gas and the plastic used. Molecules do not permeate through crystalline zones; thus, an increase in crystallinity of the material should decrease permeability. Two polyethylene materials may therefore give different permeability values at various temperatures.

Materials such as nylon, which are hydrophilic in nature, are poor barriers to water vapor, while such hydrophobic materials as polyethylene provide much better barriers.

Studies have also revealed that formulations containing volatile

ingredients might change when stored in plastic containers because one or more of the ingredients pass through the walls of the containers. Often, the aroma of cosmetic products becomes objectionable, owing to transmission of one of the ingredients, and the taste of medicinal products changes for the same reason.

The plastic container also may have an influence on the physical system making up the product. For example, certain water-in-oil emulsions cannot be stored in a hydrophobic plastic bottle, since there is a tendency for the oil phase to migrate and diffuse into the plastic.

A clearance of a plastic material from one manufacturer does not necessarily mean that the same type plastic from another blow-molder is equally satisfactory. Each bottle manufacturer has combined its own additives with the basic plastic, and each method of processing can be sufficiently different to seriously affect the stability of a product. Thus, after ascertaining that a plastic from one bottle manufacturer is satisfactory for the product, the drug or toiletry manufacturer should insist that the bottle manufacturer does not alter the components of the plastic bottle formulation or in any way alter the blow-molding process for fabricating the container.

## *Leaching*

Since most plastic containers have one or more ingredients added in small quantities to stabilize or impart a specific property to the plastic, the prospect of leaching, or migration from the container to the drug product, is present. Problems may arise with plastics when coloring agents in relatively small quantities are added to the formula. Particular dyes may migrate into a parenteral solution and cause, a toxic effect. Release of a constituent from the plastic container to the drug product may lead to drug contamination and necessitate removal of the product from the market.

## *Sorption*

This process involves the removal of constituents from the drug product by the packaging material. Sorption may lead to serious consequences for drug preparations in which important ingredients are in solution. Since drug substances of high potency are administered in small doses, losses due to sorption may significantly affect the therapeutic efficacy of the preparation. A problem commonly encountered in practice is the loss of preservatives.

These agents exert their activity at low concentration, and their loss through sorption may be great enough to leave a product unprotected against microbial growth.

Factors that influence characteristics of sorption from product are chemical structure, pH, solvent system, concentration of active ingredients, temperature, length of contact, and area of contact.

## *Chemical Reactivity*

Certain ingredients that are used in plastic formulations may react chemically with one or more components of a drug product. At times, ingredients in the formulation may react with the plastic. Even micro-quantities of chemically incompatible substances can alter the appearance of the plastic or the drug product.

## *Modification*

The physical and chemical alteration of the packaging material by the drug product is called *modification*. Phenomenon like permeation, sorption, and leaching play a role in altering the properties of the plastic and may also lead to its degradation. Deformation in polyethylene containers is often caused by permeation of gases and vapors from the environment or by loss of content through the container walls. Some solvent systems have been found to be responsible for considerable changes in the mechanical properties of plastics. Oils, for example, have a softening effect on polyethylene; fluorinated hydrocarbons attack polyethylene and polyvinyl chloride. Changes in polyethylene caused by some surface-active agents have been noted. In other cases, the content may extract the plasticizer, antioxidant, or stabilizer, thus changing the flexibility of the package. Polyvinyl chloride is an excellent barrier for petroleum solvents, but the plasticizer in polyvinyl chloride is extracted by solvents. This action usually leaves the plastic hard and stiff. Sometimes, this effect is not immediately perceptible because the solvent either softens the plastic or replaces the plasticizer; later, when the solvent evaporates, the full stiffening effect becomes apparent.

Plastic containers, when used for emulsion preparations, must be thoroughly evaluated for physical and chemical changes of the emulsion, as well as for physical changes in the container. It has been reported that certain materials in an emulsion have a tendency to migrate toward the polyethylene

wall, causing either a change in the emulsion or a collapse in the container. Since polyethylene has a tendency toward elastic recovery, air is continuously drawn into the plastic container, increasing the chance of oxidation and drying out of the preparation. In the case of an emulsion, the air can cause the emulsion to break down, owing to dehydration or oxidation of the oil phase. This phenomenon of package-"breathing"-is a major cause of product deterioration during storage in plastic containers. It can also be responsible for the loss of flavor and perfume ingredients from products packaged in plastic containers.

The property of plastics to sorb materials from solution can cause loss of drug, antibacterial agent, or other materials from solution. The curves in Fig. 27.2 show the degree of binding of several commonly used antibacterial preservatives when solutions of these preservatives are in contact with the nylon barrel of a disposable syringe. In nearly all instances, after one week's storage at 30°C, over 60% of the preservative is bound. When these same preservatives have been evaluated with polyethylene and polystyrene barrels, essentially no binding has been found. This clearly illustrates the importance of selecting the correct plastic material for a particular use.

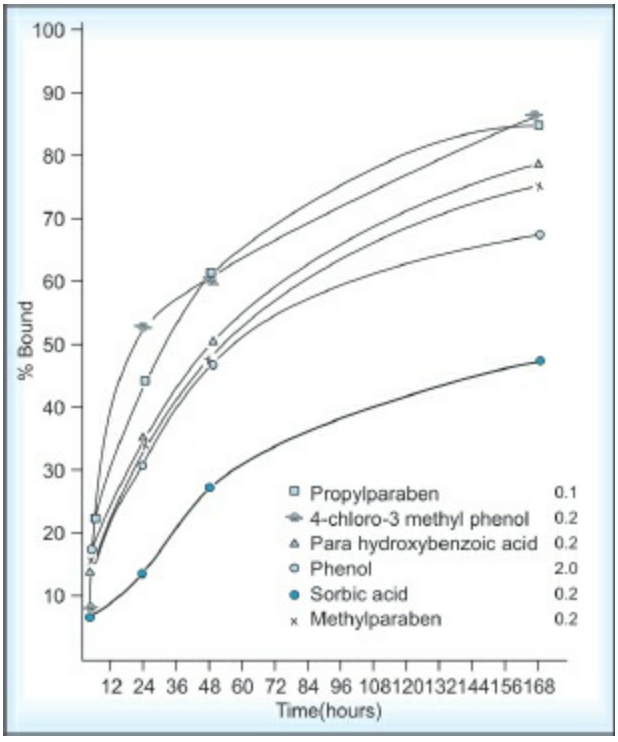



**Fig. 27.2:** Binding of antibacterial preservatives by nylon as a function of time

Stability tests performed on Vioform Lotion packaged in plastic containers showed that the Vioform was being sorbed by the container. Lining the container with an epoxy resin eliminated this problem. This is illustrated by the data summarized in Table 27.7, which presents reflectance data of the inside wall of the lined and unlined container. A decrease in the reflectance of the container wall is indicative of vioform sorption by the plastic since a yellow surface reflects less light than a white surface. This is further substantiated by the increase in reflectance of the lotion with storage, indicating that the lotion is becoming lighter because of the loss in vioform content.

**Table 27.7:** Sorption of vioform by the plastic container from the lotion when stored at 40°C

| Time (weeks) | Reflectance in % | | | |
|---|---|---|---|---|
| | Unlined container | | Lined container | |
| | Lotion | container | Lotion | container |
| 0 | 60.0 | 55.0 | 63.0 | 54.5 |
| 1 | 60.0 | 48.0 | 61.5 | 53.0 |
| 2 | 61.0 | 49.5 | 63.0 | 53.0 |
| 3 | 65.0 | 50.0 | 61.5 | 54.0 |
| 4 | — | 49.0 | 63.0 | 53.0 |

Although the epoxy lining was effective in eliminating the problem with vioform lotion, it was not effective in preventing the sorption of the antibacterial agent, phenylmercuric acetate, from solution as illustrated by the data in Table 27.8.

**Table 27.8:** Loss of phenylmercuric acetate from solution stored in lined and unlined polyethylene bottles

| Time (weeks) | % Preservative lost | | | |
|---|---|---|---|---|
| | Unlined container | | Epoxy-lined container | |
| | 25°C | 40°C | 25°C | 40°C |
| 0 | 0 | 0 | 0 | 0 |
| 1 | 10.0 | 29.3 | 12.4 | 24.6 |
| 2 | 11.5 | 38.5 | — | 27.0 |
| 3 | 20.0 | 35.8 | 15.4 | 37.7 |
| 4 | 20.0 | 46.2 | 23.9 | 43.9 |
| 8 | 33.2 | 57.7 | 27.7 | 58.5 |
| 12 | 33.7 | 70.0 | 33.2 | 68.8 |

It is evident from the results presented in Tables 27.7 and 27.8 that linings are not always effective and must be evaluated separately for each product. The ineffectiveness of the lining could be attributed to sorption of the lining itself by the preservative or to the presence of imperfections, such as pinholes in the lining, which permit penetration of the lining to the plastic.

## Selection of Proper Material

Some of the items to be considered in choosing the plastic material are listed in Tables 27.9 and 27.10.

**Table 27.9:** Plastic bottle chart

| Properties | Acrylic multipolymer | Nitrile polymers | Polyethylene | | PP | Polystyrene | SAN | Vinyl (PVC) |
|---|---|---|---|---|---|---|---|---|
| | | | Low density | High density | | | | |
| Resin density | 1.09–1.14 | 1.10–1.17 | 0.91–0.925 | 0.95–0.96 | 0.89–0.91 | 1.0–1.1 | 1.07–1.08 | 1.2–1.4 |
| Clarity | Clear | Clear | Hazy transparent | Hazy transparent | Clear | Clear | Clear | Clear |
| Water absorption | Moderate | Moderate to low | Low | Low | Low | Moderate to high | High | Low |
| Permeability to: | | | | | | | | |
| Water | High | Moderate | Low | Very low | Very low | High | High | Moderate to low |
| Oxygen | Low | Very low | High | Moderate to high | Moderate to high | High | High | Low |
| $CO_2$ | Moderate | Very low | High | Moderate to high | Moderate to high | High | High | Low |
| Resistance to: | | | | | | | | |
| Acids | Poor to good | Poor to good | Fair to very good | Fair to very good | Fair to very good | Fair to good | Fair to good | Very good |
| Alcohol | Fair | Fair | Good | Good | Good | Poor | Poor | Very good |
| Alkalies | Poor to | Good | Good | Good | Very good | Good | Good | Good |
| Mineral oil | Good | Very good | Poor | Fair | Fair | Fair | Fair | Good |
| Solvents | Poor | Fair | Good | Good | Good | Poor | Poor | Fair |
| Heat | Good | Fair | Poor | Fair | Good | Fair | Fair | Fair to poor |
| Cold | Poor | Poor | Excellent | Excellent | Poor to fair | Poor | Poor | Very poor |
| Sunlight | Good | Good | Fair | Fair | Fair | Fair to poor | Fair to poor | Good |
| High humidity | Fair | Excellent | Excellent | Excellent | Excellent | Excellent | Excellent | Excellent |
| Stiffness | Moderate to high | Moderate to high | Low | Moderate | Moderate to high | Moderate to high | Moderate to high | Moderate to high |
| Resistance to: | | | | | | | | |
| Impact | Poor to good | Good | Excellent | Good | Fair to good | Poor to good | Poor to good | Fair to excellent |
| Unit cost | Moderate | High | Low | Low | Low | Low | Moderate | Moderate |
| Typical uses | Foods, drugs, cosmetics | Foods, drugs, carbonated beverages, cosmetics, household chemicals, aerosols | Cosmetics, personal, products, foods | Detergents, bleaches, milk, other foods; industrial cleaning powders; drugs, cosmetics | Drugs, cosmetics, syrups, juices | Dry drugs, petroleum, jellies | Dry drugs | Shampoos, bath oils detergents; whiskey, wine; floor waxes; vinegar salad oil |

**Table 27.10:** Plastics used for drug bottle packaging

| Material (density) | Inertness | Water vapor rate g/100 Sq in/24 hr | Oxygen perm. cc/100 sq. in | Bottle cost relative to HDPE | Color clarity |
|---|---|---|---|---|---|
| High density PE (0.955) | Outstanding | 0.5 | 120 | 1 | Colorless, translucent |
| Low density PE (0.920) | Excellent | 1.1 | 450 | 1.5 | Colorless, hazy |
| Polystyrene (1.05) | Very poor | 10.0 | 380 | 1.1 | Colorless, clear |
| Rigid PVC (1.35) | Poor | 2.7 | 15 | 2 | Colorless, clear |
| Polypropylene (0.90) | Good to excellent | 0.4 | 230 | 1.3 | Colorless, clear |
| Nylon 6, 10 (1.10) | Good | 2.9 | 3 | 3.0 | Colorless, hazy |
| Polycarbonate (1.20) | Poor | 14.0 | 230 | 3.5 | Colorless, clear |
| Acrylic multi-polymers (1.10) | Fair | 11.5 | 30 | 2.1 | Colorless, hazy |

| Material (density) | Toughness/ Impact | Advantages | Disadvantages |
|---|---|---|---|
| High density PE (0.955) | Excellent | Inertness, low cost, low WVT, toughness | Semi-opaque, transfer of taste ingredients, high dilute solution absorption |
| Low density PE (0.920) | Excellent | Squeeze property, inertness, WVT, toughness | Flexibility, relatively poor barrier to nonpolars, high WVT |
| Polystyrene (1.05) | Poor | Clarity, stiffness, low cost | High WVT, susceptibility to cracking, poor impact |
| Rigid PVC (1.35) | Fair | Clarity, stiffness, $O_2$ barrier, retention of nonpolar materials | 10–12 ingredients are possible; difficult processability, solvent susceptibility |
| Polypropylene (0.90) | Good (poor at 40°F and | Inertness, low cost, ESCR resistance | Low-temperature brittleness, tendency to unzip, highly stabilized content |
| Nylon 6, 10 (1.10) | Excellent | Good barrier for nonpolars, tough, good $O_2$ barrier | Cost, water absorption, borderline for water-based materials |
| Polycarbonate (1.20) | Outstanding | Very tough, clear, good oxygen barrier | Cost, susceptibility to solvent cracking, poor WVT, poor barrier |
| Acrylic multi-polymers (1.10) | Good | Clarity, fair oxygen barrier, good for oils | Poor WVT, blushes, poor barrier |

Also to be taken into consideration is the potential effect of wall thickness, which relates to barrier requirements of the package for its intended shelf-life.

## METAL

In addition to glass and plastic, certain metals are also employed as containers for pharmaceutical products. Like glass, metal is nearly totally impermeable to gas and water. In addition, metal containers are extremely strong and shatterproof. For applications requiring malleability such as collapsible tubes, metal offers relatively easy manufacturing and use. Metals can also be fashioned into more complex delivery systems such as dry powder inhalers, metered-dose inhalers, aerosol containers, and even ready-to-use needles.

Dispersible systems, having a consistency of a soft paste, gel, cream, or ointment, can be conveniently packaged into collapsible metal tubes. Metals commonly used for such tubes are tin, plastic-coated tin, tin-coated lead, aluminum, and steel. All these metals carry special advantages and uses. Steel and aluminium containers are useful for bulk packing where very high strength is required.

## Tin

Tin containers are preferred for food, pharmaceuticals, or any product for which purity is of paramount consideration. Tin is frequently used in the production of aerosol cans by electroplating it onto sheet steel to improve corrosion resistance and facilitate soldering. Tin is the most chemically inert of all collapsible tube metals. It offers a good appearance and compatibility with a wide range of products. Tin and tin-coated tubes are usually employed because of their unreactive properties, although it has been reported that tin tubes can be corroded by chlorides or acid conditions. Vinyl and cellulose lacquers are applied to tin to increase their utility.

## Aluminum

Aluminum foil can be formed into rigid containers, semirigid containers, blister construction, or laminates. Aluminum tubes offer significant savings in product shipping costs because of their lightweight. They provide the attractiveness of tin at somewhat lower cost. Coated and uncoated aluminum tubes are being used, but are not always satisfactory. It has been reported that aluminum reacts with fatty alcohol emulsions to form a white encrustation. Such tubes were also found to be unstable for mercury-containing compounds. It has been stated that uncoated aluminum tubes are deleteriously affected when used for preparations outside the pH range of 6.5 to 8.0. The application of an epoxy lining to the internal surfaces of aluminum tubes was found to make them more resistant to attack.

## Lead

Lead has the lowest cost of all metals and is widely used for nonfood products such as adhesives, inks, paints, and lubricants. Lead should never be used alone for anything taken internally because of the risk of lead poisoning. With internal linings, lead tubes are used for products like fluoride toothpaste.

## Linings

If the product is not compatible with bare metal, the interior can be flushed with wax-type formulations or with resin solutions, although the resins or lacquers are usually sprayed on. An epoxy lining costs about 25% more than the uncoated.

Wax linings are most often used with water-base products in tin tubes, and phenolics, epoxides, and vinyls are used with aluminum tubes, giving better protection than wax, but at a higher cost. Phenolics are most effective with acid products; epoxides protect better against alkaline materials.

The primary disadvantages of metals relate to their cost and quality control. Metals are inherently more expensive to purchase. Metals are also prone to the development of “pinhole” defects during manufacturing that can drastically compromise their barrier properties. Not only can these defects be deleterious to the container, but they can also compromise the quality of the drug itself. Drug product should be monitored to assure that no metallic packaging defects have been transferred to the preparation—especially in ophthalmic drug products. Much like copolymerization of plastics, metals can be alloyed to enhance their characteristics as well.

## RUBBER

Rubber of varying composition is used in pharmaceuticals and biologicals as stoppers, cap liners, and parts of dropper assemblies. A major use of the rubber stopper is that of a closure for multiple-dose vial solutions for injections. Natural rubber is resistant to dilute mineral acids, alkalies and salts, but oxidising media, oils, benzene and ketones attack it. Hard rubber, made by adding 25% or more of sulphur to natural or synthetic rubber, is frequently used for handling acids, especially dilute aqueous solutions. Synthetic rubbers such as polychloroprene (Neoprene), polysiloxanes (Silicon rubber), polyisoprene, polybutadiene, chlorinated polyethylene elastomer, epichlorhydrin elastomers, styrene-butadiene rubber, nitrile rubbers (Buna N), ethylene propylene diene terpolmers and polyurethane are becoming of increasing importance due to their superiority to natural rubber in many properties such as resistance to oxidation, solvents, oils and many chemicals.

The problems that can be encountered by having rubber stoppers in contact with the liquid in the vial are the sorption of active ingredient, antibacterial preservative, or other materials into the rubber, and the extraction of one or more components of the rubber into the vial solution.

The data in Table 27.11 illustrate the influence of various rubber closures on the loss of the antimicrobial preservative chlorobutanol from vial solutions. It is evident from these data that the closures have a marked deleterious effect on preservative concentration at all temperatures. At 60°C, after 12 weeks of storage, the vials stoppered with neoprene closures and stored in an inverted position show only 8.5% residual concentration of chlorobutanol, whereas the ampuls under the same conditions of storage contain 72.4%.

**Table 27.11:** Influence of various closures on residual concentration of chlorobutanol in vial solutions after storage

| Temp (°C) | Storage (wk) | Ampule (%) | Stopper composition: Natural rubber Upright (%) | Natural rubber Inverted (%) | Neoprene polymer Upright (%) | Neoprene polymer Inverted (%) | Butyl polymer Upright (%) | Butyl polymer Inverted (%) |
|---|---|---|---|---|---|---|---|---|
| 25 | 12 | 100 | 81.00 | 78.7 | 87.3 | 81.0 | 68.1 | 63.8 |
| 40 | 2 | 100 | 76.6 | 63.8 | 61.7 | 59.6 | 68.1 | 61.7 |
| — | 12 | 97.9 | 74.5 | 57.5 | 53.2 | 46.8 | 66.0 | 66.0 |
| 50 | 2 | 97.9 | 59.6 | 57.5 | 57.5 | 46.8 | 66.0 | 61.7 |
| — | 12 | 91.5 | 57.5 | 42.5 | 46.8 | 38.3 | 59.6 | 57.5 |
| 60 | 2 | 95.8 | 57.5 | 53.2 | 51.0 | 46.8 | 51.0 | 48.9 |
| — | 12 | 72.4 | 57.5 | 42.5 | 29.8 | 8.5 | 46.8 | 46.8 |

The curves in Fig. 27.3 show that extractives from the natural rubber closures used to stopper vials containing water of injection were leached into the water when the vials were autoclaved. It is evident from the absorption spectra of the extractives in the vial solutions that even normal autoclaving at 115°C, 10 psi for 30 min causes appreciable extractive to enter the water for injection. Subsequent storage of these vials at room temperature for the shelf-life of the pharmaceuticals, which at present is an average of three years, would no doubt cause further leaching of extractives in solution.

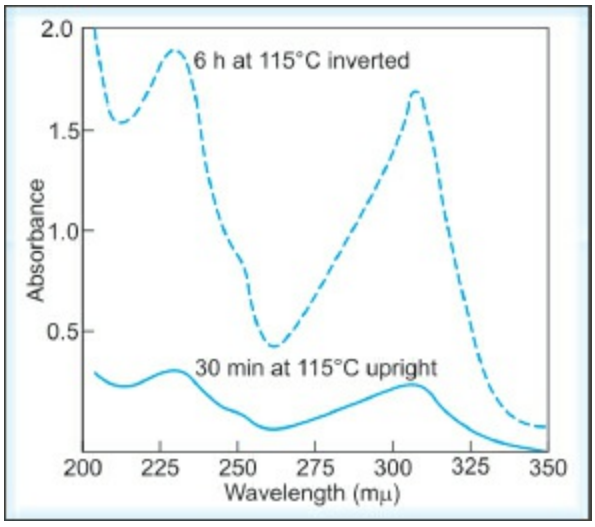



**Fig. 27.3:** Leaching of extractives from rubber stoppers on multiple dose vial solutions

If an epoxy lining is applied to rubber stoppers, a considerable reduction results in the amount of extractive leached from the stopper by the water for injection in the vials, but there is essentially no effect on the sorption of preservative from solution. However, the use of Teflon-coated rubber

stoppers essentially prevents sorption and leaching of the rubber stopper.

The presence of rubber closure extractives in the vial solutions could interfere with the chemical analysis of the active ingredient, affect the toxicity or pyrogenicity of the injectable preparation, interact with the drug or preservative to cause inactivation or loss of stability, and cause physical instability to the preparation, owing to the presence of particulate matter in the solution.

# PACKAGING DESIGN

## Collapsible Tubes

The collapsible tube is an attractive container that permits controlled amounts to be dispensed easily, and has good reclosure, and provides adequate protection of the product. The risk of contamination of the portion remaining in the tube is minimal, because the tube does not “suck back”. It is lightweight and unbreakable, and it lends itself to highspeed automatic filling operations. These tubes are perfect package for viscous products such as dentifrices, pharmaceutical creams and ointments cosmetic creams and gels.

### *Metal*

Any ductile metal that can be worked cold is suitable for collapsible tubes, but the most common ones in use are tin (15%), aluminum (60%), and lead (25%). Tin is the most expensive, and lead is the cheapest. Since tin is the most ductile of these metals, small tubes are often made more cheaply of tin even though the metal cost is higher (mostly used for ophthalmic creams and ointments). Laminates of tin-coated lead provide the appearance and oxidation resistance of straight tin at lower prices. The tin that is used for this purpose is alloyed with about 0.5% copper for stiffening. When lead is used, about 3% antimony is added to increase hardness. Aluminum hardens when it is formed into a tube, and must be annealed to give it the necessary pliability. Aluminum also hardens during use, sometimes causing tubes to develop leaks.

### *Plastic*

This distinctive style of package, like its counterpart, the metal collapsible tube, excels in functional characteristics. Plastic tubes have a number of inherent practical advantages over other containers or dispensers. They are (1) low in cost, (2) light in weight, (3) durable, (4) pleasant to touch, (5) flexible, facilitating product dispensing, (6) odorless and inert to most chemicals, (7) unbreakable, (8) leak proof, (9) able to retain their shape throughout their use and (10) high machinability. In addition, (11) they have a unique “suck-back” feature, which prevents product ooze. If too much product is dispensed with one squeeze, relaxation of hand pressure permits the product to be sucked back into the tube. If this feature is undesirable for

fear of contamination, plastic tubes designed to avoid suck-back are available. Thus, the suck-back feature of plastic tubes can be an advantage or a disadvantage. When the tube is partly empty, however, this feature is a nuisance, because the air must be expelled before the product can be dispensed.

The sidewall of a plastic tube is extruded under heat and pressure as continuous hoselike tubing and then is cut to length. The neck and shoulder are molded and joined to the tube in a separate automatic process. The tube is then decorated, mostly by offset printing. The tubes are capped with the bottoms remaining open for filling, after which they are heat-sealed to become as strong as any other part of the tube.

The most common types of material currently employed in plastic tubes are LDPE and HDPE. The former is less expensive and used more extensively. HDPE offers more protection than the uncoated low density type. Coated high density polyethylene is only slightly more protective than the coated low density type, because in both instances, the coatings serve as a prime barrier.

Other materials include PP and vinyl, but relatively high costs and various technical problems have limited their use. Both HDPE and PP are much stiffer than LDPE for tube sidewalls, making LDPE more suitable. The neck of a plastic container must be compatible with the shoulder to produce a good bond. Neck of LDPE can be headed with shoulder of LDPE or HDPE, but PP neck must be headed with PP shoulder.

## *Laminations*

Permeation problems associated with plastic tubes, and corrosion and breakage problems experienced with metal tubes, have led to the emergence of a third type of collapsible tube, the *laminated* tube. This tube, constructed of a lamination containing several layers of plastic, paper, and foil, is fabricated from flat, printed stock (Fig 27.4). This lamination, which is specifically tailored to the product requirements, is welded into a continuous tube by heat-sealing the edges of the lamination together in a machine called a “sideseamer”. The tube is cut to length, and the head is injection-molded onto the tube. The head is typically molded of LDPE. Since some permeation through this molded head is possible, a head insert, made of urea formaldehyde, can be molded into the head to reduce product permeation.

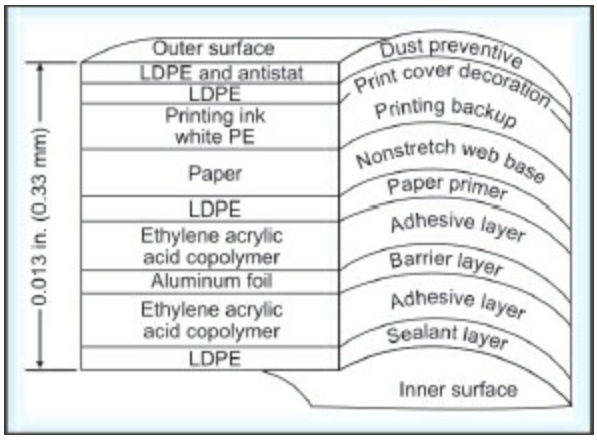



**Fig. 27.4:** Structure of laminate tube

Although not as impermeable as metal tubes, the laminated tube does provide a satisfactory level of protection for a number of products. Laminated tubes initially served the dentifrice market but today are also found in pharmaceuticals, depilatories, hand creams, hair care products, and denture adhesives.

## Closures

The closure is normally the most vulnerable and critical component of a container insofar as stability and compatibility with the product are concerned. An effective closure must prevent the contents from escaping and allow no substance to enter the container. The adequacy of the seal depends on a number of things, such as the resiliency of the liner, the flatness of the sealing surface on the container, and most important, the tightness or torque with which it is applied. In evaluating an effective closure system, the major considerations are the type of container, the physical and chemical properties of the product, and the stability-compatibility requirements for a given period under certain conditions.

Closures are available in five basic designs (1) screw-on, threaded, or lug, (2) crimp-on (crowns), (3) press-on (snap), (4) roll-on and (5) friction. Many variations of these basic types exist, including vacuum, tamperproof, safety, child-resistant, and linerless types, and dispenser applicators.

### *Threaded Screw Cap*

When the screw cap is applied, its threads engage with the corresponding threads molded on the neck of the bottle. A liner (generally a plastisol inner gasket) in the cap, pressed against the opening of the container, seals the product in the container by overcoming sealing surface irregularities, and provides resistance to chemical and physical reaction with the product being sealed.

The screw cap is commonly made of metal or plastics. The metal is usually tinplate or aluminum, and in plastics, both thermoplastic and thermosetting materials are used. Metal caps are usually coated on the inside with an enamel or lacquer for resistance against corrosion.

Almost all metal crowns and closures are made from electrolytic tinplate, a tin-coated steel on which the tin is applied by electrolytic deposition.

### *Lug Cap*

The lug cap is similar to the threaded screw cap and operates on the same principle. It is simply an interrupted thread on the glass finish, instead of a continuous thread. It is used to engage a lug on the cap sidewall and draw the

cap down to the sealing surface of the container. Unlike the threaded closure, it requires only a quarter turn.

The lug cap is used for both normal atmospheric pressure and vacuum pressure closing. The cap is widely used in the food industry because it offers a hermetic seal and handles well in sterilization equipment and on production lines. The lug cap design allows application and removal with a one-quarter turn. This not only provides consumer convenience, but also quick capping.

## *Crown Caps*

These are friction fitting closure types. This style of cap is commonly used as a crimped closure for beverage bottles and has remained essentially unchanged for more than 50 years. Crown is made of tin free steel and tin plate. The crown has a short skirt with 21 flutes that are crimped into locking position on the bottle head. Crowns also contain compressible lining material for efficient closing.

## *Roll-on Closures*

The aluminum roll-on cap can be sealed securely, opened easily, and resealed effectively. It finds wide application in the packaging of food, beverages, chemicals, and pharmaceuticals. The roll-on closure requires a material that is easy to form, such as aluminum or other light-gauge metal.

Resealable, nonresealable, and pilfer-proof types of the roll-on closure are available for use on glass or plastic bottles and jars. The packager purchases these closures as a straight-sided, threadless shell and forms the threads on the packaging line as an integral part of the filling operation.

The roll-on technique allows for dimensional variation in the glass containers; each roll-on closure precisely fits a specific container.

## *Pilfer-proof Closures*

The pilfer-proof closure is similar to the standard roll-on closure except that it has a greater skirt length. This additional length extends below the threaded portion to form a bank, which is fastened to the basic cap by a series of narrow metal “bridges”.

When the pilfer-proof closure is removed, the bridges break, and the bank remains in place on the neck of the container. The user can reseal the closure,

but the detached band indicates that the package has been opened. The torque necessary to break the bridges and remove the cap is nominal.

## *Non-reusable Roll-on Closures*

In some packaging applications a reusable cap is not desired. Non-reusable caps require unthreaded glass finishes. The skirts of these closures are rolled under retaining rings on the glass container and maintain liner compression. Closures of this type have tear-off tabs that make them tamperproof and pilfer-proof.

## *Plastic Closures*

The two basic types of plastics generally used for closures are thermosetting and thermoplastic materials. They differ greatly in physical and chemical properties, and fundamentally different manufacturing methods are used for each type.

**Thermosetting resins:** Phenolic and urea thermosetting plastic resins are widely used in threaded closures. The thermosetting plastic first softens under heat and then cures and hardens to a final state. Shaping must occur in the first stage of softening, because after curing there is no further mobility, even upon reapplication of heat and pressure. During the molding process, thermosetting materials undergo a permanent chemical change, and unlike thermoplastic materials, they cannot be reprocessed. Since parts that are improperly molded must therefore be discarded, thermosetting materials are usually fabricated by compression molding. The manufacturing process is relatively slow, but allows better control and quick response to change in temperature and material flow.

**Phenolics:** Phenolic molding compounds are available in different grades and in dark colors, usually black or brown. Phenolics are used when a hard sturdy piece is needed, and when dark colors can be tolerated.

Rigidity, heat, chemical resistance, and strength are the outstanding properties of the phenolics. Color limitation is the main drawback, although coatings are available at a premium price. As a closure, the phenolic can withstand the torquing forces of the capping machines and maintains a tight seal over a long period of time.

The phenolics are resistant to some dilute acids and alkalies and are attacked by others, especially oxidizing, acids. Although organic acids and

reducing acids usually do not have any effect but strong alkalies decompose phenolics.

**Urea:** This thermosetting resin is a hard translucent material that takes coloring well. It is more expensive than the phenolics, but the heat resistance and other properties of urea make it suitable for premium items. Elegant colors are obtainable with urea because the translucency gives brightness and color depth. Urea plastic is available in an unlimited range of colors and is a hard, brittle material that is odorless and tasteless. Being a thermosetting plastic, urea can withstand high temperatures without softening, but it chars at about 390°F. It absorbs water under wet conditions, but such absorption has no serious effect on the plastic.

Urea is not affected by any organic solvent, but it is affected by alkalies and strong acids. It has good resistance to all types of oils and greases. Although urea can withstand elevated temperature but it cannot be steam-sterilized. Parts may shrink as much as 0.003 inches after molding.

**Thermoplastic resins:** Since their introduction, thermoplastics have become widely used in the manufacture of closures. Polystyrene, polyethylene, and polypropylenes are the materials used in 90% or more of all thermoplastic closures. Each material has specific performance advantages, and the particular resin used depends on the physical and chemical properties desired for the application and on the particular product being packaged.

## Closure Liners

A liner may be defined as any material that is inserted in a cap to make a seal between the closure and the container. Liners are usually made of a resilient backing and a facing material. The backing material must be soft enough to take up any irregularities in the sealing surface and elastic enough to recover some of its original shape when removed and replaced. It is usually glued into the cap with an adhesive, or the cap can be made with an undercut, so that the liner snaps into place and is free to rotate.

### *Factors in Selecting a Liner*

Many factors have to be considered before an effective liner can be selected (Tables 27.12 and 27.13). The most important consideration is that the liner be chemically inert with its product, so that the latter is protected against any possible change in purity or potency.

**Table 27.12:** Factors in selecting a liner

1. Compatibility (chemical resistance)
2. Appearance
3. Gas- and vapor-transmission rates-WVTR, oxygen, $CO_2$, etc.
4. Removal torque
5. Heat resistance
6. Shelf-life
7. Economics

**Table 27.13:** Permeation rate of liner facings

| Facing | Water | Alcohol |
|---|---|---|
| Polyethylene, 2 mil | 0.07 | 0.06 |
| Saran, 75 gauge | 0.07 | 0.08 |
| Aluminum foil, 1 mil | 0.04 | 0.12 |
| Tinfoil, 1mil | 0.09 | 0.20 |
| Polyester, 50 gauge | 0.12 | 0.10 |

| | | |
|---|---|---|
| Vinylite | 0.20 | 0.03 |
| Solvent-resistant | 0.23 | 0.61 |
| Yellow oil | 0.28 | 0.85 |

Gas and vaportransmissionrates areusually relative and depend chiefly on the shelf life required for the product. If the period between packing and consumer use is expected to be long, low transmission rates are must. Representative permeation rates are presented in Table 27.13.

## *Homogeneous Liner*

These one-piece liners are available either as a disk or as a ring of rubber or plastic. Although they are more expensive and more complicated to apply, they are widely used for pharmaceuticals because their properties are uniform and they can withstand high-temperature sterilization.

## *Heterogeneous or Composite Liners*

These are composed of layers of different materials chosen for specific requirements. Generally, the composite liner consists of two parts: A face and a back. Usually, the face is in contact with the product, and the back provides the cushioning and sealing properties required.

## Stoppers

Rubber is used in the pharmaceutical industry to make stoppers, cap liners, and bulbs for dropper assemblies. The rubber stopper is used primarily for multiple-dose vials and disposable syringes. The rubber polymers most commonly used are natural, neoprene, and butyl rubber. In the manufacture of rubber closures, certain performance expectations require certain ingredients. The typical ingredients of a rubber closure are rubber, vulcanizing agent, accelerator/activator, extended filler, reinforced filler, softener/plasticizer, antioxidant, pigment, special components and waxes. Since the composition of rubber stoppers is complex and the manufacturing process complicated, it is common to encounter problems with certain rubber formulas. For example, when the rubber stopper comes in contact with parenteral solution, it may absorb active ingredient, antibacterial preservative, or other materials, and one or more ingredients of the rubber may be extracted into the liquid. These extractives could (1) interfere with the chemical analysis of the active ingredient, (2) affect the toxicity or pyrogenicity of the parenteral preparation, (3) interact with the drug preservative to cause inactivation and (4) affect the chemical and physical stability of the preparation so that particulate matter appears in the solution.

## TAMPER-RESISTANT PACKAGING

In 1982, the vulnerability of over-the-counter (OTC) or nonprescription drug products to malicious adulteration was dramatically and tragically demonstrated with the tainting of Tylenol capsules with cyanide, which led to the deaths of several people. Public concerns about the safety of drug packaging led to various proposals by local municipalities for legislation regarding tamper-resistant packaging. To respond as quickly as possible to the safety concerns raised regarding OTC pharmaceutical packaging, and to assist the FDA in providing a consistent national standard governing tamper-resistant packaging, the Proprietary Association, a trade association for the Proprietary Pharmaceutical Industry, provided recommendations to the FDA, which were subsequently incorporated into FDA regulation 21 CFR parts 211, 314, and 700, which covers tamper-resistant packaging of OTC drugs.

The legislation enacted has had one of the greatest impacts on the packaging of pharmaceuticals in recent history. The requirement for tamper-resistant packaging is now one of the major considerations in the development of packaging for pharmaceutical products. As defined by the FDA, “a tamper-resistant package is one having an indicator or barrier to entry which, if breached or missing, can reasonably be expected to provide visible evidence to consumers that tampering has occurred. Tamper-resistant packaging may involve immediate-container/closure systems or secondary-container/carton systems or any combination thereof intended to provide a visual indication of package integrity when handled in a reasonable manner during manufacture, distribution, and retail display”.

The visual indication is required to be accompanied by appropriate illustrations or precautionary statements to describe the safeguarding mechanism to the consumer. To reduce the possibility that the security mechanism can be restored after tampering, the FDA also requires either that the tamper-resistant feature be designed from materials that are generally not readily available (e.g. an aerosol system), or that barriers made from readily obtainable materials carry a distinctive design or logo that cannot be readily reproduced by an individual attempting to restore the package.

The following package configurations have been identified by the FDA as examples of packaging systems that are capable of meeting the requirements of tamper-resistant packaging as defined by FDA regulation 21 CFR parts

211, 314, and 700:

1. Film wrappers
2. Blister package
3. Strip package
4. Bubble pack
5. Shrink seals and bands
6. Foil, paper, or plastic pouches
7. Bottle seals
8. Tape seals
9. Breakable caps
10. Sealed tubes
11. Aerosol containers
12. Sealed cartons

The use of any of these concepts does not necessarily constitute compliance with the FDA ruling. The manufacturer must determine that the particular package concept provides tamper resistance for the manufacturer's specific product. By the same token, manufactures are not limited to the packaging concepts listed here, but have the flexibility of pursuing new packaging technologies that meet the objectives of the FDA ruling.

## Film Wrapper

Film wrapping has been used extensively over the years for products requiring package integrity or environmental protection. Although film wrapping can be accomplished in several ways and varies in configuration from packaging equipment to packaging equipment, it can be generally categorized into the following types:

- End-folded wrapper
- Fin seal wrapper
- Shrink wrapper

### *End-folded Wrapper*

The end-folded wrapper is formed by pushing the product into a sheet of over wrapping film, which forms the film around the product and folds the edges in a gift-wrap fashion as indicated in Fig. 27.5A. The folded areas are sealed by pressing against a heated bar. Because of the overlapping folding sequence of the seals, the film used must be heat-sealable on both surfaces. Materials commonly used for this application are cellophane and polypropylene. Cellophane, which is regenerated cellulose, is not inherently heat-sealable but requires a heat-seal coating to impart heat-sealing characteristics to the film. This is usually accomplished by coating the cellophane with either polyvinylidene chloride (PVdC) or nitrocellulose. Since PVdC also provides a durable moisture barrier, PVdC coated cellophane is often used for the over wrapping of products that are sensitive to moisture. Cellophane offers excellent machinability and crystal clarity, and for many years was the only clear film available for this type wrapping. In the early 1970s, polypropylene came onto the scene and represented a lower-cost alternative to cellophane. Because of the different handling characteristics of polypropylene, initial problems were experienced in obtaining satisfactory machinability. Machine modifications and the redesign of equipment led to the gradual replacement of cellophane by polypropylene, so that polypropylene now dominates the marketplace for this application. Heat-sealing characteristics are imparted to polypropylene by the use of heat-sealable acrylic coatings or by the addition of heat-sealable modifiers to the resin itself.

To be tamper-resistant, the over-wrap must be well sealed and must be printed or uniquely decorated to exclude the possibility of having an alternate overwrap substituted in its place. The printed surface of the carton being overwrapped may also be coated with a heat-sensitive varnish, which causes the overwrap to bond permanently to the paperboard carton during the sealing of the overwrap. The removal of the overwrap would accordingly deface the carton, making the carton unsuitable for re-use.

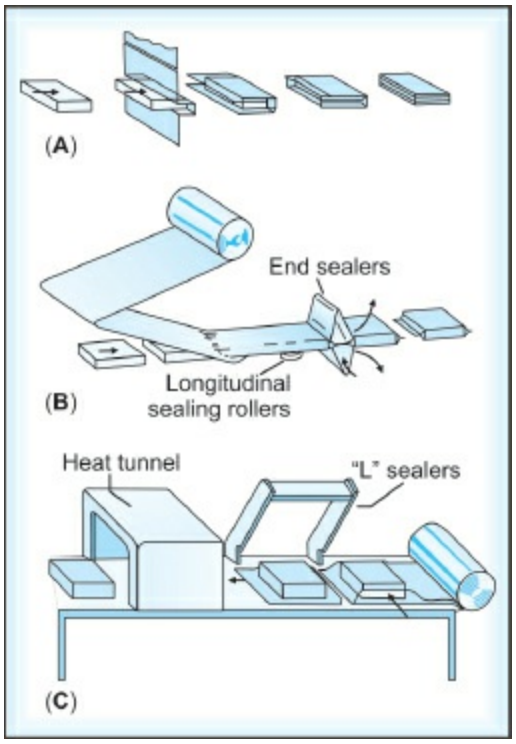
## *Fin Seal Wrapper*

Unlike the end-folded wrapper configuration, fin seal packaging does not require the product to act as a bearing surface against which the overwrap is sealed. The seals are formed by crimping the film together and sealing together the two inside surfaces of the film, producing a “fin” seal (Fig. 27.5B). Since the seals are formed by compressing the material between two heated bars (or nipping the film between pressure rollers) rather than sealing against the package, much greater and more consistent sealing pressure can be applied, and consequently, better seal integrity can be accomplished. For this reason, fin sealing has primarily been used when protective packaging is critical. Since the surface of the heat seal does not come in contact with the heated sealing bars on the packaging equipment, much more tenacious heat sealants, such as polyethylene or Surlyn can be used. With good seal integrity, the overwrap can be removed or opened only by tearing the wrapper.

## *Shrink Wrapper*

Film overwrapping can also be accomplished with the use of a shrink wrapper. The shrink wrap concept involves the packaging of a product in a thermoplastic film that has been stretched and oriented during its manufacture and that has the property of reverting back to its unstretched dimensions once the molecular structure is “unfrozen” by the application of heat. The shrink-wrap concept has a diversity of uses in packaging, one of which is its use as an overwrap. In this case, the shrink film is used in roll form, with the center folded in the direction of winding (Fig. 27.5C). As the film unwinds on the overwrapping machine, a pocket is formed in the center fold of the sheet, into which the product is inserted. An L-shaped sealer seals the remainder of the overwrap and trims off the excess film. The loosely wrapped product is then moved through a heated tunnel, which shrinks the overwrap into a tightly

wrapped unit. The materials commonly used for this application are heat-shrinkable grades of polypropylene, polyethylene, and PVC. Since the various heat-shrinkable grades of film have different physical characteristics such as tear and tensile strength, puncture resistance, and shrinking forces, selection of the particular material used must be based upon specific product considerations so that the shrink wrap provides suitable integrity without crushing or damaging the product. The major advantages of this type of wrapper are the flexibility and low cost of the packaging equipment required.




**Figs 27.5A to C:** Film wrapper systems: (A) End-folded wrapper; (B) Fin seal wrapper; (C) Shrink wrapper

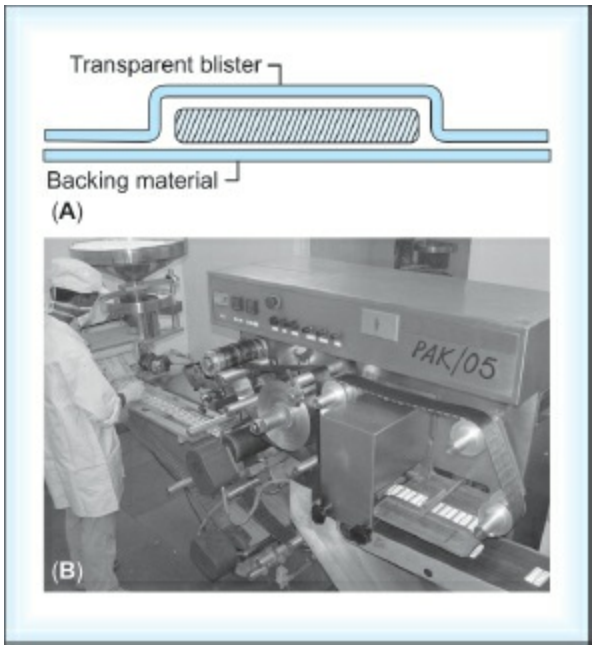
## Blister Package

When one thinks of unit dose in pharmaceutical packaging, the package that invariably comes to mind is the blister package. This packaging mode has been used extensively for pharmaceutical packaging for several good reasons. It is a packaging configuration capable of providing excellent environmental protection, coupled with an esthetically pleasing and efficacious appearance. It also provides user functionality in terms of convenience, child resistance, and now, tamper resistance.

The blister package is formed by heatsoftening a sheet of thermoplastic resin and vacuum-drawing the softened sheet of plastic into a contoured mold. After cooling, the sheet is released from the mold and proceeds to the filling station of the packaging machine. The semi-rigid blister previously formed is filled with product and lidded with a heat-sealable backing material (Fig. 27.6A). The backing material, or lidding, can be of either a push-through or peelable type. For a push-through type of blister, the backing material is usually heat-seal-coated aluminum foil. The coating on the foil must be compatible with the blister material to ensure satisfactory sealing, both for product protection and for tamper resistance. Peelable backing materials have been used to meet the requirements of child-resistant packaging. This type of backing must have a degree of puncture resistance to prevent a child from pushing the product through the lidding and must also have sufficient tensile strength to allow the lidding to be pulled away from the blister even when the lidding is strongly adhered to it. To accomplish this, a material such as polyester or paper is used as a component of the backing lamination. Foil is generally used as a component of backing lamination if barrier protection is a critical requirement; however, metallized polyester is replacing foil for some barrier applications. A peelable sealant compatible with the heat-seal coating on the blister is also required since the degree of difficulty of opening is a critical parameter for child-resistant packaging. The use of peelable backing materials for blister packaging must be carefully evaluated to ensure that peel strengths are sufficient to meet tamper-resistance objectives. Figure 27.6B shows working of an automatic blister packaging machine.

Materials commonly used for the thermoformable blister are PVC, PVC/polyethylene combinations, polystyrene, and polypropylene. For

commercial reasons and because of certain machine performance characteristics, the blisters on most unit dose packages are made of polyvinyl chloride. For added moisture protection, polyvinylidene chloride (saran) or polychlorotrifluoro-ethylene (Aclar) films may be laminated to PVC. The moisture barrier of PVC/Aclar is superior to that of saran-coated PVC, especially under prolonged and extremely humid storage conditions. Listed in Tables 27.14 and 27.15 are several commonly used thermoformable blister materials and a comparison of their protective qualities.




**Figs 27.6A and B:** (A) Blister pack; (B) Automatic blister packaging machine (*Courtesy* of Unicure)

**Table 27.14:** Barrier properties of laminations

| Material | Oxygen transmission* | Water-vapor transmission |
|---|---|---|
| 0.002 Saran/0.006 PVC | 0.6 | 0.092 |
| 0.0015 Aclar/0.002 PE/0.0075 PVC | 1.0 | 0.034 |
| 0.0015 Aclar/0.0075 PVC | 1.1 | 0.035 |
| 0.002 PE/0.0075 PVC | 1.3 | 0.170 |
| 0.0075 PVC | 1.9 | 0.330 |

| | | |
|---|---|---|
| 0.002 PE/0.005 PVC | 2.6 | 0.200 |
| 0.005 PVC | 2.7 | 0.520 |
| 0.001 Nylon | 25.0 | 19.000 |

* cc/24 hr/100 sq in. at 77T, 50% RH.
g/24 hr/100 sq in. at 95°F, 90% RH.

**Table 27.15:** Moisture barrier properties of flexible materials

| Material | Water-vapor transmission* |
|---|---|
| *Aclar (fluorohalocarbon)* | |
| 22A, 1 mil | 0.055 |
| 22A, ½ mil | 0.046 |
| 22C, 1 mil | 0.045 |
| 22C, 2 mil | 0.028 |
| 33C ½ mil | 0.040 |
| 33C, 1 mil | 0.025 |
| 33C, 2 mil | 0.015 |
| Cellulose acetate, 1 mil | 80.000 |
| *Cellophane* | |
| 140K | 0.400 |
| 195K | 0.450 |
| 195M | 0.650 |
| Polyester, 1 mil | 2.000 |
| *Polyethylene* | |
| Low density, 1 mil | 1.300 |
| High density, 1 mil | 0.300 |
| Polypropylene, 1 mil | 0.700 |
| Polyvinyl chloride, 1 mil | 4.000 |
| Rubber hydrochloride, 1.2 mil | 1.000 |
| Saran (PVdC), 1 mil | 0.200 |

| | |
|---|---|
| Two-ply waxed glassine paper | 0.500 |
| Waxed glassine paper | 3.000 |
| Waxed sulfite paper | 4.000 |

* g loss/24 hr/100 sq in./mil at 95°F, 90% RH.

## Strip Package

A strip package is a form of unit dose packaging that is commonly used for the packaging of tablets and capsules. A strip package is formed by feeding two webs of a heat-sealable flexible film through either a heated crimping roller or a heated reciprocating platen. The product is dropped into the pocket formed prior to forming the final set of seals. A continuous strip of packets is formed, generally several packets wide depending on the packaging machine's limitations. The strip of packets is cut to the desired number of packets in length (Fig. 27.7). The strips formed are usually collated and packaged into a folding carton. The product sealed between the two sheets of film usually has a seal around each tablet, with perforations usually separating adjacent packets. The seals can be in a simple rectangular or "picture-frame" format or can be contoured to the shape of the product.

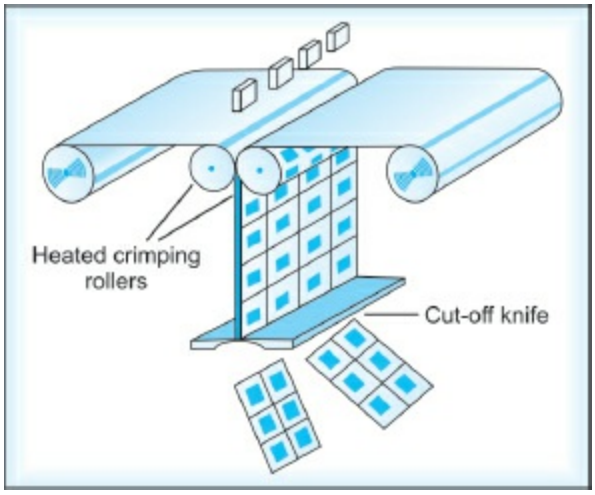



**Fig. 27.7:** Strip packaging system

Since the sealing is usually accomplished between pressure rollers, a high degree of seal integrity is possible. The use of high-barrier materials such as foil laminations or saran-coated films, in conjunction with the excellent seal formation, makes this packaging mode appropriate for the packaging of moisture-sensitive products.

A number of different packaging materials are used for strip packaging. For high-barrier applications, a paper/polyethylene/foil/polyethylene lamination is commonly used. When product visibility is important, a heat-sealable cellophane or a heat-sealable poly ester can be used. Also, the front

and back of the package may use dissimilar materials. The choice of material used depends on both product and equipment requirements.

## Bubble Pack

The bubble pack can be made in several ways but is usually formed by sandwiching the product between a thermoformable, extensible, or heat-shrinkable plastic film and a rigid backing material. This is generally accomplished by heat-softening the plastic film and vacuum-drawing a pocket into the film in a manner similar to the formation of a blister in a blister package. The product is dropped into the pocket, which is then sealed to a rigid material such as heat-seal-coated paperboard. If a heat-shrinkable material is used, the package is passed through a heated tunnel, which shrinks the film into a bubble or skin over the product, firmly attaching it to the backing card.

## Shrink Banding

The shrink band concept makes use of the heat-shrinking characteristics of a stretch-oriented polymer, usually PVC. The heat-shrinkable polymer is manufactured as an extruded, oriented tube in a diameter slightly larger than the cap and neck ring of the bottle to be sealed. The heat-shrinkable material is supplied to the bottler as a printed, collapsed tube, either pre-cut to a specified length or in roll form for an automated operation. The proper length of PVC tubing is slid over the capped bottle far enough to engage both the cap and neck ring of the bottle (Fig. 27.8). The bottle is then moved through a heat tunnel, which shrinks the tubing tightly around the cap and bottle, preventing the disengagement of the cap without destroying the shrink band. For ease of opening, the shrink bands can be supplied with tear perforations.

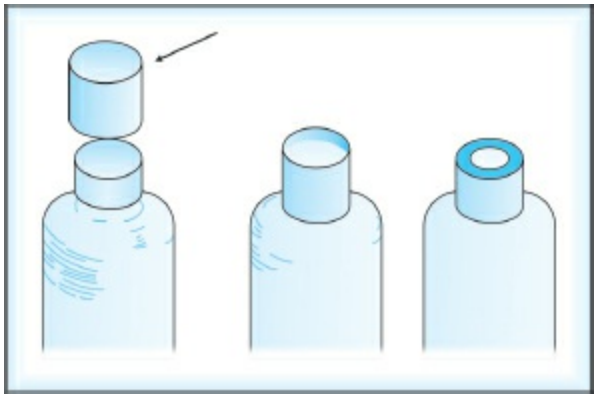


**Fig. 27.8:** Shrink banding (tubing)

## Foil, Paper, or Plastic Pouches

The flexible pouch is a packaging concept capable of providing not only a tamper-resistant package but also, by proper selection of material, a package with a high degree of environmental protection. A flexible pouch is usually formed during the product filling operation by either vertical or horizontal forming, filling, and sealing (f/f/s) equipment.

The vertical f/f/s operation, a web of film is drawn over a metal collar and around a vertical filling tube, through which the product is dropped into the formed package (Fig. 27.9). The metal filling tube also acts as a mandrel, which controls the circumference of the pouch and against which the longitudinal seal is made. The formation of this seal, which can be either a fin seal or an overlap seal, converts the packaging film into a continuous tube of film. Reciprocating sealers, orthogonal to the longitudinal seal, crimp off the bottom of the tube, creating the bottom seal of the package. The product drops through the forming tube into the formed package. The reciprocation sealer moves up the film tube a distance equal to the length of the package and forms the top and final seal of the package. This top seal of the package becomes the bottom seal of the next package and the process repeats itself. Since vertical f/f/s machines are gravity-fed, they are primarily used for liquid, powder, and granular products.

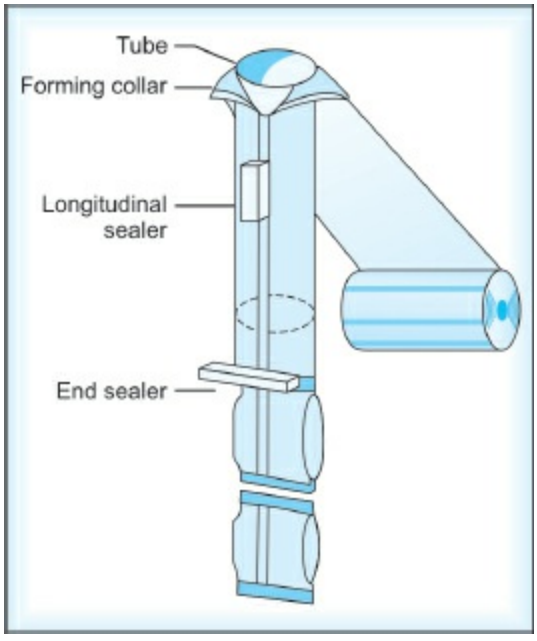


**Fig. 27.9:** Vertical form/fill/seal system

The horizontal f/f/s system is generally used for products of smaller volume, which are more amenable to the flatter format of the packages produced by this type of equipment. In this system, the web of film is folded upon itself rather than around a tube. As the folded film is fed horizontally through the equipment, a reciprocating platen creates pockets in the film by making vertical separation seals. The product is then placed into each pocket and the final top seal is made (Fig. 27.10). Packages formed on horizontal f/f/s equipment typically have a three-sided perimeter seal, but other variations are possible, depending on the type equipment used. The fin-seal wrapper illustrated in Fig. 27.5B can be considered a type of horizontal f/f/s packaging system.

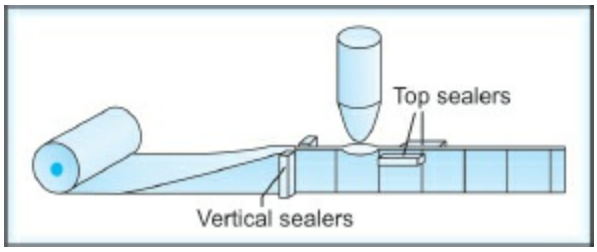



**Fig. 27.10:** Horizontal form/fill/seal system

To provide the degree of package integrity required for tamper-resistant packaging on both horizontal and vertical equipment, inner-surface-to-inner-surface sealing should be used. This permits the use of such effective sealants as polyethylene, ethylene vinyl acetate (EVA), and Surlyn, which when properly sealed must be torn apart to access the product. These sealants must be used as part of a lamination constructed to provide the necessary characteristics for the proper performance of the packaging material. The outside surface of the lamination should provide a good printing surface and should be thermally stable since it comes into direct contact with the heater bars. The outside surface material is also used as the substrate carrier, which gives the lamination the mechanical characteristics necessary for the machining and handling of the package. The most commonly used film for a substrate carrier is paper. Polyester, nylon, and cellophane are also used if transparency, puncture resistance, or gloss is desired. For moisture- and oxygen-sensitive products, foil is commonly used as part of the film lamination, with the foil sandwiched between the outer ply and the heat seal

layer. Such laminations as paper/polyethylene/foil/polyethylene and polyester/polyethylene/foil/polyethylene are commonly used for high-barrier applications. Metallized polyester is replacing foil for some high-barrier packaging applications because of its lower cost, excellent appearance, and flexural endurance.

## Bottle Seals

A bottle may be made tamper-resistant by bonding an inner seal to the rim of the bottle in such a way that access to the product can only be attained by irreparably destroying the seal. Various inner seal compositions may be used, but the structures most frequently encountered are glassine and foil laminations.

Typically, glassine liners are two-ply laminations using two sheets of glassine paper bonded together with wax or adhesive. The inner seals are usually supplied inserted in the bottle cap and held in place over the permanent cap liner by either a friction fit into the cap or a slight application of wax, which temporarily adheres the seal to the permanent cap liner. If glue-mounted inner seals are to be used, glue is applied to the rim of the bottle prior to the capping operation. The application of the cap forces the inner seal into contact with the glued bottle rim and maintains pressure during glue curing and until the cap is removed. When the bottle cap is removed, the inner seal is left securely anchored to the bottle rim.

Pressure-sensitive inner seals can also be used. The pressure-sensitive adhesive is coated on the surface of the inner seal as an encapsulated adhesive. During the capping operation, the torque pressure ruptures the encapsulated adhesive, which then bonds the inner seal to the rim of the bottle. One type of pressure-sensitive inner seal is constructed of thin-gauge styrene foam inner seal material coated on one side with a specially formulated torque-activated adhesive. The adhesive has minimal surface tack, but when applied with a properly torqued cap, it provides excellent adhesion to both glass and plastic bottles. A third method of application uses a heat-sensitive adhesive that is activated by high-frequency induction. This type of application requires the use of aluminum foil as part of the inner seal composition. Once the cap is applied, the bottle is passed under an induction coil, which induces high-frequency resonation in the foil. The frictional heat that is generated activates the heat-seal coating and bonds the liner to the bottle. This type of seal can only be used with plastic caps since metal caps would interfere with the induction sealing of the inner seal. To meet the tamper-resistant criteria, the inner seals must be printed or decorated with a unique design. The seal must also be bonded sufficiently to ensure that its removal would result in destruction of the seal.

## Tape Seals

Tape sealing involves the application of a glued or pressure-sensitive tape or label around or over the closure of the package, which must be destroyed to gain access to the packaged product. The paper used most often is a high-density lightweight paper with poor tear strength. Labels made of self-destructing paper are available; these cannot survive any attempt at removal once they have been applied. To reduce further the possibility of removing the label intact, perforation or partial slitting of the paper can be made prior to application so that the label tears readily along those weak points if any attempt is made to remove it.

## Breakable Caps

Breakable closures come in different designs. The roll-on cap design used in the past for carbonated beverages uses an aluminum shell, which is placed over the bottle neck during the capping operation. The cap blank is held on the bottle under pressure while rollers crimp and contour the bottle tread into the cap blank. The bottom portion of the cap is rolled around and under the locking ring on the bottle neck finish. This lower portion of the cap blank is usually perforated so that it breaks away when the cap is unscrewed, which serves as a visible sign of prior opening. A ratchet-style plastic cap is also commonly used for a number of different products. In this design, the bottom portion of the closure has a tear-away strip, which engages a ratchet on the bottle neck. To remove the closure, the bottom portion of the closure must be torn away to disengage the ratchet and allow the removal of the cap.

## Sealed Tubes

Collapsible tubes used for packaging are constructed of metal, plastic, or a lamination of foil, paper, and plastic. Metal tubes are used for those products that require high degree of barrier protection afforded by metal. Most of these are made of aluminum and are usually coated to eliminate compatibility problems between product and package. Puncture inserts, which are usually made of aluminum 3 to 5 mil thick, are used to seal the tube opening for tamper resistance. These inserts have to be punctured and pried out to gain access to the product. Extruded plastic tubes are widely used for those products that are compatible with the limited barrier characteristics of plastic. These tubes are usually constructed of polypropylene or polyethylene. For high-barrier packaging, metal or laminated tubes are used. Laminated tubes are constructed of a multilayer lamination made of foil, paper, and plastic specifically tailored to the product requirements. The lamination is used for the body of the tube with the head injection molded onto the tube. Since the head is injection-molded, any number of designs are available that must be cut or broken to gain access to the product. These seal end designs are usually molded of low density polyethylene. The tubes are filled from the other end and are sealed either by crimping the end in the case of metal tubes, or by induction sealing in the case of plastic or laminated tubes. Additional information can be found under the heading “Collapsible Tubes”, a previous section in this chapter.

## Aerosol Containers

The aerosol container used for pharmaceutical products is usually made of drawn aluminum. The inside of the container can be specially coated if product compatibility is a problem. A hydrocarbon propellant in its cooled liquid phase is added to the container along with the product, and a spray nozzle contained in a gasketed metal ferrule is crimped over the opening of the aerosol container. A length of polyethylene tubing, called a dip tube, is attached to the inside of the spray nozzle and dips into the product, drawing product into the spray nozzle when the sprayer is activated. The spray nozzles are usually metered to allow a specific dose to be dispensed with each spray. The design of the aerosol package makes it inherently tamper-resistant. *See* Chapter 21 for a further discussion on aerosol containers.

## Sealed Cartons

Folding paperboard cartons have been used as a secondary package for OTC products for many years. The popularity of this packaging mode is based on both functional and marketing considerations. With the advent of mass marketing of OTC products in the self-service sections of larger stores, shelf presence and product stack-ability became a dominant consideration in the package design. The mass distribution of fragile products also placed a requirement on the secondary package for the prevention of breakage during distribution. Labeling requirements in many cases exceeded the limited copy area provided by the label on the primary container and consequently required additional copy area to be provided as either inserts or carton panels. All of these considerations were addressed with the use of a folding carton to contain the primary package.

The closure of folding cartons can be accomplished in a number of ways. The most prevalent method has been the use of the “tuck end” design. The tuck end design feature allowed the ends of the carton to be held closed by the physical engagement of the side tabs at the open end of the carton, with the slits placed in the carton tuck or lid. This design feature, which has been prevalent in the folding carton industry because of its functionality and compatibility with highspeed packaging equipment, is no longer considered an acceptable closure mechanism for OTC products. If tuck end cartons are to be used, they must be augmented with some other form of tamper-resistant packaging such as film over-wrapping, tape sealing or glue sealing the carton. Seal end cartons differ from tuck end cartons in that rather than using the mechanical interlocking design of the tuck end to close the carton, externally applied glue or hot melt is used to provide carton sealing.

## PRODUCT-PACK VALIDATION

Compatibility studies between the product and pack are necessary to ensure that product degradation does not occur during the shelf life of the product. Also, the packaging material is such that it is capable of protecting the product from the environment. Stability studies will identify such requirements. The importance of the validation work required on the product and pack characteristics is shown using nonsterile and sterile product, respectively.

## Nonsterile Product-Pack Validation

For a tablet in a polypropylene screw-cap bottle, the following validation work will be required.

### *Water Vapor Permeability*

Permeation through the bottle wall will depend on whether the product has low or high affinity for water. For a moisture sensitive tablet permeation through the bottle wall and cap should be performed with packs containing the full quantity of tablets.

For determination of permeation through the bottle wall, each bottle is filled with tablets and the mouth of each bottle is immediately sealed with polyethylene coated aluminum foil. The samples are stored under controlled temperature and humidity conditions, that is, the highest expected market storage or that recommended by ICH. Sample of the product sealed into the bottle is analyzed immediately after sealing (control) and periodically analyzed after 3 or 6 months for moisture content and assay of active ingredient. To show whether just the capped bottle (without seal) is sufficient to protect the product from moisture uptake, the same experiment is repeated with cap in place but without aluminum seal over the bottle mouth. The product should be analyzed more frequently, possibly on a daily basis, since it is likely that the moisture uptake through the bottle cap threads will be high.

### *Light Transmission*

This test is to determine the effect of light passing through the bottle wall on the product appearance and stability. Although exposure of product to direct sunlight is not expected, it is likely that storage by customer or by pharmacist may result in limed storage of the product in daylight. Therefore, samples stored in both daylight and artificial light should be checked for discoloration or product degradation.

### *Stability*

It is unlikely that a compatibility problem, particularly with tablet, will occur, although it is still necessary to check up the full life of the product. There is the possibility that either the taste or smell of the tablets will be affected. The

complete packs along with product should be stored at recommended temperatures and analyzed periodically throughout storage. The analysis should include close examination at the tablet-container contact points, and of taste and smell.

## Sterile Product-Pack Validation

For a parenteral product in a glass vial with an inert atmosphere, the following validation work will be required.

### *Compatibility*

The pack components must be washed and sterilized through the validated procedures and then the vials must be filled with sterile product under aseptic conditions. Terminal sterilization is done if it is the part of the intended process operation. The product and pack compatibility test involves storage at each of market temperatures for one year beyond the expected market life of the product. Samples containing the amount of product to be marketed in each pack should be stored in an upright and inverted position. This is to determine both glass vial—product and rubber plug—product compatibility. Stability trials under dark should be performed since it is expected to be used on the market. Some samples must also be stored under daylight conditions to determine the maximum exposure time to daylight.

Component performance should be monitored during the compatibility trials to ensure that deterioration has not occurred.

### *Seal Integrity/Immersion Test*

The immersion test involves terminally sterilizing broth—or product-filled packages, followed by immersion of the seal area or the entire package in an aqueous suspension of microorganisms, as shown in Fig. 27.11, for an extended period of time. Alternatively, packages may be used that have been aseptically filled with culture media. In designing an immersion test, it should be remembered that the sterilization cycle should duplicate the one actually used for the product. The seals of each vial should be examined prior to microbial challenge to ensure that there are no defects. Often, the microbial identity of any positive samples is confirmed to screen out any false positives. After examination, each vial should be inverted into a tray containing the challenge microbe. Examples of those reportedly used for such tests include *P. aeruginosa, P. diminuta, E. coli, and S. marcescens*. Size, motility, and viability in the product or culture media are all factors in microorganism selection. Pressure cycles may be included in the challenge test to further stress the package or to mimic anticipated product processing,

distribution, and storage conditions. Incubation temperatures appropriate for growth of the challenge microorganism should be selected. Microbial growth in the package, as evidenced by cloudiness, may be detected visually or with instrumentation. In the case of product-filled packages, verification of nonsterility may require aseptic filtration and filter plating for microorganism identification. Any nonsterile package contaminants are generally identified to verify the challenge microorganism as the source of contamination.

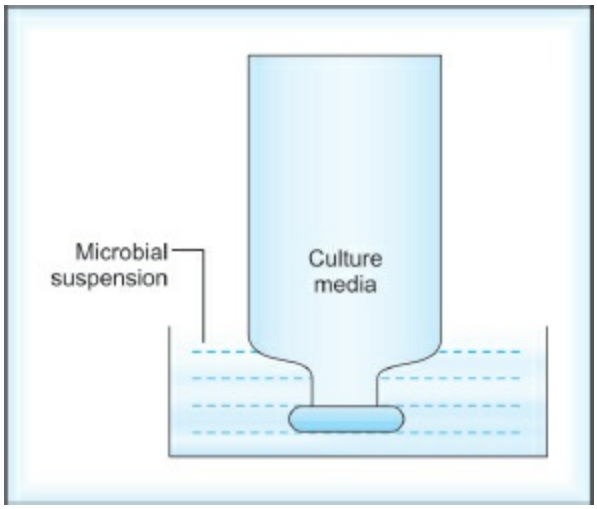



**Fig. 27.11:** Immersion test for seal integrity

## REGULATORY REQUIREMENTS

When the FDA evaluates a drug, the agency must be firmly convinced that the package for a specific drug will preserve the drug's efficacy as well as its purity, identity, strength, and quality for its entire shelf life. The specific FDA regulation relative to drugs states "containers, closures and other component parts of drug packages, to be suitable for their intended use, must not be reactive, additive or absorptive to an extent that the identity, strength, quality or purity of the drug will be affected". Under the provisions of the Federal Food, Drug and Cosmetic Act, however, no specifications or standards for containers or container closures are provided. Under the Act, it is the responsibility of the manufacturer to prove the safety of a packaging material and to get approval before using it for any food or drug product. The FDA does not approve containers as such, but only the materials used in the container. A list of substances considered "Generally Recognized As Safe" (GRAS) has been published by the FDA. The FDA has published regulations (part 133) that implement the Current Good Manufacturing Practice requirements of section 501 (a) of the Act. Part 133.9 of these regulations sets forth criteria with respect to product containers, which manufacturers, processors, packers, or holders of drugs use as guidelines. Any drug container must be approved for such use, along with the drug, before going on the market. The drug manufacturer must include data on the container and package components in contact with the pharmaceutical product. The main requirements are as follows:

**1. New drug application (NDA):** NDA must be submitted and approved prior to the sale of the product on the market. The details that need to be included are as follows:

a. Primary component formulation details, critical dimensions and drawings
b. Primary and secondary component artwork such as labels, leaflets and cartons.
c. Primary component quality standards, testing methods and compatibility data.
d. Product—pack validation data.
e. In-process operations and facility details of filling and packaging operations.

f. List of primary component manufacturers.

**2. Supplemental NDA type I:** This is submitted when major process changes are required. The supplemental NDA require FDA approval prior to implementation. The type of changes include:

a. Relaxing or deletion of primary component specification limit or critical dimensions that will not affect component quality.

b. Changing a testing method of a primary component.

c. Any change in artwork test.

d. The use of different facility for filling or packaging where the facility has not received a cGMP inspection.

**3. Supplemental NDA type II:** These changes can be implemented prior to FDA approval. Adding a new testing methodology to ensure better quality control, changes in artwork to strengthen safe use of a product, deletion of artwork giving false indications and small changes (cGMP approved) in filling or packaging process are included in supplemental NDA Type II changes.

**4. Annual NDA reports:** Annual NDA reports detail the minor changes (that do not require a supplemental application) that take place within a year after approval of NDA.

**5. Drug master file (DMF):** A DMF is a submission to FDA that provide detailed information about processes and facilities used in manufacturing, packaging and storing of pharmaceutical products. The complete file in relation to packaging component consists of two types (Type I and III), out of five types of DMF. DMF type I details manufacturing site and facility whereas, DMF type III details data supporting the suitability of the component for their intended use.

In the case of packaging materials, most packaging material manufacturers maintain DMF with the FDA. Upon request from the manufacturer, the FDA uses this file as a reference to support a New Drug Application that which a drug manufacturer files.

# 28:
# Kinetic Principles and Stability Testing

Degradation studies are an important aspect of preformulation evaluation of drugs. Preformulation scientists induce degradation of test materials by raising temperature, increasing humidity, by subjecting materials to sheer or compressive forces, exposing the test materials to various pH conditions, UV-visible light or by adding other reactants.

Degradative reactions in pharmaceutical formulations take place at definite rates and are chemical in nature. They depend on conditions like concentration of reactants, temperature, pH, radiation, and catalysts. An effective and efficient study of these reactions requires the application of chemical kinetic principles.

Degradation kinetics provides predictive information. It aims to predict the intrinsic stability of a drug in order to anticipate problems that may arise during development. Cost saving can be realized if these activities are coordinated.

Under certain circumstances, such as hydrolysis in solution, prediction of shelf life may be achieved with reasonable accuracy. Early signs of drug instability help to identify potential developmental issues, and thereby fosters development of potential stabilization strategies. It also suggests ways to optimize manufacturing processes.

This chapter will discuss the basic treatments of drug degradation studies, including kinetics, pathways, important factors, and typical practices used for assessing stability of pharmaceutical drug products.

# RATE, ORDER AND MOLECULARITY OF REACTIONS

In a chemical reaction, reactants yield products. When a reaction starts, the concentrations of reactants and products change with time until the reaction reaches completion or equilibrium. The concentrations of the reactants decrease, while those of the products increase over time. The velocity with which a reactant or a product undergoes chemical change is called the rate of a chemical reaction. Therefore, the rate of a reaction can be represented either by decreasing change in the concentration of reactant or increasing change in the concentration of product with respect to time.

*Consider a chemical reaction*

$$a\,\mathrm{A} + b\,\mathrm{B} \rightarrow c\,\mathrm{C} + d\,\mathrm{D} \quad \ldots (1)$$

Here *a*, *b*, *c* and *d* are the stoichiometric coefficients indicating the molar ratio of the reactants and products of the reaction. The rate of change of concentration of each species can differ, depending on the stoichiometric coefficients.

$$\text{Rate of reaction} = -\frac{1}{a}\frac{d[\mathrm{A}]}{\mathrm{dt}} = \frac{1}{b}\frac{d[\mathrm{B}]}{dt} = \frac{1}{c}\frac{d[\mathrm{C}]}{\mathrm{dt}} = \frac{1}{d}\frac{d[\mathrm{D}]}{dt} \quad (2)$$

Or

$$\text{Rate or reaction} = k\,[A]^{\alpha}\,[B]^{\beta} \quad \ldots (3)$$

Here *k*, the proportionality constant, is called the specific *rate constant*, or the *rate constant*, *α* and *β* are the *order of reaction* with respect to A and B. The order of the overall reaction, $n = \alpha + \beta$. The *order of reaction* is defined as the manner in which the rate of a reaction varies with the concentration of the reactants. For the most part, the degradation of pharmaceuticals can be treated as zero-order, first-order, or pseudo-first-order reactions, even though many of the pharmaceutical compounds degrade by complicated mechanisms. Consequently, the lower-order reaction are treated in detail here, and only minor consideration is given to higher-order reactions.

Regardless of the order of the reactions, the rate of reaction has the units of concentration/time (i.e. mole/L sec). The units of the rate constant are dependent on the overall order of the reaction:

$$k = (\text{concentration})^{1-n}\,(\text{time})^{-1} \quad \ldots (4)$$

When the rate equation corresponds stoichio-metrically, the reaction is called aan elementary reaction, and when it does not the reaction is called a non-elementary reaction (e.g. the thermal decomposition of nitrous oxide to nitrogen and oxygen; $N_2O \rightarrow N_2 + ½\ O_2$). When a reaction takes place via a single stoichiometric equation that has a single rate expression, it is called single reaction and when more than one stoichiometric equations are used to express the rate of reaction of all constituents then it is called multiple reaction.

The *molecularity* of a single elementary reaction is the number of molecules engaged in the reaction. A simple elementary reaction is referred to as uni-, bi-, or termolecular if one, two, or three chemical species are involved in the chemical reaction, respectively:

*Unimolecular:* $A \rightarrow B + C$

(e.g. bromine decomposition; $Br_2 \rightarrow 2\ Br$).

*Bimolecular:* $A + B \rightarrow C + D$

(e.g. formation of Hydroiodic acid; $H_2 + I_2 \rightarrow 2\ HI$).

*Termolecular:* $A + 2B \rightarrow C + D$

(rarely observed in pharmaceutical science).

Order and molecularity of a reaction are generally identical for elementary reactions. However complex reactions, involving multiple steps may have different order and molecularity of reaction.

## Zero-order Reaction

When the reaction rate is independent of the concentration of the reacting substance, then rate depends on the zero power of the reactant (Rate = k°C) and therefore, the reaction is considered to be zero-order reaction. In this type of reaction, the limiting factor is something other than concentration, for example, solubility or absorption of light in certain photochemical reactions. When solubility is the factor, only the amount of drug, which is in the solution undergoes degradation. This can be depicted as follows:

$$A \text{ (solid)} \rightleftarrows A \text{ (sol)} \rightarrow B \ldots (5)$$

As drug is consumed in the degradative reaction, more drug goes into solution until all solid A has reacted. Until this occurs, the degradative reaction does not depend on the total concentration of drug, but only on the portion that is in solution, resulting in a zero-order reaction.

The rate of decomposition of the drug with $C_0$ as initial concentration can be described mathematically as follows:

$$\text{Rate of concentration decrease} = \frac{-dC}{dt} = k \quad \ldots (6)$$

where; $C$ = concentration remaining at time t
$k$ = proportionality factor
= reaction rate
$t$ = time

$$\int_{C_0}^{C} dc = -k\int_{0}^{t} dt \quad \ldots (7)$$

Integration of equation (7) yields:

$$C - C_0 = -kt \ldots (8)$$

If the data from a stability study followeda zero-order reaction, a plot of *C* versus *t*, as shown in Fig. 28.1, results in straight line with the slope equal to −k. The value of $-k$ would indicate the amount of drug that is degrading per unit time, and the intercept of the line at time zero is equal to the constant in Eq. (48). An example of zero-order kinetics is shown by the loss of color of a liquid multisulfa preparation at elevated temperatures at 500 mμ.

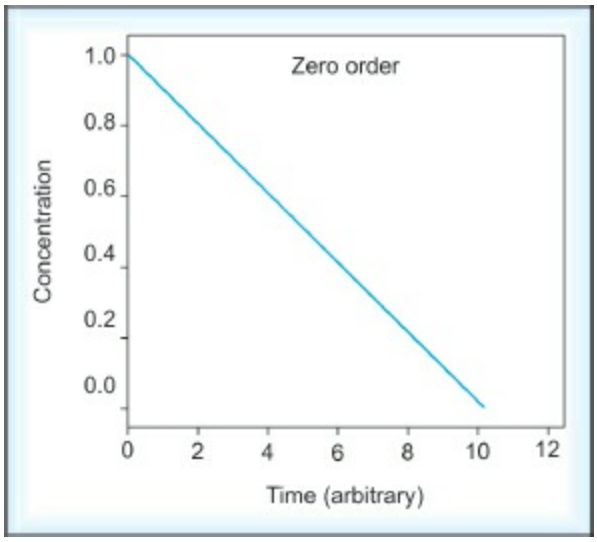



**Fig. 28.1:** A representative zero-order plot of the amount of drug reacting vs. time

## First-order Reaction

When the reaction rate depends on the first power of concentration of a single reactant (rate = $kC$), the reaction is considered to be first-order. In this type of reaction, a substance decomposes directly into one or more products (A → products). The rate of reaction is directly proportional to the concentration of the reacting substance and can be expressed mathematically in the following form:

$$\text{Rate of concentration decrease} = \frac{-dC}{dt} = kC \quad \ldots (9)$$

Integrating equation (9) in the following form:

$$\int_{C_0}^{C} \frac{dC}{C} = -k \int_0^t dt \quad \ldots (10)$$

We obtain:

$$\ln C - \ln C_0 = -k\,(t - 0) \ldots (11)$$

Converting to common logarithm yields:

$$\log C = \log C_0 - \frac{kt}{2.303} \quad \ldots (12)$$

Using the above equation for a first-order reaction, a straight line is produced when the logarithm of the concentration $C_a$ is plotted against time, as shown in Fig. 28.2. The velocity or reaction rate constant, ($k$), can be calculated by multiplying the slope of the line by 2.303. Higher the temperature, greater is the k value, as evidenced by the steepness of the slopes.

Integration of Eq. (9) between the limits $C_1$ and $C_2$ and $t_1$ and $t_2$ results in the following:

$$-\int_{C_1}^{C_2} \frac{dC}{C} = k \int_{t_1}^{t_2} dt \quad \ldots (13)$$

$$-(\ln C_2 - \ln C_1) = k(t_2 - t_1) \quad \ldots (14)$$

$$k = \frac{1}{t_2 - t_1} \ln \frac{C_1}{C_2} = \frac{2.303}{t_2 - t_1} \log \frac{C_1}{C_2} \quad \ldots (15)$$

These equations permit the calculation of the rate of decomposition of a substance between any time interval ($t_2$–$t_1$) if the concentration of drug at

these two times is known.

When $t_1$ is the time at the beginning of reaction, $t_0$ is the time at concentration $C_0$, and $t_1$ is any time t at concentration $C$, then Eq. (15) can be expressed as follows:

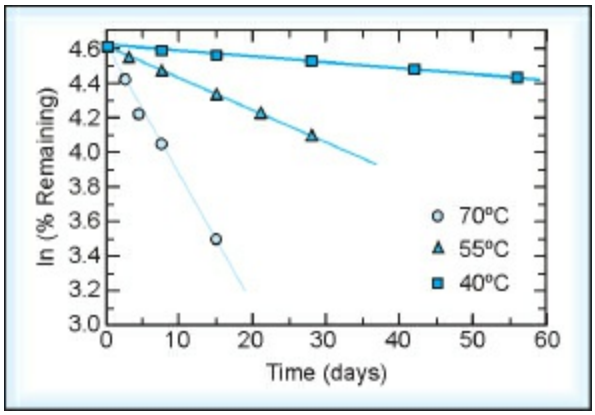



**Fig. 28.2:** Representative degradation curves for a material deteriorating according to first-order kinetics

$$k = \frac{2.303}{t} log \frac{C_0}{C} \qquad \text{... (16)}$$

Use of this expression permits the calculation of the rate of reaction $k$, by determining the concentration of drug remaining at any time t. Eqs (15) or (16) can also be used to ascertain whether the reaction is following first-order kinetics by determining k at several time intervals and noticing whether the values are essentially constant. Equation (16) is also written as follows:

$$k = \frac{2.303}{t} log \frac{a}{(a-x)} \qquad \text{... (17)}$$

Where $a = C_0$, $x$ = amount reacting in time $t$, and $(a–x)$ the amount remaining after time $t$.

The constant $k$ is called the reaction velocity constant, or more frequently, the specific reaction rate. For a first-order reaction, it is a number that expresses the fraction of the material reacting in a unit of time and may be expressed in reciprocal seconds, minutes, or hours. For example, when k has a value of 0.001 $sec^{-1}$, the material is decomposing at a rate of 0.1% per second.

The time necessary for a fraction of the material to degrade can be readily

calculated. The *half-life*, $t_{1/2}$, of a drug is the time required for 50% of the drug to degrade and can be calculated as follows:

$$t_{½} = \frac{2.303}{k} log \frac{C_0}{C} = \frac{2.303}{k} log \frac{100}{50} \quad \text{... (18)}$$

$$\text{Therefore, } t_{½} = \frac{0.693}{k} \quad \text{... (19)}$$

In the pharmaceutical field, *shelf life*, the time required for 90% of drug to remain (or 10% of the drug to degrade) is important, since it represents a reasonable limit of degradation of active ingredients. The $t_{90\%}$ value can be calculated as follows:

$$t_{90\%} = \frac{2.303}{k} \log \frac{100}{90} = \frac{0.105}{k} \quad \text{... (20)}$$

It is important to note here that the $t_{1/2}$ or $t_{90\%}$ is concentration independent. In other words, it takes the same time to reduce the concentration of drug from 0.1 moles to 0.05 moles as it would to go from 0.001 moles to 0.0005 moles.

$$t_{90\%} = 0.152\ t_{½} \quad \text{... (21)}$$

## Second-order Reaction

If a reaction rate depends on the concentration of two reactant species (rate = $k\ C_a\ C_b$ or $k = C_a\ C_a$ or $k\ C_a^2$), the reaction would be second order. Second order reactions are of two types:

Type 1: $\rightarrow$ A + A P (rate = $k\ C_a\ C_a$ or = $k\ C_a^2$)

Type 2: $\rightarrow$ A + B P (rate = $k\ C_a\ C_b$).

When the initial concentrations of A and B are identical, the rate equation can be simplified as:

$$\frac{d[C]}{dt} = -k[C]^2 \qquad \text{... (22)}$$

Where $k_2$ is second-order rate constant. Integration of Eq. (22) yields:

$$\int_{[C]_0}^{[C]} \frac{d[C]}{[C]^2} = -k\int_0^t dt \ \text{ or } \ \frac{1}{[C]} - \frac{1}{[C_0]} = -kt \qquad \text{... (23)}$$

For such reactions the plot of 1/[*C*] vs. t gives a straight line with slope of *k* as shown in Fig. 28.3. When the initial concentrations of *A* and *B* are identical, the concentrationtime profile can be simplified as:

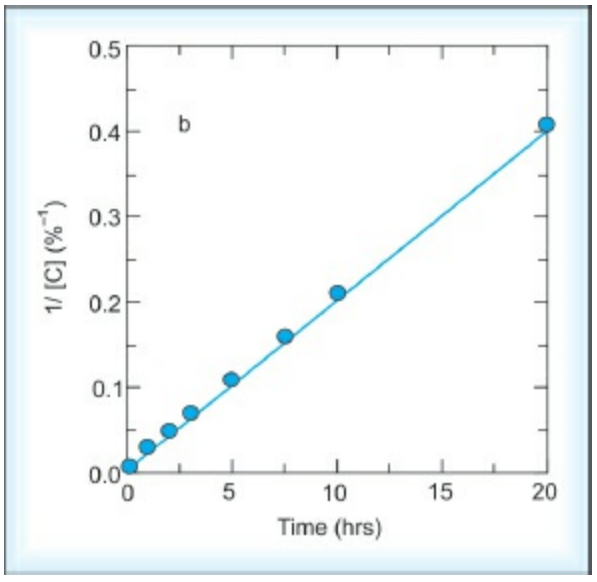



**Fig. 28.3:** Linear plot of kinetic data for second-order reaction

Figure 28.4 plots the reactant concentrationtime profiles for theoretical

zero-, first-, and second-order kinetics. Table 28.1 summarizes the rate equations, formulae for calculating reactant concentration—time profiles, halflives and shelf life for previously discussed simple order kinetics. The rate constants used to generate Fig. 28.4 were assumed to be numerically identical in all cases. Identical initial reactant concentrations were assumed for the second-order reaction in both Fig. 28.4 and Table 28.1.

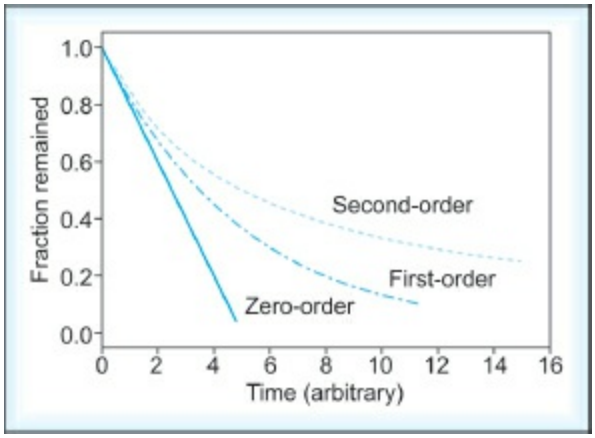



**Fig. 28.4:** Reactant concentration-time profiles for theoretical zero, first, and second-order reactions

**Table 28.1:** Summary of the rate equations, concentration—time profiles the simple order kinetics

| | | | Second-order | |
|---|---|---|---|---|
| | **Zero-order** | **First-order** | **a = b = $C_0$** | **a ≠ b** |
| Differential rate expression | $-\frac{dc}{dt} = k$ | $-\frac{dc}{dt} = kc$ | $-\frac{dc}{dt} = kc^2$ | $-\frac{dc}{dt} = kc_a c_b$ |
| Concentration time profile | $c = c_0 - kt$ | $c = c_0\, exp^{(-kt)}$ | $-\frac{dc}{dt} = kc_a c_b$ | $k = \frac{1}{t(a-b)} \ln \frac{b(a-x)}{a(b-x)}$ |
| **Unit of rate constant** | **moles/liter second** | **1/second** | **liter/mole second** | |
| Half life ($t_{½}$) | $\frac{c_0}{2k}$ | $\frac{0.693}{k}$ | $\frac{1}{c_0 k}$ | (i) When $x = 0.5a$ $\frac{1}{k(a-b)} \ln \frac{0.5ab}{a(b-0.5a)}$ (ii) When $x = 0.1b$ $\frac{1}{k(a-b)} \ln \frac{b(a-0.5b)}{0.5ab}$ |
| Shelf-life ($t_{90\%}$) | $\frac{c_0}{10k}$ | $\frac{0.105}{k}$ | $\frac{0.11}{c_0 k}$ | (i) When $x = 0.1a$ $\frac{1}{k(a-b)} \ln \frac{0.9ab}{a(b-0.1a)}$ (ii) When $x = 0.1b$ $\frac{1}{k(a-b)} \ln \frac{b(a-0.1b)}{0.9ab}$ |

## Pseudo-zero-order Reaction

In the solid state, many drugs decompose according to pseudo-zero-order rates as reactions occur between the solid drug and moisture. The system behaves as a suspension, and because of the presence of excess solid drug, the first-order reaction rate actually becomes a pseudo-zero-order rate, and the drug loss rate is linear with time. The rate expression becomes similar to that in Eq. (7) except that $k$ is $K'$, which indicates the pseudozero-order reaction rate. Pseudo-zero-order rates of reactions frequently occur in drugs formulated as pharmaceutical suspensions.

## Pseudo-first-order Reaction

A pseudo-first-order reaction can be defined as a second-order or bimolecular reaction that is made to behave like a first-order reaction. This is found in the case in which one reacting material is present in great excess or is maintained at a constant concentration as compared with the other substance. Under such circumstances, the reaction rate is determined by one reactant even though two are present (since the second reactant does not exhibit a significant change in concentration during the degradative reaction). An example of such a situation is the hydrolysis of an ester catalyzed by hydroxyl ion. If the hydroxyl ion concentration is high as compared with the concentration of the ester, the reaction behaves as a first-order reaction and can easily be followed by assay for residual ester. A similar approach, and one more frequently employed, is to keep the pH constant through the use of appropriate buffers.

An example of a drug that obeys pseudofirst-order kinetics is cefotaxime sodium. As shown in Fig. 28.5, semilogarithmic plots of cefotaxime sodium concentration versus time results in linear relationships at various pH levels demonstrating first-order rate of hydration (log $k_1$ vs. $t$). Acid catalysis occurs at ≤pH 4, and base catalysis occurs at pH 8, in which the concentrations of $H^+$ and OH respectively are high compared with the concentration of cefotaxime sodium.

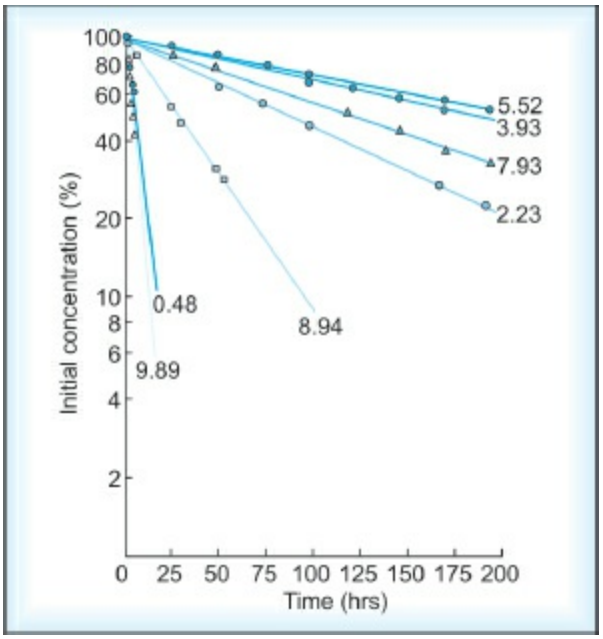


**Fig. 28.5:** Observed pseudo-first-order plots for the degradation of cefotaxime at various pH values, 25°, and μ = 0.5. Values indicated in figure are of pH

## COMPLEX REACTIONS

Although most degradative reactions occurring in pharmaceutical systems can be treated by zero-order, first-order, and pseudo-first-order kinetics, there are certain pharmaceutical formulations that exhibit more complicated reactions. These have opposing, consecutive, and side reactions along with the main reaction. In most instances, the extent of the simultaneous reactions is small in comparison with the main reaction and can be neglected. Some of these are briefly described.

## Opposing/Reverse Reactions

The simplest case of a reversible reaction is that in which both reactions are of the first order, as illustrated by the following:

$$A \underset{k'}{\overset{k}{\rightleftharpoons}} B \qquad \text{... (24)}$$

A somewhat more complicated reversible reaction is one in which the forward reaction is of a first-order type and the reverse reaction of a second-order type, as demonstrated by the following:

$$A \underset{k'}{\overset{k}{\rightleftharpoons}} B + C \qquad \text{... (25)}$$

When the forward and reverse reactions are both of the second-order type, the reaction takes on the following form:

$$A + B \underset{k'}{\overset{k}{\rightleftharpoons}} C + D \qquad \text{... (26)}$$

Reversible reactions of this type are quite common, but usually, the reverse reaction is ignored because the concentration is not significantly affected. An example of this is expressed by the following:

$CH_3COOH + C_2H_5OH \rightleftarrows CH_3COOC_2H_5 + H_2O$

Initially, the reverse reaction can be ignored, but as the reaction proceeds, and the concentration of water and ethyl acetate increases, both reactions influence Eq.(26).

Since this has been given as a first example of a second-order reaction, a brief discussion of this type of reaction is presented. For Eq. (26), the rate of reaction is proportional to the concentration of the two reacting substances A and B for the forward reaction and C and D for the reverse reaction. For the forward reaction, if a and b represent the initial concentration of the two reacting substances, and if $x$ denotes the moles of A and B in each liter reacting in the interval of time t, then the velocity of the reaction is expressed by the equation:

$$\frac{dx}{dt} = k\,(a - x)\,(b - x) \qquad \text{... (27)}$$

When A and B are present in equal concentrations, $a = b$:

$$\frac{dx}{dt} = k\,(a - x)^2 \qquad \text{... (28)}$$

Integrating equation (28) yields:

$$\frac{1}{k}\frac{dx}{(a-x)^2} = dt$$

$$\frac{1}{k}\frac{1}{(a-x)} = t + constant$$

$$\text{for } t = 0, \text{constant} = \frac{1}{ka} \text{ (since } x = 0 \text{ at } t = 0)$$

$$k = \frac{1}{t}\frac{x}{a(a-x)} \qquad \text{... (29)}$$

The half-life or time for 50% degradation ($t_{½}$) can be calculated by substitution.

Since $x$ = ½a at the half-life, substituting into Eq. (29) results in the following equation:

$$t_{½} = \frac{1}{ka} \qquad \text{... (30)}$$

Integrating Eq. (27), in which concentrations of A and B are not equal, the following equation results:

$$kt = \frac{1}{a-b}\ln\frac{b(a-x)}{a(b-x)}$$

or

$$k = \frac{2.303}{t(a-b)}\log\frac{b(a-x)}{a(b-x)} \qquad \text{... (31)}$$

In such reaction, plotting the log $\frac{b(a-x)}{a(b-x)}$ versus time ($t$), a straight line is obtained, and $k$ is thenmen obtained by multiplying the slope of the line by 2.303/($a$ – $b$).

## Consecutive Reactions

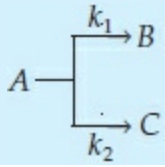
When the stages of a consecutive reaction occur at rates of approximately same magnitude, each stage must be considered in the kinetics of the overall reaction. The simplest case is one in which both the consecutive processes are of first order, as illustrated by the following equation:

$$A \xrightarrow{k_1} B \xrightarrow{k_2} C \qquad \ldots (32)$$

In the reaction, if $k_2$ is considerably greater than $k_1$, $B$ can be considered as an unstable intermediate, and the rate determining step for the overall reaction would be the conversion of $A$ to $B$. The overall reaction could then be treated by first-order kinetics.

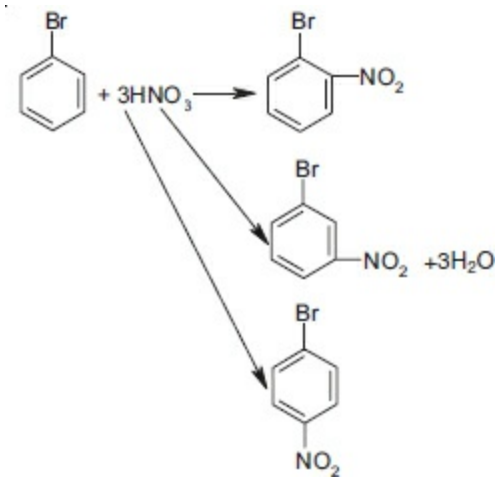
## Side Reactions

In some processes, the reacting substance can be removed by two or more reactions occurring simultaneously, as depicted by the following equation:

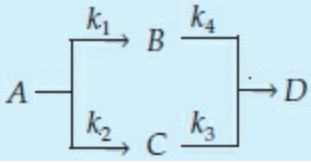



... (33)

In general, side or competing reactions are more common to organic chemistry. The organic chemist routinely deals with the production of several compounds from these two reactants; however, through the proper manipulation of conditions (e.g. pressure, temperature, concentration) the desired product predominates. An example of a competing reaction is the nitration of bromobenzene to produce ortho, meta, and para nitrobenzene as follows:




Purified insulin degrades by two mechanisms—deamidation and polymerization. Degradation reactions may occur as consecutive and side reactions according to the following scheme:




... (34)

where:, *A* is insulin, *B* is desamido insulin, *C* is polymerized insulin and *D* is polymerized desamido insulin.

The relative rates of deamidation and polymerization are pH and temperature-dependent. This example of a complex reaction is probably representative of the complexity of degradation mechanisms that are seen with polypeptides produced by genetic engineering and developed as pharmaceutical dosage forms.

# FACTORS INFLUENCING REACTION RATES

## Temperature

In order for the rate constants or velocity of degradation to be of use in the formulation of pharmaceutical products, it is necessary to evaluate the temperature dependency of the reaction. According to rule-of-thumb methods, the rate of reaction is said to double for each 10° rise in temperature. Although this rule may serve as a fairly accurate estimate for certain preparations, it is not generally applicable. Therefore, to assign an overall factor for the influence of temperature on the acceleration of reactions is foolhardy. Some deterioration reactions are not measurably influenced over a 10° temperature range, while others undergo rapid degradative changes. The recommended procedure is to set up a planned schedule of accelerated tests for each formulation in order to ascertain the temperature dependency of the chemical changes in the product undergoing evaluation.

The most satisfactory method for expressing the influence of temperature on reaction velocity is the quantitative relation proposed by arrhenius:

$$k = Ae^{-Ea/RT} \quad \text{... (35)}$$

where:

- $k$ = specific rate of degradation
- $A$ = frequency factor or Arrhenius factor
- $E_a$ = activation energy
- $R$ = gas constant (1.987 calories degree$^{-1}$ mole$^{-1}$)
- $T$ = absolute temperature

The constant of integration (A) or frequency factor in the ***Arrhenius equation*** is a measure of the frequency of collisions that can be expected between the reacting molecules for a given reaction. Logarithmically, it may be expressed as follows:

$$\log k = \log A - \frac{E_a}{2.303RT} \quad \text{... (36)}$$

where log A can be considered a constant.

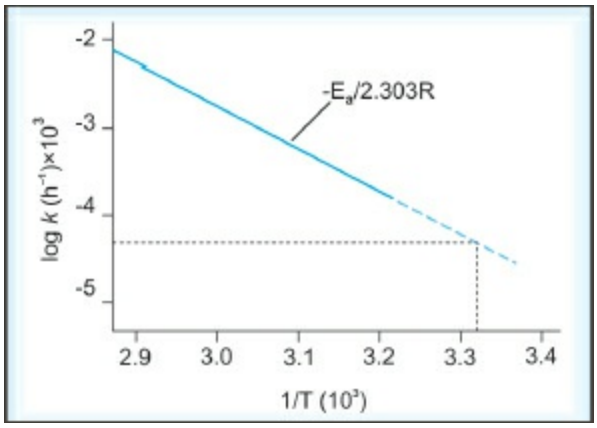



**Fig. 28.6:** Dependence of reaction rate on temperature

From Eq. (36), a plot of log k versus $\frac{1}{T}$ yields a slope equal to – $E_a$/2.303 $R$ from which the value for the activation energy can be calculated (Fig. 28.6).

The activation energy ($E_a$) represents the energy the reacting molecules must acquire to undergo reaction. Higher the value for the heat of activation, more the stability is temperature-dependent. $E_a$ can also be obtained by determining k at two different temperatures ($T_1$ and $T_2$).

Upon integration of the Eq. (35), between the limits $k_1$ and $k_2$ and $T_1$ and $T_2$, the following equation results:

$$\log\left[\frac{k_2}{k_1}\right] = \frac{E_a}{2.303R}\,\frac{(T_2 - T_1)}{T_2 T_1} \qquad \ldots (37)$$

The utility of the temperature dependency relationship depends on the controlling mechanisms of degradation. Preparations that degrade through solvolytic processes, e.g. reactions in solution, usually have heats of activation in the range of 10 to 30 kcal/mole. Here, considerable advantage may be taken of the significant increases in rate of reaction that result with temperature elevation. On the other hand, if diffusion or photolysis are the rate determining steps of the reaction, the heat of activation is only of the magnitude of 2 to 3 kcal/mole, and little advantage is gained by accelerated temperature studies in prediction, since the temperature effect on rate is small. For reactions such as pyrolysis of polyhydroxylic materials, in which the heat of activation can be of the magnitude of 50 to 70 kcal/mole, the rate of degradation, which may be great at elevated temperatures, may not be of

any practical significance at the temperatures of marketing and storage of the pharmaceutical preparation. The manner in which temperature affects reaction rates may be understood by considering two theories: Classic collision theory and transition-state theory.

## *Classic Collision Theory*

The reaction is possible, when reactant molecules collide with each other. However, not every collision leads to a successful reaction. The classic collision theory postulates that the reaction takes place only if the involved molecules exceed a certain minimum energy (the activation energy).

The fraction of molecules with kinetic energies exceeding $E_a$ can be derived using the Boltzmann distribution law:

$$f(E_k > E_a) = \exp\left(\frac{E_a}{RT}\right) \quad \text{... (38)}$$

Here *f* is the fraction reacting (or probability of reaction), and $E_k$ is the kinetic energy. Therefore, the rate can be expressed as:

$$rate = ZPC_{E_k > E_a} = ZPC_T \exp\left(-\frac{E_a}{RT}\right) \quad \text{... (39)}$$

Where *Z* is the collision frequency, *P* is the steric or probability factor ($0 < P < 1$), because not every collision leads to a reaction, $C_{Ek} > E_a$ is the concentration of the reactant molecules with kinetic energy exceeding the activation energy, and $C_T$ is the total reactant concentration. By comparison with the general rate law, one can immediately see that the expression of the rate constant, *k*, is similar to the Arrhenius equation:

$$k = ZP \exp\left(\frac{-E_a}{RT}\right) \quad \text{... (40)}$$

Thus *ZP* is identified with the Arrhenius frequency factor, while $E_a$ is the activation energy.

## *Transition-state Theory*

The transition state theory also referred to as the absolute rate theory postulates that in addition to collision, a transition state (or activated complex) must exist in which the reacting molecules have an appropriate

configurational geometry for the reaction to occur; furthermore, the transition state is in equilibrium with the reactants. In order for the activated complex to form, a certain energy barrier (the activation energy) needs to be overcome. Decomposition of the activated complex leads to products.

For an elementary bimolecular reaction, a simplified scheme can be represented as follows:

$$A + B \underset{}{\overset{K^{**}}{\rightleftharpoons}} [A...B]^{**} \rightarrow C$$

The double asterisk designates the transition state. The transition state theory assumes that the decomposition of the activated complex is the rate-determining step of the reaction. Among all the molecular motions of the activated complex, there is one mode of vibration (frequency designated as $\upsilon$) that will lead to the formation of the product. Combined with the equilibrium assumption, the rate of the reaction can be written as:

$$Rate = \frac{d[C]}{dt} = v[A \cdots B]^{***} = vK^{**}[A][B] \quad \text{... (41)}$$

Thus, the rate constant of the reaction, k, has the form:

$$k = \upsilon K^{**} \text{ ... (42)}$$

The transition state theory provides a better picture of chemical reactivity, and is still used today. It accounts for the influences on reaction rate by various factors, such as solvent, ionic strength, and dielectric constant. In general, enhancement of the reaction rate is expected if a factor stabilizes the transition state. However, sometimes changing one factor might also change the effects of other factors. Therefore, caution must be taken while performming such exercises. Figure 28.7 shows the schematic diagram of the transition state for an exothermic reaction.

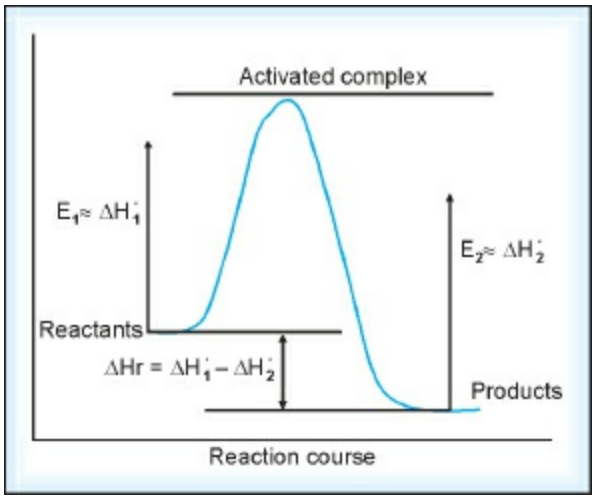



**Fig. 28.7:** Schematic diagram of the transition state for an exothermic reaction

## pH

The magnitude of the rate of hydrolytic reactions catalyzed by hydrogen and hydroxyl ions can vary considerably with pH. Hydrogen ion catalysis predominates at the lower pH range, whereas hydroxyl ion catalysis operates at the higher pH range. At the intermediate pH range, either the rate can be independent of pH or catalyzed by both hydrogen and hydroxyl ions. The rate constants in this pH range are usually less, however, than those at higher or lower pH values. To determine the influence of pH on the degradative reaction, the decomposition is measured at several hydrogen ion concentrations. The pH of optimum stability can be determined by the plotting the logarithm of the rate constant versus pH, as illustrated by the pH profile in Fig. 28.8. The point of inflection of such a plot represents the pH of optimum stability. Knowledge of this point is extremely useful in the development of a stable dosage form, provided the pH is within safe physiologic limits. Studies of this type can be performed at elevated temperatures so that data can be obtained in as short a time as possible. The shift of this point of inflection caused by temperature elevation is usually not of sufficient magnitude to affect seriously the conclusions drawn from such data. The plot in Fig. 28.9 gives an actual example in which the point of inflection for methyl-DL-a-phenyl-2-piperidylacetate served as a guide in the development of a stable injectable solution.

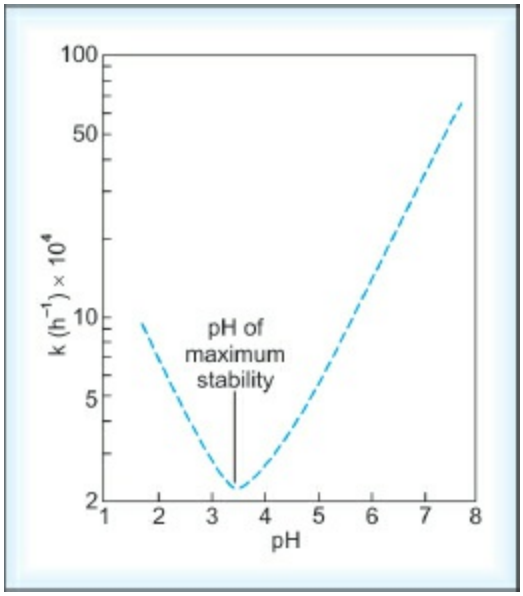



**Fig. 28.8:** pH inflection plot of maximum stability

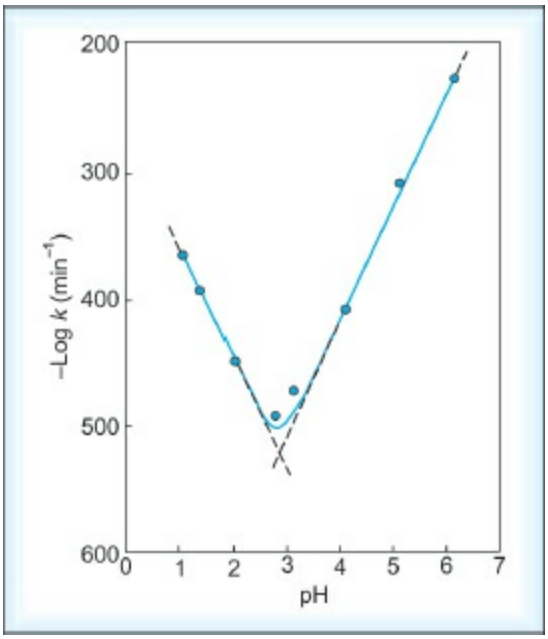



**Fig. 28.9:** pH dependency of the hydrolysis of methyl DL-a-phenyl-2-piperidylacetate at 80°

## General Acid-base Catalysis

Buffer salts are commonly used in the formulation of pharmaceutical liquids to regulate the pH of the solution. Although these salts tend to maintain the pH of the solution at a constant level, they can also catalyze the degradation. Therefore, it is necessary to evaluate the effect of buffer concentration on the stability of the preparation in addition to the effect of hydrogen and hydroxyl ion concentrations. Common buffer salts such as acetate, phosphate, and borate have been found to have catalytic effects on the degradation rate of drugs in solution. As examples, Fig. 28.10 illustrate the catalytic effects of phosphate buffers on cefadroxil degradation at various pH.

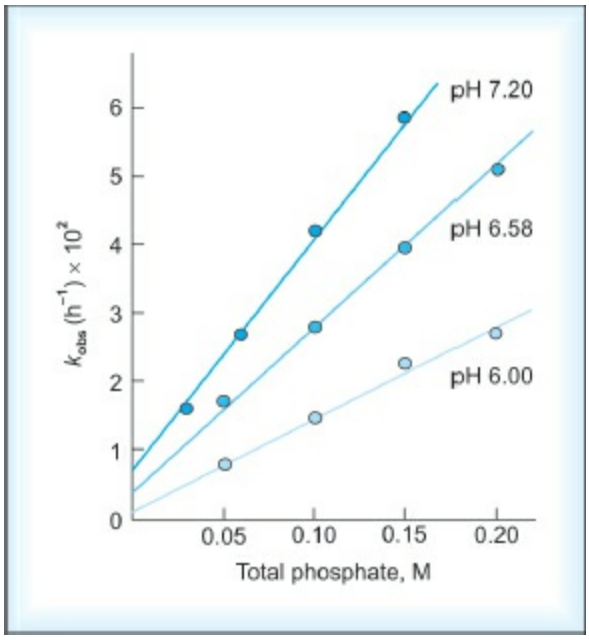



**Fig. 28.10:** Plots of pseudo-first-order rate constant vs. total phosphate buffer concentration for cefadroxil degradation at various pH values, 35°, and ionic strength 0.5

To determine whether a particular formulation is catalyzed by the employed buffer system, the ionic strength of solution is kept constant, and the concentration of buffer is altered, while the ratio of the buffer salts is kept constant to maintain the pH. If the degradation reaction is found to be influenced by the different concentrations of buffer, then the reaction is considered to be general acid and base catalyzed. In such a case, the concentration of the buffer ratio should be kept as low as possible to diminish

this catalytic effect.

## *Ionic Strength*

The rate of reaction can be influenced by the ionic strength of the solution in accordance with the following equation:

$$\log k = \log k_0 + 1.02 Z_A Z_B \sqrt{\mu} \quad \ldots (43)$$

where $Z_A + Z_B$ are the charges carried by the reacting species in solution, $\mu$, the ionic strength, $k$, the rate constant of degradation, and $k_0$, the rate constant at infinite dilution. The ionic strength ($u = \frac{1}{2}\Sigma C_i Z_i^2$) is defined as half the sum of the terms obtained by multiplying the concentration of each of the ionic species present in the solution by the square of its valence. Plotting the logarithm of the reaction rates versus the square root of the ionic strength, as illustrated in Fig. 28.11, can determine whether an increase in ionic strength increases, reduces, or has no effect on the degradation rate.

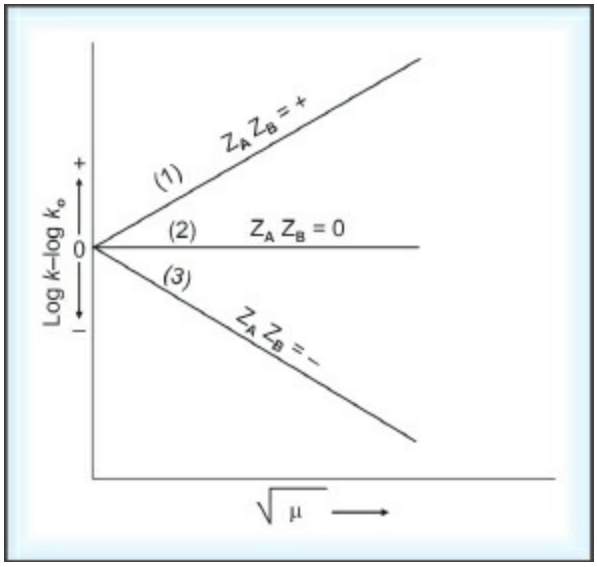



**Fig. 28.11:** Dependence of reaction rates on ionic strength

The concentration of salt employed in a liquid pharmaceutical formulation can increase or decrease the degradation rate of the drug in solution or have no effect. When the drug is positively charged and is undergoing hydrogen ion catalysis, an increase in ionic strength caused by the addition of a salt, such as sodium chloride, causes an increase in the rate of degradation, as shown in curve 1, Fig. 28.11. A decrease in the rate of

degradation results if the positively charged drug is undergoing hydroxyl ion catalysis, and the ionic strength is increased by addition of a salt as shown in curve 3, Fig. 28.11. If the drug undergoing degradation is a neutral molecule, changes in ionic strength by the addition of a salt would have no effect on the rate of gradation, as shown in curve 2, Fig. 28.11.

## Dielectric Constant

The dielectric constant (or relative permittivity) of a solvent is a measure of its polarity. Water has a high dielectric constant (approximately 78 at room temperature); other solvents have much lower values (e.g. for example, approximately 24 for ethanol). The equation that describes the effect of the dielectric constant, $\varepsilon$, on the rate constant of an ionic reaction is:

$$\log k = \log k_{\varepsilon=\infty} - KZ_AZ_B \times \frac{1}{\varepsilon} \quad \ldots (44)$$

where $k_\varepsilon = \infty$ is the rate constant in a solvent of infinite dielectric constant, $K$ is a constant for a particular reaction at a given temperature, $Z_A$ and $Z_B$ are the charge numbers of the two interacting ions. This equation predicts that a plot of log k against the reciprocal of the dielectric constant of the solvent should be linear with a gradient ($-KZ_AZ_B$). The intercept when $1/\varepsilon = 0$ (i.e. when $\varepsilon = \infty$) is equal to the logarithm of the rate constant, $k_\varepsilon = \infty$, in a theoretical solvent of infinite dielectric constant (Fig 28.12):

The positive gradient (curve 1) indicates that, the drug ion and the interacting ion are of opposite signs and therefore the choice of a non-polar solvent will only result in an increase of decomposition. The negative gradient (curve 3) indicates that, the charges on the drug ion and the interacting species are the same. This means that if we replace the water with a solvent of lower dielectric constant then we will achieve the desired effect of reducing the reaction rate.

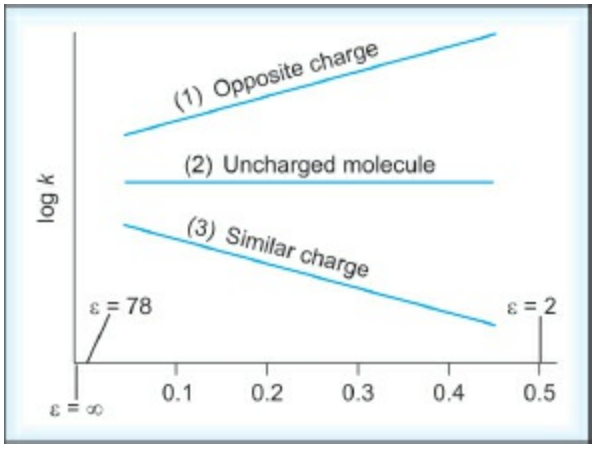



**Fig. 28.12:** Influence of dielectric constant on the rate constant for reactions

between ions of different charges

## DEGRADATIVE PATHWAYS

Although the decomposition of active ingredients in pharmaceutical dosage forms occurs through several pathways, i.e. hydrolysis, oxidation-reduction, racemization, decarboxylation, ring cleavage, and photolysis, those most frequently encountered are hydrolysis and oxidation-reduction. Consequently, this section treats these two important degradation processes in detail and only briefly reviews the others.

## Hydrolysis

Many pharamaceuticals contain ester or amide functional groups, which undergo hydrolysis in solution. Examples of drugs that tend to degrade by hydrolytic cleavage of an ester or amide linkage are anesthetics, antibiotics, vitamins, and barbiturates.

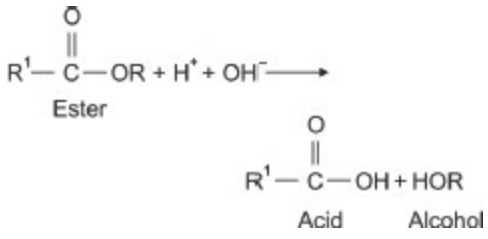
## *Ester Hydrolysis*

The hydrolysis of an ester into a mixture of an acid and alcohol essentially involves the rupture of a covalent linkage between a carbon atom and an oxygen atom. Although some hydrolyses can be effected in pure water, in the majority of cases, the presence of a catalyst is needed to promote the reaction. These catalysts are invariably substances of a polar nature, such as mineral acids, alkalies, or certain enzymes, all of which are capable of supplying hydrogen or hydroxyl ions to the reaction mixture. The alkaline hydrolysis of an ester does not differ essentially from an acid-catalyzed hydrolysis, except that it is irreversible, and therefore quantitative, because the resultant acid is at once neutralized. On the other hand, the acid-catalyzed hydrolysis of esters is reversible and may be made essentially complete in either direction by an excess of water or alcohol.

The numerous schemes presented to represent the hydrolysis of esters by either alkali or acid, the one given by Walters is perhaps the clearest to visualize.

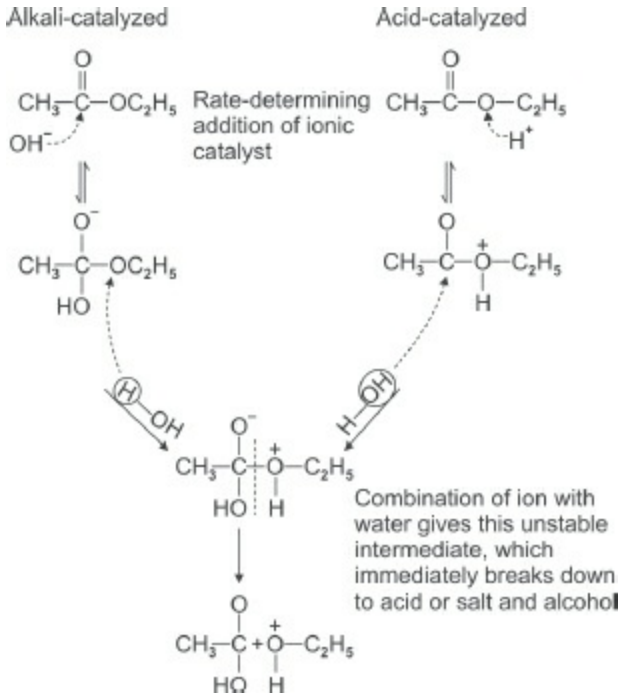
For both the alkali and acid-catalyzed hydrolysis, it is evident that the ester is cleaved at the acyl-oxygen linkage, that is, between the carbonyl carbon $\left(\begin{matrix}O\\ \|\\ C\end{matrix}\right)$ and the oxygen of $C_2H_5$ ($O–C_2H_5$). This type of cleavage takes place for most ester hydrolytic reactions.

In practice, the general scheme employed to denote ester hydrolysis is as follows:



This holds true for both acid- or alkaline-catalyzed reactions.

The general form of the kinetic equations to express acid or base-catalyzed hydrolysis is as follows:




$$\frac{d(ester)}{dt} = k(ester)(H^+)$$

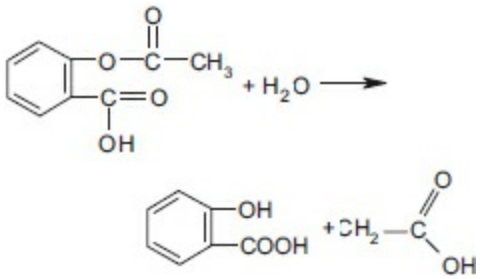
$$\frac{d(ester)}{dt} = k(ester)(OH^-)$$

These equations denote second-order reactions, but in studying degradation reactions of this type, it is possible to treat them as pseudo-first-order reactions. This is done by keeping the $OH^-$ or $H^+$ at a considerably higher concentration than the ester concentration or by keeping the $H^+$ or $OH^-$ concentrations essentially constant through the use of buffers. This would cause the previous equation to reduce to:

$$\frac{d(ester)}{dt} = -k(ester)$$

which represents a kinetic expression for a first-order reaction. Whenever possible, first-order kinetic expressions have been employed in the study of the degradation of drugs by ester hydrolysis, but at times, second-order kinetic expressions have been employed.

A number of reports in the literature deal with detailed kinetic studies of the hydrolysis of pharmaceutical ingredients containing an ester group in the molecule. Probably one of the earliest and most thorough studies was performed on aspirin by Edwards. He studied the degradation of aspirin in various buffer solutions by treat the overall reaction as pseudo-first-order.



The data in Table 28.2 represent a summary of the rates of degradation over a wide pH.

**Table 28.2:** Aspirin hydrolysis at varying pH at 17°C

| pH | k ($days^{-1}$) | pH | k ($days^{-1}$) |
|---|---|---|---|
| 0.53 | 0.578 | 6.0 | 0.120 |
| 1.33 | 0.0835 | 6.98 | 0.10 |
| 1.80 | 0.045 | 8.00 | 0.13 |
| 2.48 | 0.0267 | 9.48 | 0.321 |
| 2.99 | 0.0343 | 10.5 | 1.97 |
| 4.04 | 0.088 | 11.29 | 13.7 |
| 5.03 | 0.130 | 12.77 | 530 |

From a plot of the log *k* versus pH as presented in Fig. 28.13, Edwards was able to postulate a reaction mechanism and determine the influence of pH on the degradation. The pH of optimum stability is at 2.4. At a pH of 5 to 7, the degradation reaction was essentially pH-independent, and at a pH above 10, the stability of aspirin was found to decrease rapidly with increase in pH. In the area in which the degradation is pH-independent, there are several reactions going on, each causing an effect of its own resulting in a cancellation of the effect of $H^+$ and $OH^-$, which gives a uniform rate over this pH range.

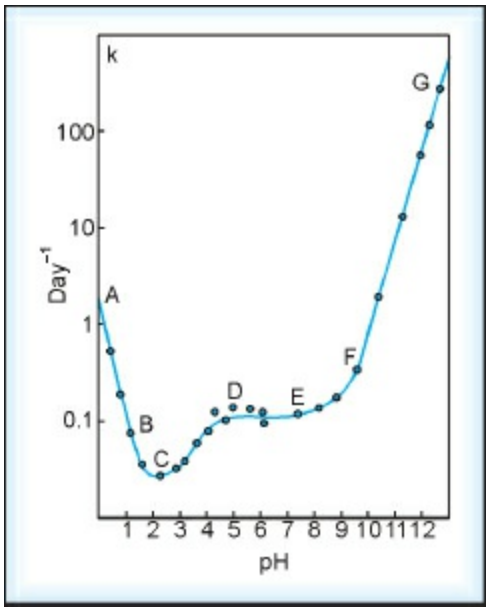



**Fig. 28.13:** Overall velocity constant for aspirin hydrolysis at 17°C as a function of pH

Although the use of pseudo-first-order kinetics is sufficient to define and study the degradation of aspirin, the hydrolysis of aspirin proceeds through a complex mechanism over the pH range studied, consisting of six different degradative pathways as shown below.

Other pharmaceutical materials that have been reported to degrade through ester hydrolysis are procaine, atropine, and methyl p-aminobenzoate. These examples serve to illustrate the importance of chemical kinetic studies in evaluating the degradative pathways and overall stability of pharmaceutical compounds containing an ester group in the molecule.

$CH_3COOC_6H_4COOH + H_3O^+ \xrightarrow{k_1} HOC_6H_4COOH + CH_3COOH + H^+$ (at low pH)

$CH_3COOC_6H_4COOH + H_2O \xrightarrow{k_2} HOC_6H_4COOH + CH_3COOH$ (uncatalyzed)

$CH_3COOC_6H_4COOH + OH^- \xrightarrow{k_3} HOC_6H_4COOH + CH_3COO^-$

$HOC_6H_4COO^- + CH_3COOH$ (pH indipendent)

$CH_3COOC_6H_4COO^- + H_3O^+ \xrightarrow{k_4} HOC_6H_4COOH + CH3COOH$

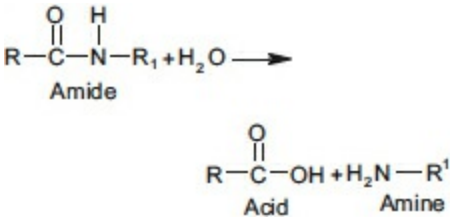
$CH_3COOC_6H_4COO^- + H_2O \xrightarrow{k_5} HOC6H4COOH + CH3COO^-$

$HOC_6H_4COO^- + CH_3COOH$ (uncatalyzed)

$CH_3COOC_6H_4COO^- + OH^- \xrightarrow{k_6} HOC_6H_4COO^- + CH_3COO^-$ (at high pH)

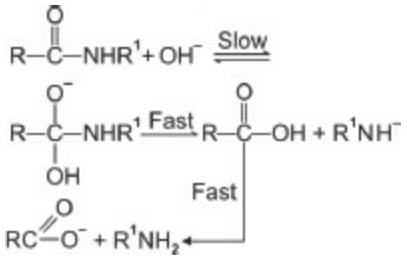
## *Amide Hydrolysis*

Pharmaceutical compounds containing an amide group can undergo hydrolysis in a manner similar to that of an ester type compound. Instead of the acid and alcohol that form as a result of ester hydrolysis, hydrolytic cleavage of an amide results in the formation of an acid and an amine.



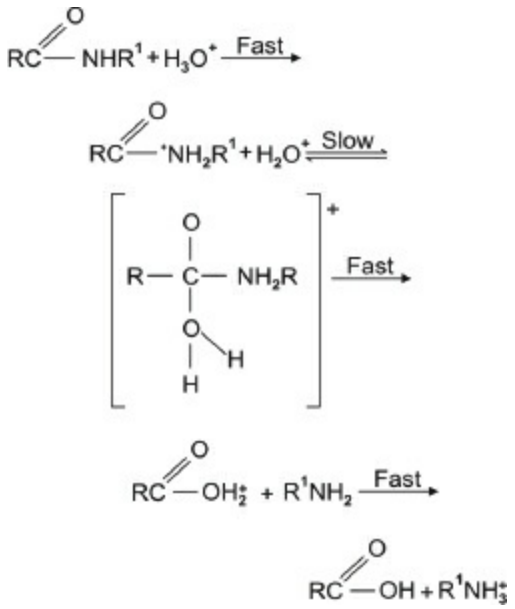

Because of the relatively greater stability of amides as compared with structurally similar esters, there is considerably less information in literature on quantitative chemical kinetic studies pertaining to the hydrolytic stability of such compounds. Pharmaceuticals such as niacinamide, phenethicillin, barbiturates, and chloramphenicol degrade by amide hydrolysis.

In a report on the stability of salicylamide and some N-substituted derivatives, Kosky postulated both basic and acid hydrolysis as mechanisms for degradation. The basic hydrolysis proceeded as follows:




The rate-determining step in the hydroxideion-catalyzed reaction is the nucleophilic attack by the hydroxide ion.

The acid hydrolysis is as follows:



The mechanism for acid hydrolysis of amides requires that substituents should exert only weak polar effects, but suitably situated, they should exert strong steric effects. The effect of alkyl and aminoalkyl substituents on the amide nitrogen in retarding the rate of acid hydrolysis of salicylamide appears to be due primarily to steric hindrance.

As can be seen in Fig. 28.14 in the acid medium, salicylanilide was more stable than salicylamide, which in turn was more stable than benzamide. Aminoalkyl substituents on the nitrogen increased the stability of benzamide. Salicylamide was more stable in basic than acidic medium, probably owing to the protection afforded by the negative charges on the phenolate ion. The N-alkyl and N-amino alkylsalicylamides were highly resistant to acid and base hydrolysis. This appeared to be due to combined steric hindrance by the hydroxyl group in the ortho position and the alkyl and aminoalkyl group on the nitrogen.

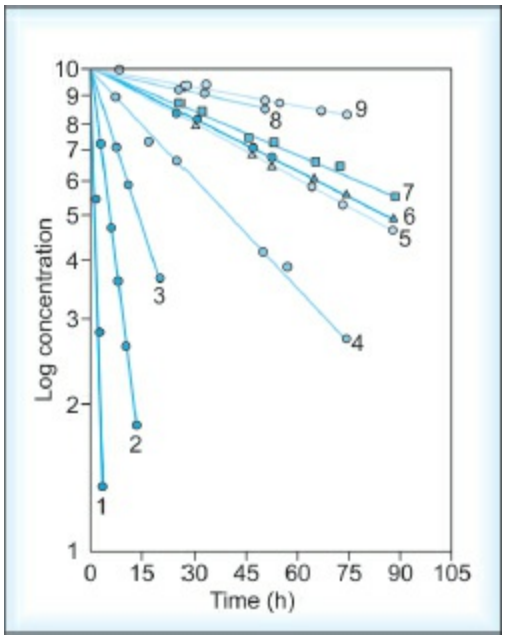
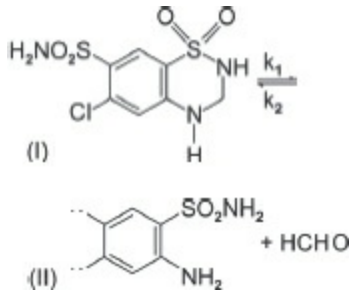



**Fig. 28.14:** Pseudo-first-order plot of the hydrolysis of amides in 1.0 N.ON perchloric acid at 90°. Key: 1. Benzamide, 2. Salicylamide, 3. N-(2-diethylaminoethyl) benzamide, 4. Salicylanilide, 5. N-(2-diethylaminoethyl) salicylamide hydrochloride, 6. N-(2-dimethylaminoethyl)-salicylamide hydrochloride, 7. N-(2-disopropylamino-ethyl)-salicylamide hydrochloride, 9. N-isopentyl-salicylamide and 9. N-propylsalicylamide

## *Ring Alteration*

A hydrolytic reaction can proceed as a result of ring cleavage with subsequent attack by hydrogen or hydroxyl ion. Examples of drugs that have been reported to undergo hydrolysis by this mechanism include hydrochlorothiazide, pilocarpine, and reserpine. Quite often, equilibrium kinetics is associated with such mechanisms.

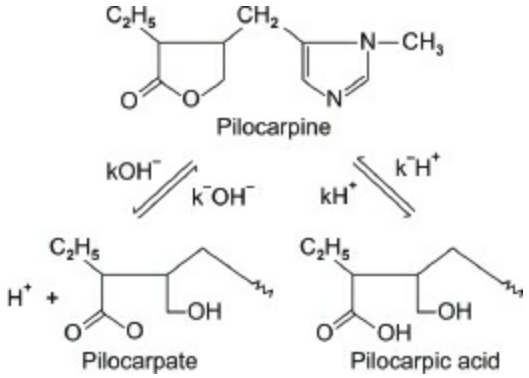



Mollica et al. reported that the hydrolysis of hydrochlorothiazide involved

reversible kinetics in which the rate of forward reaction was influenced by pH, but the equilibrium constant was independent of pH. The hydrolytic reaction was reported to proceed by ring opening to form an imine, which undergoes attack by water or hydroxide ion to yield a carbinolamine intermediate, which further decomposes to formaldehyde and 4-amino-6-chloro-m-benzenedisulfonamide, as shown by the following scheme:

$$I \rightleftarrows R - N = CH_2 \rightleftarrows R - NH - CH_2OH \rightleftarrows II + HCHO$$

The hydrolysis of pilocarpine in aqueous solution has been reported to involve a cyclic equilibrium process, which is catalyzed by hydrogen ion and hydroxyl ion. Although uncertainty exists as to whether both the hydrogen ion and hydroxide ion catalysis are equilibrium processes, the concentration of pilocarpate and pilocarpic acid are influenced by pH. One of the schemes postulated for the cyclic mechanism is as follows:




Pilocarpine is relatively stable in solutions of acidic pH, $k_H = 1.35 \times 10^{-1}$ liters/mole/hr. As the pH increases, pilocarpine progressively becomes unstable, $k_{OH} = 7.56\ 10^2$ liters/mole/hr. Phosphate and carbonate buffers catalyze the degradation, whereas borate does not. The addition of methylcellulose improves the stability slightly.

## *Protection From Hydrolysis*

As a result of the realization that a considerable number of drugs degrade through hydrolysis, methods to enhance the stability of pharmaceuticals undergoing this type of degradation have been under study. The following are factors to be considered:

## *1. pH*

If physiologically permissible, the solution of the drug should be formulated as close as possible to its pH of optimum stability. In the event that the hydrolytic degradation of the drug is general acid and base catalyzed, that is, that the degradation is catalyzed by the acid and basic species of the buffer salt in addition to $H^+$ and $OH^-$, the buffer concentration should be kept at a minimum.

## *2. Type of Solvent*

Partial or full replacement of water with a solvent of lower dielectric constant generally causes considerable decrease in the velocity of hydrolysis. Examples of these nonaqueous solvents are ethanol, glycols, glucose, and mannitol solutions and substituted amides.

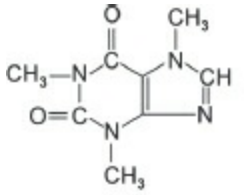
Contrary to this generalization, Marcus and Taraszka found that aqueous solutions of chloramphenicol containing up to 50% propylene glycol had no effect on improving the stability of this antibiotic over that obtained with solutions of the antibiotic in water. In fact, a slight increase in the rate of reaction was observed. Consequently, as illustrated in this study, it is unwise to make a blanket assumption that replacement of all or part of the water in a pharmaceutical preparation enhances the stability of an active ingredient. Instead, each situation must be individually evaluated with due consideration given to the mechanism of degradation. An instance in which the use of propylene glycol was found to retard amide hydrolysis is given in a report by Bodin and Taub, who show that the stability of pentobarbital in solution is effectively enhanced by the use of a propylene glycol solvent system.

## *3. Complexation*

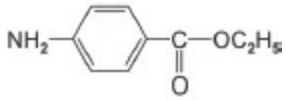
The hydrolytic rates may be influenced in two ways by complex formation, namely, by either steric or polar effects. Obviously, the attachment of a large caffeine molecule, for example, on a benzocaine molecule, can greatly affect the frequency and ease of encounter of the ester with various catalytic species ($H^+$, $OH^-$) through steric hindrance. The reaction also may be affected by the electronic influence of the complexing agent, which can alter the affinity of the ester carbonyl ion for the catalytic species. In general, the steric effect would be expected to decrease the hydrolytic rate, whereas the electronic effect may increase or decrease the reaction velocity.

There have been several reports on the influence of complexing agents in

retarding the hydrolytic deterioration of esters. Higuchi and Lachman, Lachman et al., and Lachman and Higuchi have shown that caffeine complexes with local anesthetics, such as benzocaine, procaine, and tetracaine, cause a reduction of the velocity of their hydrolytic degradation. These investigators have also shown that the complexed fraction of the ester undergoes essentially no degradation.



Caffeine



Benzocaine

Consequently, if it were possible to the total amount of drug in solution, it might be possible to stabilize it completely. Because of the limited solubility of caffeine, it has not been possible to accomplish this in the studies employing the hydrochloride salts of the local anesthetics. Guttman reported that the velocity of the base-catalyzed decomposition of riboflavin was decreased by the presence of caffeine in solution. It was found that the vitamin in its complexed form with caffeine possessed negligible reactivity toward alkaline hydrolysis.

## *4. Surfactants*

Using benzocaine as an example, Riegelman studied the effect of surfactants on the rate of hydrolysis of esters. He found that nonionic, cationic, and anionic surfactants stabilize the drug against base catalysis, as evidenced by the data in Table 28.3.

**Table 28.3:** Influence of surfactant on benzocaine degradation at 30°C using 0.04N NaOH

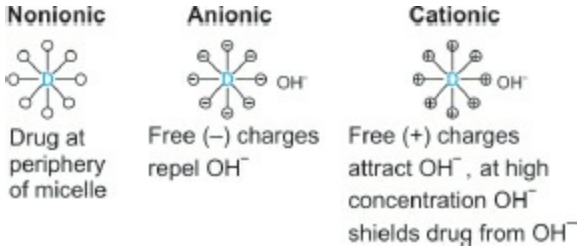
| Half-life ($t_{1/2}$) in minutes | Nonionic (%) | Cationic (%) | Anionic (%) |
|---|---|---|---|
| 64 | 0 | 0 | 0 |
| 188 | 1.33 | | |
| 324 | 3.3 | | |
| 57 | | 0.067 | |
| 425 | | 1.34 | |
| 650 | | 2.46 | |
| 420 | | | 1 |
| 1150 | | | 5 |

A 5% sodium lauryl sulfate solution (anionic) caused an 18-fold increase in the halflife of benzocaine. The association of benzo-caine close to the anionic head group of the surfactant made a definite barrier to the approach of the hydroxyl group into the micelle and attack on the ester linkage. When 2.46% cetyl trimethyl ammonium bromide in solution (cationic) is used, a tenfold increase in the half life of benzocaine results. This effect is rather interesting, but can possibly be explained by the fact that although the negatively charged hydroxyl ion is attracted by the cationic group, it apparently cannot penetrate beyond the polar head into the deeper confines of the micelle wherein the benzocaine appears to be held. When a nonionic surfactant at 3.3% concentration is used, only about a fourfold to fivefold increase in half life was obtained for benzocaine, indicating that the nonionic surfactant is a less effective stabilizer than the anionic or cationic ones. Because of the relatively high degree of hydration at the surface of the nonionic surfactant micelle, it would appear that considerable hydrolytic attack could take place within the micelle, as well as in the aqueous phase. However, this explanation of micelle protection against hydrolytic degradation of pharmaceutical compounds warrants further exploration.

## *5. Modification of Chemical Structure*

A number of reports in literature show that certain substituents added to the alkyl or acyl chain of aliphatic or aromatic esters or to the benzene ring of aromatic esters cause a decrease in the hydrolytic rate. This may be attributed to a steric and/or polar effect of the substituent group. For example, by increasing the length of, or by branching, the acyl or alkyl chain, the rate of hydrolysis of the ester usually decreases owing to steric hindrance. However, if an electrophilic or nucleophilic group is introduced into the acyl or alkyl side chain of aliphatic or aromatic esters, or on the benzene ring of aromatic

esters, the rate of hydrolysis can be increased or decreased by the electronic effect of these groups. For example, alkaline hydrolysis of aromatic esters is promoted by the presence of electrophilic groups on the benzene ring (halogen or $NO_2$), which attract electrons away from the reaction site (ester groups). The hydrolysis is retarded, on the other hand, by nucleophilic groups ($CH_3$, $OCH_3$ and $NH_2$), which cause electrons to move toward the point of reaction. The reverse effect would be found in the case of hydrogen-ion-catalyzed hydrolysis of aromatic esters.




In general, base-catalyzed hydrolytic reactions are more affected by polar effects of substituents than is acid-catalyzed hydrolysis. On the other hand, the steric retardation of acid-catalyzed hydrolysis caused by substituents is greater than for base-catalyzed hydrolysis. The total effect produced by substituents in alkaline hydrolysis, however, is considerably greater than the effect produced in acid ester hydrolysis. This is probably accounted for by the fact that in alkaline ester hydrolysis, both polar and steric effects of the substituents occur, whereas in acid ester hydrolysis, the polar effect is almost negligible.

In pharmaceutical practice, it is generally not possible to employ substituents on drug molecule for improving stability against hydrolytic cleavage of the ester group, because in most cases, these substituents also have an effect on the physiologic activity of the drug molecule. The dipivalate ester of epinephrine, however, helps to protect the catechol ring from undergoing oxidation, thus enhancing the stability of the topical ophthalmic solution of epinephrine.

## *6. Salts and Esters*

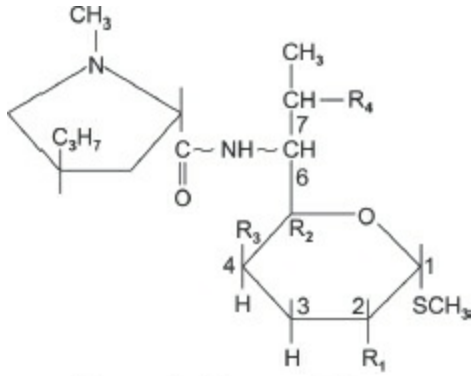
Another technique that is sometimes employed to increase the stability of pharmaceuticals undergoing degradation through hydrolysis is to reduce their solubility by forming less soluble salts or esters of the drug. Usually, only the

fraction of the drug that is in solution undergoes hydrolytic degradation. Garrett, in his study with acyl-salicylates, found that a compound that shows rapid hydrolysis in solution may be made to exhibit better stability than a more stable analog by reducing its solubility.

Transient derivatives are nontoxic additions to drug molecules, such as hydrolyzable esters, which remain intact long enough to improve the drug bioavailability and then cleave to allow the parent compound to exert its recognized biologic activity. A transient derivative is a more soluble and/or stable form of the parent compound and permits better absorption for improved and more reproducible bioavailability. In this case, the drug modification undergoes biotransformation or hydrolysis at physiologic pH to yield the active form of the drug.

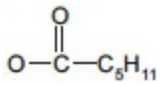
Monoesters of the antibiotics lincomycin and clindamycin have been prepared to render soluble and stable compounds suitable for injection. At pH 7.4, the antibiotics undergo biomodification to yield the active undissociated forms.

Monoesters of lincomycin with faster rates of hydrolysis were found to have greater in vivo antibacterial activity.



Lincomycin, R groups = OH

Substitution of a hexanoate group:




The 2nd position gives rise to hydrolysis rates at pH 7.4 in intestinal fluid, which are significantly greater than hexanoate substitution in the 3, 4 or 7 positions. The difference in enzymatic hydrolysis rates is attributed to the

different steric and electronic environments of the four hydroxyl groups. Steric hindrance at the 2 position was also demonstrated by the observation that (3,3-dimethyl) butyrate hydrolyzed much more slowly than the hexanoate ester.

The phosphate esters of clindamycin undergo biomodification to release the active undissociated form. The inactive compound, clindamycin phosphate, on treatment with dephosphorylating enzymes affords the active compound clindamycin. Figure 28.15 shows the percentage of hydrolysis of clindamycin phosphate esters in various enzymatic systems. The 3-phosphate ester was found to hydrolyze much more slowly and much less extensively than the 2-phosphate ester. An in vivo study in rats revealed lower blood levels for the 3-phosphate, which are probably related to the rate and degree of hydrolysis.

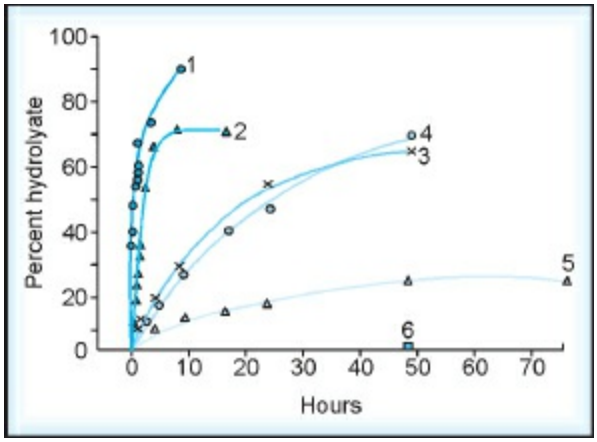



**Fig. 28.15:** Percentage of hydrolysis of clindamycin phosphate esters in enzyme systems

1. Clindamycin–2$PO_4$ in alkaline phosphatase.
2. Clindamycin–2$PO_4$ in rat liver homogenate.
3. Clindamycin–2$PO_4$ in human plasma.
4. Clindamycin–3$PO_4$ in alkaline phosphatase.
5. Clindamycin–3$PO_4$ in rat liver homogenate.
6. Clindamycin–3$PO_4$ in human plasma.

## *Oxidation-reduction*

The oxidative decomposition of pharmaceutical compounds is responsible for the instability of a considerable number of pharmaceutical preparations. For example, steroids, vitamins, antibiotics, and epinephrine undergo oxidative degradation. These reactions are mediated either by free radicals or by molecular oxygen. Because of the complexity of oxidation and their sensitivity to trace metal and other impurities, it is difficult to reproduce them or to establish mechanisms for the reactions. Consequently, many reports dealing with oxidation-reduction reactions are qualitative in nature rather than quantitative.

A substance is said to be oxidized if electrons are removed from it. Thus, a substance is oxidized when it gains electronegative atoms or radicals or loses electropositive atoms or radicals. Oxidation often involves the addition of oxygen or the removal of hydrogen. The simplest type of oxidation is, therefore, the elimination of an electron, as in the process:

$$Fe^{++} \rightarrow Fe^{+++} + e^{-}$$

where the ferrous ion is oxidized to the ferric ion.

The most common form of oxidative decomposition occurring in pharmaceutical preparations is autoxidation, which involves a free radical chain process. In general, autoxidation may be defined as the reaction of any material with molecular oxygen. Free radicals are produced by reactions involving homolytic bond fission of a covalent bond, so that each atom or group involved retains one of the electrons of the original covalent bond. This may be depicted as follows:

$$A{:}B \rightarrow A\cdot + B\cdot$$
$$CH_3{:}\,CH_3 \rightarrow 2CH_3$$

These radicals are highly unsaturated and readily take electrons from other substances, causing oxidation. The autoxidation of an organic substance RH by a free radical chain process can be simply described as follows:

*Initiation*

$$RH \xrightarrow[light,\,heat]{activation} R\cdot + (\dot{H})$$

*Propagation*

$R\cdot + O_2 \rightarrow RO_2$

$RO_2\cdot + RH \rightarrow ROOH + R\cdot$

Hydroperoxide Decomposition:

$ROOH \rightarrow RO\cdot + \cdot OH$

*Termination*

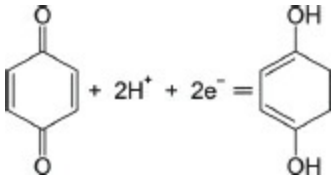
$RO_2\cdot + X \rightarrow$ inactive products

$RO_2 + RO_2 \rightarrow$ inactive products

The initiation of this reaction can be produced by the thermal decomposition of substances naturally present or added to the reaction mixture or possibly by light. As shown above, termination of the reaction may take place by combining two $RO_2\cdot$ radicals or by *X*, a free radical inhibitor. In the latter case, *X* generally converts the peroxy radical $RO_2\cdot$ to a hydroperoxide and becomes a resonance stabilized radical incapable of continuing the chain. Generally free radicals can best be terminated by a free radical inhibitor (e.g. sodium metabisulfite, thiourea, cysteine hydrochloride), since otherwise the product of recombination of radicals could contain sufficient energy to redissociate the molecule. In autoxidative reactions, only a small amount of oxygen is needed to initiate the reaction, and thereafter, oxygen concentration is relatively unimportant.

Heavy metals, particularly those possessing two or more valency states, with a suitable oxidation-reduction potential between them (copper, iron, cobalt, and nickel) generally catalyze oxidative deteriorations. These metals reduce the length of the induction period (the time in which no measurable oxidation occurs) and increase the maximum rate of oxidation. They can affect the rates of chain initiation, propagation, and termination, as well as the rate of hydroperoxide decomposition. In each case, their major function is to increase the rate of formation of free radicals.

Many oxidation processes are catalyzed by hydrogen and hydroxyl ions. This can partly be ascribed to the fact that the redox potential for many reactions depends on pH. This is particularly true for pharmaceutical compounds falling under the classification of weak acids. The system quinone/hydroquinone may be taken as a classic example to illustrate this point.



The oxidation potential may be expressed by the following simplified version of the Nernst equation:

$$E = E_0 + \frac{0.06}{2} log \frac{C_{H+}^2 \cdot C_{qutnone}}{C_{hydroqutnone}}$$

where $E_0$ is the so-called standard potential, $E$ is the actual potential, 2 equals the number of electrons taking place in the change from the ox-form to the red-form, and 0.06 is a calculated approximate constant. It can be seen from this equation that an increase in the concentration of hydrogen ions causes an increase in the value of E. This means that the red-form of the system is less readily oxidized when the pH is low. Since pharmaceuticals that undergo deterioration through oxidation are generally in the red-form, minimum decomposition is usually found in the pH range of 3 to 4.

Although oxygen concentration is of importance in the autoxidation process, its significance is usually not adequately considered. When studying the rate of the reaction for an oxidation at different temperatures, it is necessary to consider both the direct effect of the temperature and the effect of temperature on the oxygen content (concentration of oxygen) of the liquid. For example, the transfer of a preparation from storage at 15 to 5°C, with a temperature coefficient of 2 for a 10°C change, causes the rate to be reduced to half its initial magnitude, owing to the direct temperature dependence of the reaction. Simultaneously, the concentration of oxygen increases by about 25%, usually resulting in an increased rate of oxidation. Examples of pharmaceuticals that degrade through oxidative pathways are shown in Table 28.4.

For the most part, oxidative degradations of pharmaceutical compounds follow first-order or second-order kinetic expressions. Guttman and Meister studied the base-catalyzed degradation of prednisolone and found that the degradation exhibited a first-order dependency on steroid concentration. The rate of prednisolone disappearance from aqueous solutions increased with an increase in hydroxyl ion concentration under both aerobic and anaerobic

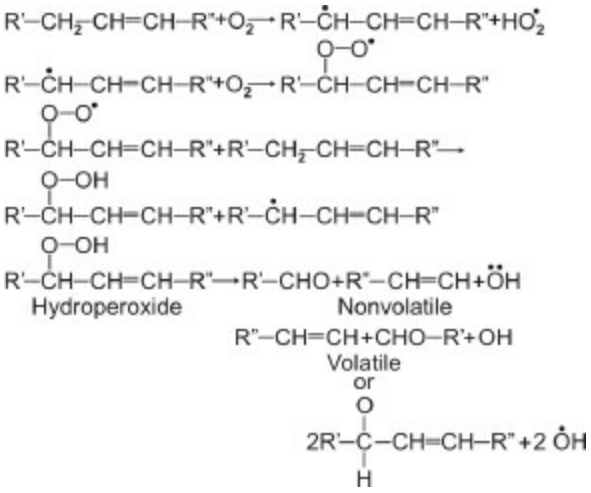
conditions; however, the reaction mixture exposed to air showed more rapid degradation of prednisolone.

**Table 28.4:** Functional groups susceptible to oxidation

| Functional group | Drugs/excipients |
|---|---|
| Aldehydes | Paraldehyde, flavor |
| Amines | Clozapine |
| Carboxylic acids | Fatty acids |
| Conjugated Dienes | Vitamin A |
| Ethers | Diethyl ether |
| Nitrites | Amyl nitrite |
| Phenols | Catecholamines |
| | Morphine |
| Thioethers | Chlorpromazine |
| Thiols | Dimercaprol |

*Rancidity*, which affects nearly all oils and fats, is a widely known term covering many typical off-flavors formed by the autoxidation of unsaturated fatty acids present in an oil or fat. These off-flavors have a more or less distinct odor and are due to the volatile compounds that are formed upon oxidation of the oils and fats. These volatile compounds are generally short-chain monomers that are formed by cleavage of the nonvolatile hydroperoxide primary oxidation product. The free radical mechanism shown here depicts the oxidation of oils and fats that takes place in the presence of atmospheric oxygen, light, and trace amounts of catalysts.

Determination of iodine numbers can be employed as an indication of whether oxidation has taken place across the double bond.




## *Protection from Oxidation*

The stability of pharmaceutical compounds undergoing oxidative degradation can be increased by several approaches.

## *1. Low Oxygen Content*

Since, in many cases, oxidative degradation of a drug takes place in aqueous solution, it is helpful to keep the oxygen content of these solutions at a minimum. In a study to determine the oxygen content of water prepared and treated by different techniques, it was found that water in equilibrium with atmospheric oxygen contains 5.75 ml per liter of oxygen at 25°C and 9.14 ml per liter of oxygen at 4°C, and that all free oxygen is expelled from water at 100°C.56 Freshly distilled water collected directly from the distillation apparatus and stored at 4°C in closed containers contains 1/4 to 3/4 the oxygen content of water saturated with oxygen from the atmosphere. Boiling and cooling this freshly distilled water to 20°C results in almost a twofold increase in the oxygen content of the water; however, if the water is cooled in an atmosphere essentially free from atmospheric oxygen, there is no increase in oxygen content. To obtain water that contains a minimal amount of free oxygen, water after it is first boiled is purged with carbon dioxide or nitrogen gas. The oxygen content of water treated with carbon dioxide is reduced to 0.45 ml per liter at 20°C.

Since most oxidative degradations of pharmaceutical compounds are probably autoxidative in nature and involve chain reactions that require only a small amount of oxygen for initiating the reaction, reduction of oxygen

concentration alone is not sufficient in many cases to prevent degradation from occurring. The traces of oxygen left may be sufficient to start a chain reaction. Consequently, it is necessary to add agents such as antioxidants and chelating agents to obtain acceptable protection against oxidative degradation.

## *2. Antioxidants*

Antioxidants are added to pharmaceutical formulations as redox systems possessing higher oxidative potential than the drug that they are designed to protect, or as chain inhibitors of radical induced decomposition.

In general, the effect of antioxidants is to break up the chains formed during the propagation process by providing a hydrogen atom or an electron to the free radical and receiving the excess energy possessed by the activated molecule.

Although the selection of an antioxidant can be made on sound theoretic grounds based on the difference in redox potential between the drug and antioxidant, electrometric measurements only rarely predict the efficiency of antioxidants in complex pharmaceutical systems. The effectiveness of an antioxidant or the comparative value of various antioxidants for a particular pharmaceutical preparation is best accomplished by subjecting the pharmaceutical system with the antioxidant to standard oxidative conditions and periodically assaying the formulation for both drug and antioxidant. Although this method may require maximum effort, it yields the most useful information.

It should be remembered that because of the complexity of free radical oxidative processes and their sensitivity to trace amounts of impurities, attempts to compare the effectiveness of antioxidants among different pharmaceutical systems are of limited validity.

The effectiveness of these antioxidants can depend on the concentration used, whether they are used singularly or in combination, the solution pH, and the package integrity and nonreactivity. The effectiveness of antioxidants can also be enhanced through the use of synergists such as chelating agents. These agents tend to form complexes with the trace amounts of heavy metal ions inactivating their catalytic activity in the oxidation of medicaments. Examples of some chelating agents are ethylenediamine tetraacetic acid derivatives and salts, dihydroxyethyl glycine, citric acid, and tartaric acid.

Trace metal impurities in buffer salts causes an accelerated decomposition of prednisolone, which was first thought to be due to buffer concentration. By studying the oxidative degradation with and without 0.1% disodium salt of ethylenediamine tetraacetic acid at different buffer concentrations, it was found that the solutions not containing any chelating agent degraded more rapidly as the buffer concentration increased, while the solutions containing chelating agent showed that the rate of degradation was independent of the concentration of the buffer. This is clearly shown by the graphs in Fig. 28.16.

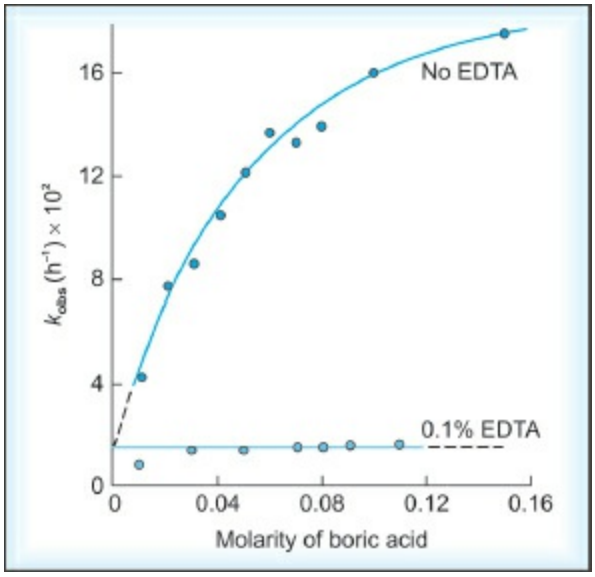



**Fig. 28.16:** The effect of buffer concentration on the rate of prednisolone degradation in the presence and absence of sequestrene $Na_2$ at 30°, pH 10, and ionic strength 0.2

## *3. pH*

It is also desirable to buffer solutions containing ingredients that are readily oxidizable to a pH in the acid range. This causes an increase of the oxidation potential of the system with a concurrent increase in stability when oxidations are catalyzed by hydrogen or hydroxyl ions. The pH of optimum stability in the acid range, however, must be determined experimentally for each drug.

## *4. Solvents*

Solvents other than water may have a catalyzing effect on oxidation reactions

when used in combination with water or alone. For example, aldehyde, ethers, and ketones may influence free radical reactions significantly.

## *Photolysis*

Consider decomposition of pharmaceutical compounds resulting from the absorption of radiant energy in the form of light, that has become more important in recent years because of the complex chemical structure of many new drugs. Degradative reactions, such as oxidation-reduction, ring rearrangement, or modification and polymerization, can be brought about by exposure to light at particular wave lengths. According to the equation $E = 2.859 \times 105/\lambda$ kcal per mole, the shorter the wave length ($\lambda$) of light, the more energy is absorbed per mole. Consequently, the radiations absorbed from the ultraviolet and violet portions of the light spectrum are more active in initiating chemical reactions than those absorbed from the other longer wave length portions of the spectrum.

In a large number of systems that are photolyzed, free radicals are products that undergo subsequent reactions. If the molecules absorbing the radiation take part themselves in the main reaction, the reaction is said to be a *photochemical* one. Where the absorbing molecules do not themselves participate directly in the reaction, but pass on their energy to other molecules that do, the absorbing substance is said to be a *photosensitizer*.

The kinetics of photochemical reactions is more complicated than the kinetics of thermal reactions because more variables are involved. The intensity and wave length of light and the size and shape of the container may greatly affect the rate of reaction. A photochemical reaction may be accompanied by a thermal reaction that is identical to the photochemical reaction opposite to it, or entirely different in character. A photochemical reaction may produce a catalyst, which then causes a thermal reaction to proceed at a measurable rate. Sometimes an induction period is necessary while a sufficient quantity of catalyst is being accumulated, to make the reaction proceed with a measurable velocity. A thermal reaction, once started, may continue after the illumination is stopped, giving an aftereffect. The energy available in a photochemical reaction is much greater than in a thermal reaction, and this fact often changes the character of the reaction.

As a result of the complexity of photolytic reactions, investigations in this area of pharmaceutical stability have been, for the most part, qualitative in

nature. Only within recent years have photodegradative studies been performed on a quantitative basis. In photodegradative reactions, second-order, first-order, and zero-order reactions are possible.

Felmeister and Dischler studied the photodecomposition of chlorpromazine hydrochloride at 253.5 mp and derived the mechanism for the free radical-mediated deterioration of chlorpromazine. The ultraviolet irradiation of chlorpromazine causes the degradation to proceed through a semiquinone free radical intermediate. The semiquinone free radical disproportionates in aqueous media in both the absence, as well as the presence of dissolved oxygen, though the loss of free radical is slightly faster in the presence of oxygen. The loss in concentration of chlorpromazine versus photons of light per liter absorbed shows that degradation follows zero-order kinetics.

Hamlin et al. studied the photolytic degradation of alcoholic solutions of hydrocortisone, prednisolone, and methylprednisolone exposed to ordinary fluorescent lighting. The plots in Fig. 28.17 show that the degradation follows first-order kinetics and that prednisolone and methylpred-nisolone show about the same rate of degradation, whereas hydrocortisone degrades about 1/7 the rate of the other two steroids. Hence, the two double bonds present in prednisolone and methylprednisolone make these steroids more susceptible to light-catalyzed degradation than the one double bond in the A ring of hydrocortisone.

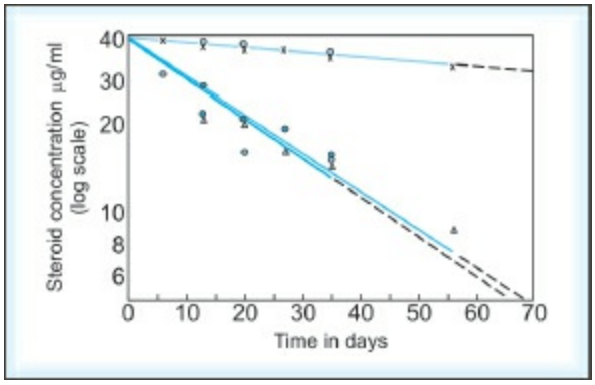



**Fig. 28.17:** First-order-plots of the photolytic degradation of steroids, when exposed to laboratory fluorescent lighting. Hydrocortisone: X, INH assay; ∘, UV assay. Prednisolone: ▲, INH assay; ▫, UV assay. Methylpredni-solone; •, INH

## *Racemization*

In such a reaction, an optically active substance loses its optical activity without changing its chemical composition. This reaction is important to the stability of pharmaceutical formulations, since the biological effect of the dextro form can be considerably less than that of the levo form. For example, levo-adrenaline is 15 to 20 times more active than dextro-adrenaline. Solutions of levo-adrenaline form a racemic mixture of equal parts of levo- and dextro-adrenaline with a pharmacologic activity just over half that of the pure levo-compound.

The kinetics of racemization may be studied in a manner similar to hydrolytic reactions. By determining the rate constant, temperature dependency of the reaction, and dependence of the reaction on pH, it is possible to establish the optimal conditions for storage of the preparation. Racemization reactions, in general, undergo degradation in accordance with first-order kinetic principles. The racemization of a compound appears to depend on the functional group bound to the asymmetric carbon atom; aromatic groups tend to accelerate the racemization process.

# STABILITY TESTING OF PHARMACEUTICAL PRODUCTS

There are legal, moral, economic, and competitive reasons, as well as reasons of safety and efficacy, to monitor, predict, and evaluate drug product stability. Pharmaceutical products (e.g. tablets, capsules, creams, injectables, and inhalations) and the components of a drug product, that is, the active pharmaceutical ingredient and the inactive ingredients should be stable in the drug product for the proposed shelf-life duration of the drug product. Shelf-life (retest or expiration date) is the time period during which active and the inactive ingredients and drug products are expected to remain within approved specification, provided that they are stored under the conditions defined on the container label. Stability applies to chemical, physical, microbiological, toxicological and therapeutic properties. Stability and storage conditions are determined by evaluation of quality parameters with time under the influence of a variety of environmental factors such as temperature, humidity, and light. A stable product is one, which retains its chemical integrity and label potency (chemical properties) within the specified limits, retains its appearance, palatability, uniformity, dissolution, and suspendability (physical properties), is resistant to microbial growth (microbiological stability) and whose therapeutic effect remain unchanged without any significant increase in toxicity, during the shelf-life.

Section 505 (b) of the CFR law as stated on the New Drug Application Form specifically describes the requirements for stability information as:

“A complete description of, and data derived from studies of the stability of, the drug, including information showing the suitability of the analytical methods used. Describe any additional stability studies under way or contemplated. Stability data should be submitted for any new drug substance, for the finished dosage form of the drug in the container in which it is to be marketed, and if it is to be put in solution at the time of dispensing, for the solution to be prepared as directed. State the expiration date(s) that will be used on the label to preserve the identity, strength, quality, and purity of the drug until it is used. If no expiration date is proposed, the applicant must justify its absence”.

Good manufacturing practice (GMP) regulations and guidelines allude to the importance of stability and shelf-life of pharmaceutical products. In

addition to GMP requirements, several guidelines are published by organizations with international mandates such as the WHO and International Conference on Harmonization (ICH), specifically to address pharmaceutical stability. The ICH has published a comprehensive series of stability guidances covering different aspects of stability: for example, ICH Q1A (R2), Stability Testing of New Drug Substances and Products; ICH Q1B, Photostability Testing of New Drug Substances and Products; ICH Q1C, Stability Testing of New Dosage Forms; ICH Q1D, Bracketing and Matrixing Designs for Stability Testing of New Drug Substances and Products; and ICH Q1E, Evaluation of Stability Data. Some of these guidelines have been adopted and published by the FDA.

In this section, the stability and shelf life of pharmaceutical drug products is reviewed, with information derived from GMP regulations and guidelines as well as technical and regulatory guidance documents on stability. The importance of photostability studies for stability—indicating method development and for appropriate container closure selection to assure stability of the marketed product through shelf-life is discussed. Recommended storage conditions for various climatic zones are discussed. The testing parameters/attributes for evaluation of stability of various drug dosage forms are outlined.

## Stability Requirements

The guidance for industry Q1A (R2), defines requirements for submission of stability data with drug application in the EU, Japan and the United States.

### *Selection of Batches and Container Closure System*

At least three primary batches of the drug product, with the manufacturing process similar to the production batches intended for marketing are used for generating stability data. The three batches of the drug product should use different batches of the drug substance. Where possible, at least two batches should be pilot scale and the third batch can be of the smaller size. For stability studies, drug product should be packaged in the same container closure system as proposed for marketing of the drug product. The strength and container size of the proposed packaging configuration should be placed on stability, unless bracketing and matrixing designs are used.

### *Matrixing and Bracketing Design*

Matrixing and bracketing designs are used to achieve reduced testing while at the same time generating enough stability data for evaluation of shelf life. At a subsequent sampling time point, different sets of samples of the total number would be tested. The design assumes that the stability of the samples tested represents the stability of all batches of the product. The differences in the samples for the same drug product should be identified as, different batches, different strengths, different sizes of the same container and closure, and possibly, in some cases different containers/closure systems. Matrixing can cover reduced testing when more than one variable is being evaluated. For example, a one-half reduction in time points eliminates one in every two time points from full study design, and one-third reduction eliminates one in every three time points. However, such a scenario must include full testing at initial, 12 month, and at the end of the long-term testing period (36 month shelf-life study).

The bracketing design assumes that the stability of any intermediate levels is represented by the stability of the extremes tested. Where a range of strengths is to be tested, bracketing is applicable if the strengths are identical or very closely related. For example, in a three batch stability study with dosage strengths of 50, 75 and 100 mg in 15, 100 and 150 ml containers,

testing for 50 and 100 mg strengths in 15 and 150 ml containers may be adequate. Matrixing design may involve elimination of some stability sample pull time points to achieve reduced testing strategy.

## *Storage Conditions and Testing Frequency*

The stability guidance provides information on storage conditions and testing frequency for the drug products under six intended storage conditions (1) drug products intended for storage at room temperature, (2) drug products packaged in impermeable containers, (3) drug products packaged in semipermeable containers, (4) drug products intended for storage in a refrigerator, (5) drug products intended for storage in a freezer and (6) drug products intended for storage below −20°C. The storage conditions are summarized in Table 28.5.

**Table 28.5:** Storage conditions for stability evaluation of pharmaceutical products

| Stability study type | Stability storage conditions | Minimum time period covered by data at submission (months) |
|---|---|---|
| *Marketed drug product intended for room temperature storage conditions* | | |
| Long-term | 25°C ± 2°C, 60% RH ± 5% RH or 30°C ± 2°C, 65% RH ± 5% RH | 12 |
| Intermediate | 30°C ± 2°C, 65% RH ± 5% RH | 6 |
| Accelerated | 40°C ± 2°C, 75% RH ± 5% RH | 6 |
| *Marketed drug product packaged in semipermeable containers* | | |
| Long-term | 25°C ± 2°C, 40% RH ± 5% RH or 30°C ± 2°C, 35% RH ± 5% RH | 12 |
| Intermediate | 30°C ± 2°C, 65% RH | 6 |

| | | |
|---|---|---|
| | ± 5% RH | |
| Accelerated | 40°C ± 2°C, not more than 25% RH | 6 |
| *Marketed drug product intended for storage in refrigerator* | | |
| Long-term | 5°C ± 3°C | 12 |
| Accelerated | 25°C ± 2°C 60%, RH ± 5% RH | 6 |
| *Marketed API intended for storage in freezer* | | |
| Long-term | –20°C ± 5°C | 12 |

The drug product is said to undergo significant change if:

a. A 5% change in assay from initial value is observed or if product fails to meet the acceptance criteria for potency.
b. Any degradation exceeding acceptance criterion is observed.
c. Unexpected change in physical attributes such as color, phase separation, resuspendability, caking, hardness is observed.
d. Product fails to meet acceptance criterion for pH.
e. Product fails to meet acceptance criterion for dissolution.

The sampling frequency and testing in stability studies should be targeted to generate data sufifcient to establish a stability profile of the drug product. For long-term storage stability studies, the guidance recommends testing every 3 months over the first year, 6 months over the second year, and annually thereafter through the proposed shelf life for drug products, where the proposed shelf life is at least 12 months. For 6 months accelerated storage stability studies, sampling at 0, 3, and 6 months is recommended. When significant changes are likely to occur under the accelerated storage condition, increased testing is required with inclusion of a fourth sampling point. If intermediate storage and testing become a necessity, a minimum of 6 months of data from this study should be submitted with the application. When significant changes are likely to occur under the accelerated storage condition, testing at the intermediate storage condition for 12 months with sampling at time 0, 6, 9 and 12 is recommended.

## Climatic Zones Concept

For convenience in planning for packaging and storage, and for stability studies, international practice identifies four climatic zones, which are described in Table 28.6. The United States, Europe and Japan are characterized by zones I and II. The values in each climatic zone are based on on observed temperatures and relative humidities from which mean kinetic temperatures (MKT) and average humidity values are calculated. Earlier to 2005, Climatic Zone IVA and IVB were combined and designated as Climatic Zone IV, which was proposed to have hot and humid climate, 30°C MKT and 75% RH. India was proposed to belong in climatic zone IV. Finally manu facturers must decide whether they are seeking a local/single country application versus a global approach. This decision will largely define the stability conditions. An example of a generic, global protocol is given in Table 28.7.

**Table 28.6:** Climatic zones and associated mean kinetic temperatures and average humidity values

| Climatic zones | Climatic conditions | Mean kinetic temperature | Yearly average relative humidity | Typical countries |
|---|---|---|---|---|
| **Zone I** | Moderate/ Temperate | 21°C | 45 % RH | Japan, UK, Canada, US, Russia, Northern Europe |
| **Zone II** | Mediterranean/ Subtropical | 25°C | 60 % RH | Japan, US, Southern Europe |
| **Zone III** | Hot and dry | 30°C | 35 % RH | Iran, Iraq, Sudan |
| **Zone IVA*** | Hot and humid/ Tropical | 30°C | 65 % RH | Brazil, Ghana, Indonesia, Nicaragua, Philippines |
| **Zone VIB*** | Hot and very humid | 30°C | 75 % RH | |

* Earlier to 2005, Climatic Zone IVA and IVB were designated as Climatic Zone IV

**Table 28.7:** Generic global stability protocol

| Test condition | Climatic zone | Temp./Humidity Temp./Humidity | Time points (months) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | 0 | 1 | 3 | 6 | 9 | 12 | 18 | 24 | 36 |
| Long-term | I and II | 25°C/60% RH | √ | | √ | √ | √ | √ | √ | √ | √ |
| | III and IVA | 30°C/65% RH | √ | | √ | √ | √ | √ | √ | √ | √ |
| | IVB | 30°C/75% RH | √ | | √ | √ | √ | √ | √ | √ | √ |
| Intermediate* | I and II | 30°C/65% RH | √ | | √ | √ | √ | √ | √ | √ | √ |
| Accelerated | I, II, III, IV | 40°C/75% RH | √ | | √ | √ | | | | | |
| Stress | — | 50°C | √ | √ | √ | | | | | | |

* No intermediate storage condition is recommended for Zones III and IV

## Stability Testing Practice

Stability testing is included at all stages of the drug product life cycle from early stages of product development to late stage follow-up stabilities. In particular the life cycle can be divided into 6 different stages:

**Stage 1:** *Early stage stress and accelerated testing with drug substances* or *API*, to gather information about physical and chemical properties (solubility profile, hygroscopicity, thermal and chemical stability) and to determine a preliminary re-test period and storage conditions.

**Stage 2:** *Stability on preformulation batches*.

**Stage 3:** *Stress testing on scale-up batches*.

**Stage 4:** *Accelerated and long-term testing*, to uncover all kinds of potential degradation products found after long-term storage and is required for registration purposes.

**Stage 5:** *On-going stability testing*.

**Stage 6:** *Follow-up stabilities*.

Depending on the stage of stability, product type and dosage form, the product is analyzed at intervals for various parameters. These parameters may include assays for the active ingredient, measurement of known degradation products, dissolution time, appearance, etc. Additionally, samples from production lots of approved products are retained for stability testing in case of product failures or production process changes. Retained samples can be tested alongside returned samples to determine whether the root cause of the problem was manufacturing or storage related. Some specific stability test for various dosage forms are as follows:

*Tablets:* Dissolution (or disintegration, if justified), water content, hardness/friability.

*Hard gelatin capsules:* Brittleness, dissolution, microbial bioburden and water content.

*Softgelatin capsules:* Dissolution, microbial bioburden, pellicle formation, pH, and leakage.

*Emulsions:* Phase separation, pH, viscosity, microbial bioburden, size and size distribution of dispersed globules.

*Solutions and suspensions:* Precipitate formation, clarity test, pH, 

viscosity, microbial bioburden, extractables, and polymorphic conversion when applicable. Additional tests for suspensions include redispersability, rheology, size and size distribution of particles.

*Powders for oral solutions and suspensions:* Water content, reconstitution time.

*Suppositories:* Softening range, dissolution at 37°C.

*Transdermal Patches:* Release rates, leakage, microbial bioburden, peel and adhesive forces.

*Metered dose inhalers and nasal aerosols:* Content uniformity, aerodynamic particle size distribution, microscopic evaluation, water content, leak rate, microbial bioburden, valve delivery, extractables, leachables from plastic and elastomeric components.

*Nasal sprays:* Clarity, microbial bioburden, pH, particulate matter, unit spray medication content uniformity, droplet and/or particle size distribution, weight loss, pump delivery, microscopic evaluation of suspensions, particulate matter, extractables, leachables from plastic and elastomeric components of container closure and pump.

*Topical, ophthalmic and otic preparations:* Clarity, homogeneity, pH, resuspendability (for lotions), consistency, viscosity, microbial bioburden, and water loss. For ophthalmic and otic products additional attributes should include sterility, particulate matter, and extractables.

*Small volume parenterals:* Color, clarity of solutions, particulate matter, pH, sterility, endotoxins. Powders for injection solutions include clarity, color, reconstitution time and water content, pH, sterility, endotoxins/pyrogens, and particulate matter. Suspensions for injection should include additional particle size distribution, redispersability, and rheological properties. Emulsion for injection should include phase separation, viscosity, mean size, and distribution of dispersed globules.

*Large volume parenterals:* Color, clarity, particulate matter, pH, sterility, endotoxin/pyrogen, and volume.

## Design of Stability Studies

Designing a stability study requires the following aspects to be taken into consideration:

*Preliminary stability data:* The preliminary stability data may provide scientific justification to the selection of the type of stability study design. During the drug product design phase, stability data is generated to choose the final package configuration and storage conditions. Such data is helpful in the development of a good stability study design that can offer as much information as possible with as minimal data as possible.

*Drug product attributes and apecifications:* All attributes that may affect the quality of the drug product have to be included in design of a stability study. These attributes have to be tested at every time period and, thus, the number of samples needed at each time should be set accordingly. The specification for each attribute of the drug product is also required to define the acceptance criteria of the stability study.

*Test methods:* All test methods used in the stability study should have been previously qualified and validated.

*Preliminary testing data:* Any data generated during the developmental and clinical trial phases of the drug may be useful to determine the expected manufacturing process variability. The process variability is an important factor to define the sampling plan to be used in the stability study.

*Design factors:* Identification of the appropriate design factors is crucial when choosing the design of the stability study. A full design requires testing the drug product for factor combinations and at all time periods. The full design provides the largest amount of information, however, the number of required combinations increases exponentially with the number of factors. On the other hand, a reduced design requires drug product testing only at a fraction of factor combinations. When variables are higher, matrixing designs are better than bracketing designs. Failure to identify a relevant design factor may cause a significant delay in the stability study completion. A design of stability study protocol is shown in Fig. 28.18.

| Stability Protocol | | |
|---|---|---|
| **Format No.:** HU/FP/NBT/SP/001 | **Protocol No.:** S/T/AC/XY/001 | **Page:** 1 of 1 |
| **Product Name:** XY Tablets | | **Strength:** 50 mg |
| **Batch No.:** NTJ-(12)-11-A | **Batch Size:** 1,00,000 Tablets | **Mfg. Date:** 02/02/2009 |
| **Packaging Details:** | **Container Material and Type:** High density polyethylene (HDPE) bottle<br>**Capacity:** 80 cc **Closure Material and Type:** HDPE child resistant caps<br>**Pack Size:** 100 tablets per bottle<br>**Desiccant Present:** Yes **Absorbent Cotton Present:** No | |
| **Storage Conditions:** 40°C temperature and 75% relative humidity | | |
| **Stability Initiation Date:** 03/02/2009 | | |
| **Stability Testing Time Points:** 0,1, 2, 3 and 6 months interval | | |

| Sr. No. | Test to be Performed | Specifications |
|---|---|---|
| 1. | Physical Appearance | White colored, round, film coated tablet |
| 2. | Assay ( by HPLC method) | 99.8 - 99.9% |
| 3. | Dissolution Testing on 12 tablets<br>Media: 900 ml of 0.1N HCl<br>Time Points: 5,10, 15 and 30 min. | >85% of drug must dissolve in 30 min or less |
| 4. | Disintegration Time | Less than 5 min |
| 5. | Moisture Content | 1.87 – 2.05% |
| 6. | Hardness | 70 N |
| 7. | Friability | 0.12% |
| 8. | Content Uniformity Test | 98.5 - 99.8% |
| 9. | Related Substance Test<br>i) known impurity A<br>ii) known impurity B<br>iii) total unknown impurity | <br>0.002%<br>0.005%<br>0.007% |

| Prepared By | Checked By | Approved By |
|---|---|---|
| Signature: Neeraj<br>Date: 1 Jan 2012 | Signature: Shadab<br>Date: 1 Jan 2012 | Signature: Gaurav<br>Date: 1 Jan 2012 |

**Fig. 28.18:** Typical stability protocol for tablet dosage forms

*Stability data generation and handling:* In the past few years, stability testing has been revolutionized by two highly technologic advancements—high pressure liquid chromatography (HPLC) methodology for analyzing drug product stability, and the computer for stability data acquisition, storage, analysis, and reporting. Most assay methods described in the official compendial drug monographs are not stability-indicating, that is, the assay is not capable of differentiating intact drug from one or more of its degradation products. In other words, the assay lacks specificity for the active ingredient. Nearly all new drug products being developed today use HPLC as the method

of choice for analyzing the stability of the dosage form. Since the FDA has become more demanding in its requirements that product expiration dates be based on stability data obtained using stability-indicating assay methods and such data analyzed by valid statistical calculations, HPLC and the computer have become essentially indispensable tools for fulfilling these requirements.

## Photostability Studies

The testing should be done in a sequential manner with fully exposed drug product and if necessary, followed by testing the product in the immediate pack and then in the marketing pack. The photostability guidance provides recommendations for the light exposure options to which the API or drug product should be exposed. Two light sources are recommended. The first light source is the artificial daylight fluorescent lamp with a combination of ultraviolet (UV) and visible outputs, xenon, or metal halide lamps. In the second light source option, the sample should be exposed to both the cool white fluorescent and near – UV lamp. It is recommended that the near – UV lamp having a spectral distribution from 320 to 400 nm with a maximum energy emission between 350 and 370 nm be used with a significant portion of UV in both bands, that is, 320–360 and 360–400 nm. For confirmatory studies, the samples should be exposed to light providing an overall illumination of not less than 1.2 million lux hours and an integrated near – UV energy of not less than 200 $Wh/m^2$. Dark controls, covered with an aluminum foil to protect from light, should also be placed side by side with the light-exposed samples. The use of validated actino-meter solution or calibrated radio-meters or lux meters are recommended to ensure that the samples are exposed to desired light exposure. Results of the photostability study should help in (1) identification of precautionary measures needed during manufacture and packaging, (2) container closure design for protection from light and (3) storage conditions and light protection required during shelf life of the marketed product.

## Expiration Dating of Pharmaceuticals

Expiration dating of pharmaceuticals corresponds to the determination of an expiration dating period or shelf-life for drug products and a retest period for drug substances. The expiration dating period or, shelf-life, of a drug product is defined as the time interval in which the drug product is expected to remain within an approved shelf-life specification, provided that it is stored according to label storage conditions and that it is in the original container closure system. The Expiry/Expiration Date is the actual date placed on the container/labels of a drug product designating the time during which a batch of the drug product is expected to remain within the approved shelf-life specification if stored under defined conditions and after which it must not be used. To arrive at an expiration date, it must be determined first for how long and under what conditions a pharmaceutical formulation can meet all of its quality specifications. In general, this issue is answered through stability testing that monitors chemical and physical product attributes as a function of time, temperature, and other environmental factors.

Good Manufacturing Practice (GMP) requirements for drug stability (Section 211.166), expiration dating (Section 211.137), and FDA guidelines for stability studies (Section 98), contain significant and specific information related to conducting stability studies and assigning expiration dates. The important features are listed in the following summary:

1. Each product’s expiration date is related to the specific storage condition stated on the label and must be based on data obtained from an appropriate stability testing program.
2. Such a stability program would include:
    a. numbers and sizes of containers per sample time.
    b. testing of drug product in the marketed container-closure system at appropriate storage condition(s).
    c. an adequate number of batches, usually at least three production batches, to be placed on long-term stability testing for a new product initially, and one production batch per year thereafter.
3. Expiration date must be derived from data obtained using reliable and specific stability-indicating assay methods for the active ingredients).
4. Tentative expiration dates can be assigned based upon data from

accelerated stability studies as long as it can be shown that such accelerated studies are scientifically sound and that long-term studies are being conducted to confirm the predicted expiration date. It may be appropriate to utilize an increased sampling frequency toward the end of the expiration dating period. A rule of thumb for solid dosage forms allows a 2-year tentative expiration date at room temperature if the drug has retained 90% of its original potency after 90 days storage at 40°C and 75% relative humidity.

5. Long-term stability studies should attempt to determine the time at which the lower 95% confidence limit curve of the mean degradation line intersects the acceptable lower limit for drug degradation.
6. If a drug product is placed in a different packaging system, it can retain its expiration date if the repackager can assure the FDA that the repackaged container-closure system is as good as the original package and that all product specifications can be maintained throughout the dating period.

## Shelf-life Determination

The most desirable stability data are from actual shelf-life studies using products in the container-closure systems stored under labeled conditions. For introducing new products to the market, however, or for making material changes in the process, formula, or container-closure system of existing products, one cannot wait until all the needed stability data at room temperature are generated. Therefore, appropriately designed and executed short-term (e.g. 3 month) accelerated stability studies have been accepted by the FDA as data bases for use in extrapolating longer room-temperature expiration dates. Use of accelerated data is obviously not a substitute for actual shelf-life study. It is a means of predicting shelf-life of a product based on scientific principles and guided by experience. This method of shelf-life prediction based on short-term accelerated stability data is currently well-utilized by pharmaceutical scientists.

### *Arrhenius Relationship for Stability Prediction*

The Arrhenius equation has been used by pharmaceutical scientists in predicting room temperature stability of drug products based on data obtained under exaggerated conditions (higher temperature rates of degradation). The main steps in the process are:

1. Drug products be stored at several elevated temperatures.
2. Determination of the order of reaction by plotting stability data at several elevated temperatures according to the equations relating decomposition to time for each of the orders of reaction, until linear plots are obtained.
3. Value of the rate constant $k$ at each temperature is calculated from the gradient of these plots, and the logarithm of $k$ is plotted against reciprocal temperature according to the Arrhenius equation $\log k = \log A - E_a/2.303RT$. The graph in Fig. 28.19 represents a plot of $k$ values obtained at several elevated temperatures. Since the plot is linear, the prediction of stability at shelf temperature is possible by extrapolating the curve to the lower temperatures and reading off the $k$ value for the lower temperature. Once the k value is obtained, it can be used to estimate the time for $t_{10\%}$ degradation with the aid of Eq. (20).

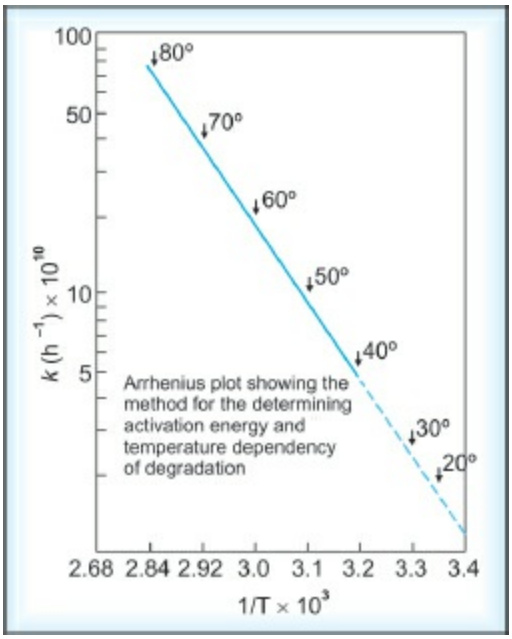



**Fig. 28.19:** Temperature dependency of degradation rates

## *Practical Example: Estimation of Degradation from Accelerated Data: First-order Case*

Table 28.8 shows the degradation data, obtained at long term, intermediate and accelerated storage conditions (*Note:* The data in this example is artificially created and will be for demonstration purpose only). All values are expressed as a percentage of the label claim (%LC). Using only the data at higher temperatures (30°C and higher), estimate the assay values at 25°C as a function of time. The long term data (25°C) is usually not available when a new drug application (NDA) is submitted and that is why it is precluded from the calculation.

**Table 28.8:** Assay degradation data for drug product

| t(months) | Assay at 25°C | t(months) | Assay at 30°C | Assay at 40°C | Assay at 50°C |
|---|---|---|---|---|---|
| 0 | 99.9 | 0 | 99.9 | 99.9 | 99.9 |
| 3 | 99.4 | 2 | 99.4 | 98.0 | 95.6 |
| 6 | 98.2 | 4 | 98.7 | 95.9 | 91.2 |
| 9 | 97.8 | 6 | 97.4 | 95.1 | 87.4 |
| 12 | 97.4 | | | | |
| 18 | 95.6 | | | | |
| 24 | 94.5 | | | | |
| 36 | 91.7 | | | | |

*Solution:* For first-order kinetics, the first step is to obtain the natural logarithm of the concentrations which is done in Table 28.9. Then, these values are graphed as a function of time where straight lines should be obtained as shown in Fig. 28.20. The slope of each line corresponds to the value of $k_1$, which is shown in the bottom part of Table 28.9. Next, a plot of ln $k_x$ versus 1/T is prepared with the three values of one for each temperature as shown in Fig. 28.21. Note that absolute temperature should be used here.

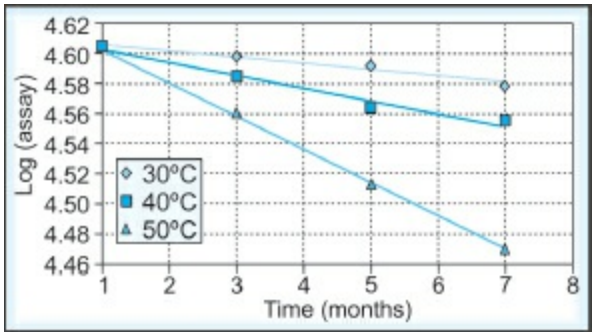



**Fig. 28.20:** Graphical representation of degradation data treated as first order kinetics

**Table 28.9:** Degradation data treated as first order kinetics

| t(months) | In(assay) at 30°C | In(assay) at 40°C | In(assay) at 50°C |
|---|---|---|---|
| 0 | 4.60 | 4.60 | 4.60 |
| 2 | 4.59 | 4.58 | 4.55 |
| 4 | 4.58 | 4.55 | 4.50 |
| 6 | 4.57 | 4.55 | 4.46 |
| $k_1$ | 0.00415 | 0.00847 | 0.02224 |
| In ($k_1$) | – 5.483 | – 4.771 | – 3.798 |
| 1/T | 0.00330 | 0.00319 | 0.00309 |

From Equation 36, the value of $k_1$ can be extrapolated to T = 25°C (1/T = 0.003354 $K^{-1}$) to then predict the drug product assay as a function of time. The values of slope and intercept are shown in Fig. 28.21. The calculation is:

In $k_{25}$ = – 8237.4 × 0.003354 + 21.639 = – 5.990 $k_{25}$ = 0.00250 $month^{-1}$

Once the specific reaction rate is calculated, the drug product shelf life can be calculated using first-order shelf-life Eq. (20).

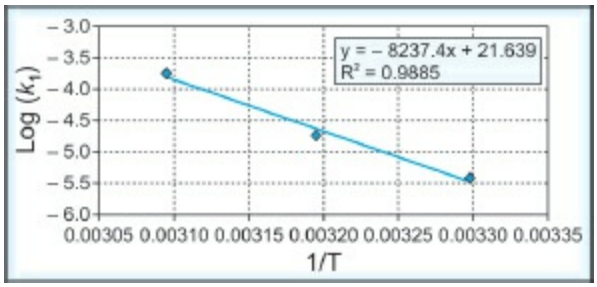



**Fig. 28.21:** Plot of In $k_1$ versus 1/T for first order kinetics

## *Practical Example: Estimation of Degradation from Accelerated Data: Zero-order Case*

*Solution:* For zero-order kinetics, the original concentrations can be graphed as a function of time (Table 28.8). Straight lines can be drawn through the points even if the actual degradation order is 1 because the degradation is small (less than 10%), as seen in Fig. 28.22. The slope of each line corresponds to the value of $k_0$. These values are shown in Table 28.10. Next, a plot of ln $k_0$ versus 1/T is prepared with the three values of $k_0$, one for each temperature, which is shown in Fig. 28.23.

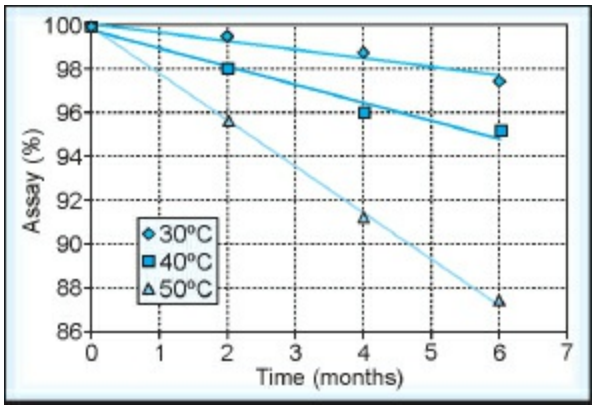



**Fig. 28.22:** Graphical representation of degradation data treated as zero-order kinetics

**Table 28.10:** Degradation data treated as zero-order kinetics

| | 30°C | 40°C | 50°C |
|---|---|---|---|
| $k_0$ | 0.410 | 0.825 | 2.095 |
| In($k_0$) | – 0.892 | – 0.192 | 0.740 |
| 1/T | 0.00330 | 0.00319 | 0.00309 |

The value of $k_0$ can be extrapolated to T = 25°C (1/T = 0.003354 $K^{-1}$) using equation 36 to then predict the drug product assay as a function of time. The values of slope and intercept are shown in Fig. 28.23; the calculation is:

$$\text{In } k_{25} = -7974.8 \times 0.0003354 + 25.369$$
$$= -1.379$$
$$k_{25} = 0.2519 \text{ month}^{-1}$$

Once the specific reaction rate is calculated, the drug product shelf-life can be calculated using zero-order shelf-life equation (Table 28.1).

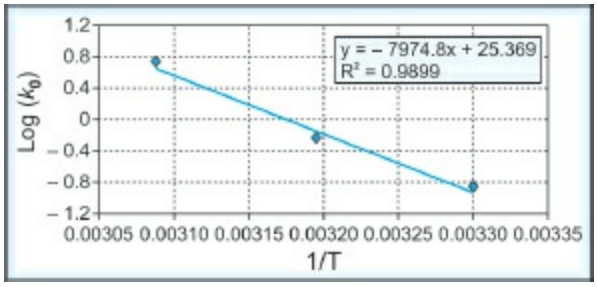



**Fig. 28.23:** Plot of In $k_0$ versus 1/T for zero-order kinetics

## *Limitations of Arrhenius Relationship for Stability Prediction*

There are a variety of situations in which Arrhenius predictions can be erroneous or invalid. Higher temperatures may evaporate solvents, thus producing unequal moisture concentrations at different temperatures. Also at higher temperatures, there is less relative humidity and oxygen solubility, thus hindering the predictability of room temperature stability of drugs sensitive to the presence of moisture and oxygen. For disperse systems, viscosity is decreased as temperature is increased, and physical characteristics may be altered, resulting in potentially large errors in prediction of stability. Different degradation mechanisms may predominate at different temperatures, thus making stability prediction marginal at best.

## *Other Techniques for Stability Prediction*

Simplified graphic techniques have been employed to predict the breakdown that may occur over prolonged periods of storage at normal shelf conditions. Free and Blythe describe such a technique for liquid products, where the decomposition behaves according to the general kinetic laws. For example,

the plots in Fig. 28.24 show that the degradation is following a first-order reaction. The time for the loss lines at several temperatures to reach 90% of the theoretic potency is noted by arrows on the curve. These time values at different temperatures are plotted in Fig. 28.25, and the time for 10% loss of potency at room temperature can be obtained from the resulting straight line by extrapolation to 25°C. If the extrapolated data in Fig. 28.25 show that the time to reach 90% potency at room temperature is too rapid to provide an adequate shelf life for the product, it is possible to determine the average required for the product to maintain at least 90% potency for a prescribed time. This is accomplished by drawing the loss line representative of the 90% potency value at room temperature, as shown in Fig. 28.26. Then a line is drawn parallel to this from the desired shelf life back to "0" days. The example shown in Fig. 28.26 indicates that by the use of a 10% overage, the product now takes about twice as long to fall below 90% of labeled claim during shelf storage.

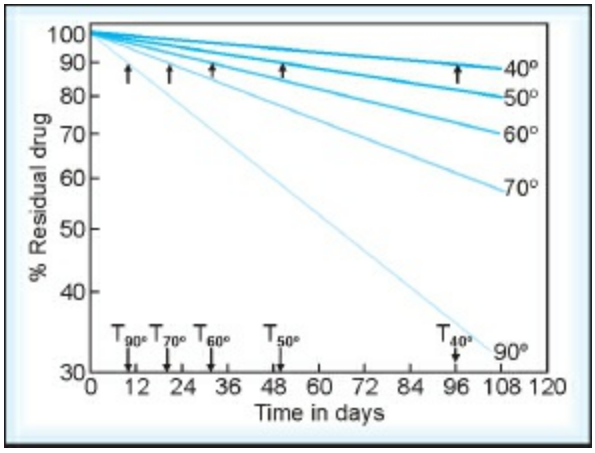



**Fig. 28.24:** Values of $t_{10\%}$ at several temperatures

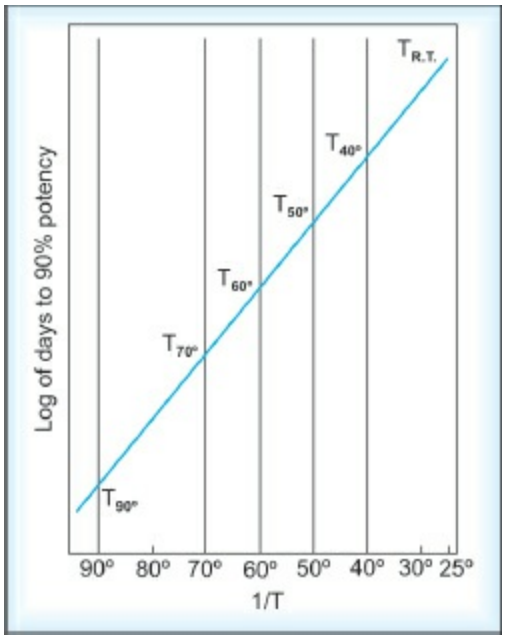



**Fig. 28.25:** Plot of $t_{10\%}$ values vs. absolute temperature$^{-1}$

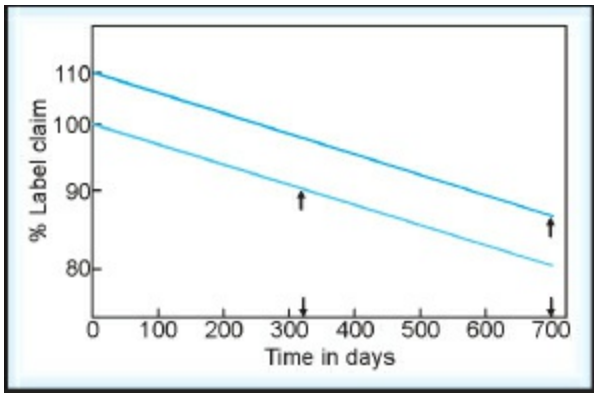



**Fig. 28.26:** Plot of average and normal loss curves

**Kennon** described the construction of certain kinetic paths, which can be used for purposes of comparison during formulation development work. Using standard kinetic equations, he calculated the paths that reactions would follow if a 10% potency loss in two years at room temperatures were permitted. By choosing activation energies of 10 and 20 kcal/mole, both of which are conservatively low, and by plotting the time in months that a formulation would take to drop to 90% potency versus 1/T, one arrives at Fig. 28.27. Table 28.11 presents the data used in Fig. 28.27.

**Table 28.11:** Maximum and minimum time at which potency must be at least

| 90% of label claim at the temperature indicated in order to predict a shelf-life of two years at room temperature | | |
|---|---|---|
| **Temperature** | **Maximum** | **Minimum** |
| 37° | 12 months | 6.4 months |
| 45° | 8.3 months | 2.9 months |
| 60° | 4.1 months | 3 weeks |
| 85° | 6 weeks | 2.5 days |

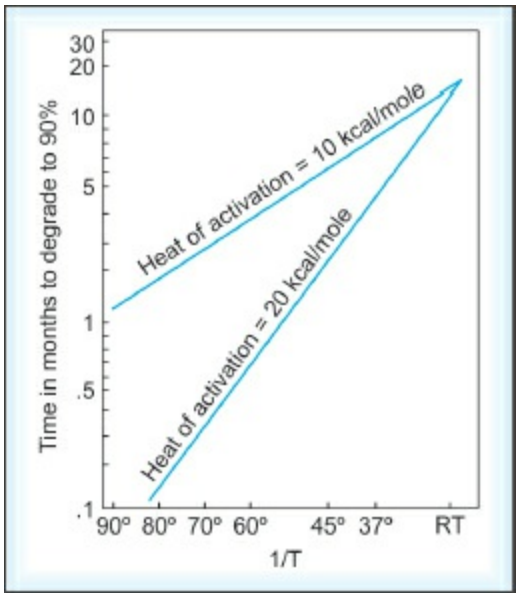



**Fig. 28.27:** Two-year shelf-life goal reference decomposition

If the potency of the formulation is found to remain above 90% of its original concentration after storage at the various temperatures for certain periods of time given in the graph and table, there is good assurance that the formulation will meet the requirement of a two-year shelf-life. Thus, if the assays are over 90% of original concentration at the minimum times shown (indicated by the 20 kcal/mole line on the graph) at the respective temperatures, in all probability, the assays will be over 90% after two years at room temperature. If the assays remain over 90% at the maximum times shown (indicated by the 10 kcal/mole line on the graph), it is a certainty (kinetically speaking) that a potency of over 90% will be maintained after two years at room temperature.

It is evident from the foregoing discussion that considerable information

can be gained on the stability characteristics of a drug through the use of certain physicochemical principles. Since most pharmaceutical preparations are complex, the degradation reaction may be complicated by possible interaction of the several ingredients in the formulation. It becomes impractical and is usually unnecessary to perform thorough basic kinetic studies on the final formulation to obtain an estimate of the shelf-life of the product. It is usually sufficient to follow the degradation or some property of the degradation as a function of time at several elevated temperatures, using the kinetic expressions presented, and then to extrapolate the data to ambient conditions to obtain an estimate of the shelf-life of the product.

Thermal stability of pharmaceutical solutions and suspensions can be estimated by applying an *accelerated non-isothermal kinetic method*. At suitable intervals, time and temperature are recorded, and samples of drug product are placed in a thermostated water bath whose temperature is increased at programmed intervals. The drug samples are assayed and plotted at log concentration versus time at a particular temperature. The points of the non-isothermal degradation curve, shown in Fig. 28.28, are fitted according to a polynomial regression equation:

$$f(t) = a + bt + ct^2 + dt^3$$

where $f(t)$ is the concentration function, and a, b, c and d are coefficients. The rate constants at each temperature are then calculated from the first derivative of this equation, which represents the slope of the tangent line at each point of the curve. Arrhenius plots are then generated, and the predicted rate constant at room temperature is calculated.

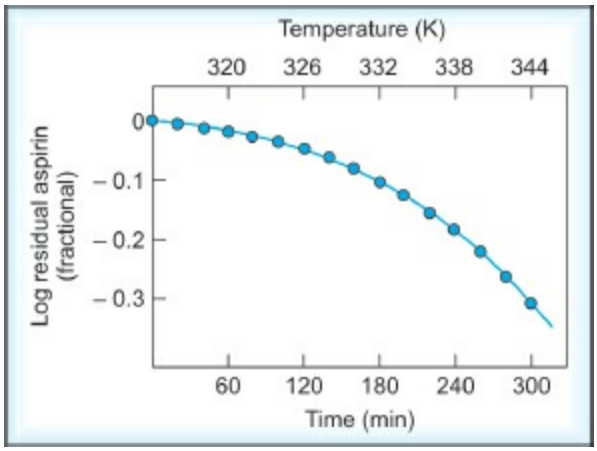


**Fig. 28.28:** The concentration-time-temperature curve for the degradation of aspirin in 0.1N HCl at 40–70°C. Key: ∘, experimental points;—, mathematical fit (correlation coefficient = 0.9999)

The advantage of non-isothermal kinetic studies is the short time period required to generate the data for estimating stability. Disadvantages include the need for programmable and sophisticated temperature control equipment and the decrease in experimental accuracy in predicting product shelf-life.

The importance of stability testing in the development of pharmaceutical dosage forms is well recognized in the pharmaceutical industry. Increased filings of Abbreviated New Drug Applications (ANDA) and Paper New Drug Applications (PNDA) by generic and nongeneric drug manufacturers has resulted in an increase in submissions of stability data to the Food and Drug Administration (FDA). With the advent of the biotechnologic age, as the bioengineered products become ready for testing in humans, stability test data for these compounds are required as part of the submissions of Investigational New Drug Applications (INDs) to the FDA to assure their quality and safety. This increase in stability testing has come at a time in which the empiric methods have, for the most part, been replaced by a more scientific approach to stability evaluation using various appropriate physical and chemical principles.

# 29:
# Quality Management: Quality Control and Assurance

In Article 2 of the World Health Organization (WHO) Constitution, one of the Organization's functions is that, it should "develop, establish and promote international standards with respect to food, biological, pharmaceutical and similar products". The quality of pharmaceuticals has been a concern of the (WHO) since its inception. The concept of total quality control refers to the process of striving to produce a perfect product by a series of measures requiring an organized effort by the entire company to prevent or eliminate errors at every stage in production. In the pharmaceutical industry today, we are concerned with the quality system's efficiency as well as its effectiveness. Rising regulatory and customer requirements make this ever more challenging. Without assurance of meeting acceptable standards of quality, safety and efficacy, any healthcare service is evidently compromised. Although the responsibility for assuring product quality belongs principally to quality assurance personnel, it involves many departments and disciplines within a company. To be effective, it must be supported by a team effort. Quality must be built into a drug product during product and process design, and it is influenced by the physical plant design, space, ventilation, cleanliness, and sanitation during routine production. The product and process design begins in research and development, and includes preformulation and physical, chemical, therapeutic, and toxicologic considerations. It considers materials, in-process and product control, including specifications and tests for the active ingredients, the excipients, the product, specific stability procedures for the product, freedom from microbial contamination and proper storage of the product, and containers, packaging, and labeling to ensure that container closure systems provide functional protection of the product against factors such as moisture, oxygen, light, volatility, and drug/package interaction. Provision for a cross referencing system to allow any batch of a product to be traced from its raw materials to its final destination in the event of unexpected difficulties is required.

## QUALITY MANAGEMENT

Traditionally in the pharmaceutical industry, the quality function is divided into two parts, quality control and quality assurance. *Quality control* (QC) involves operational techniques and activities that are used to assure product compliance to specification. In general, it is concerned with evaluating events from the past. *Quality assurance* (QA) is in general concerned with events in the present. It is focused on building quality into a product through planned and systematic activities like validation, process and environmental control, and documentation.

In the drug industry at large, quality management is usually defined as the aspect of management function that determines and implements the "quality policy". There are two schools of thought on quality management. One views quality management as the management of success and the other the removal of failure. They are both valid.

Success means not only that products, services and processes fulfil their function but also that the function is what customers desire. Failure means not only that products, services and processes would fail to fulfil their function but also that their function was not what customers desired. Ideally, if we could design products, services and processes that fulfil customer needs we would have achieved the ultimate goal. The quality management principles are comprehensive and fundamental rule or belief, for leading and operating an organization, aimed at continually improving performance over the long term by focusing on customers while addressing the needs of all other interested parties. Table 29.1 document eight principles that have emerged as fundamental to the management of quality.

**Table 29.1:** Key principles in quality management

| Principle | Explanation |
|---|---|
| Principle 1–customer focus | Organizations depend on their customers and therefore they should understand and meet the current and future customer needs and strive to exceed customer expectations. |
| Principle 2–leadership | Leaders establish unity of purpose and |

| | |
|---|---|
| | direction of an organization. They create and maintain a micro-environment in which people can become fully involved to achieve organization's objectives. |
| Principle 3–involvement of people | People at all levels are the essence of an organization and their full involvement enables their abilities to be utilized for the organization's benefit. |
| Principle 4–process approach | A desired effect is realized more efficiently when related resources and activities are managed as a process. |
| Principle 5–system approach to management | Identifying, understanding and managing a system of interrelated processes for a given objective improves the organization's efficiency and effectiveness. |
| Principle 6–continual improvement | Continual improvement should be a permanent objective of the organization. |
| Principle 7–factual approach to decision-making | Effective decisions are based on the examination of data and information. |
| Principle 8–mutual beneficial supplier relationships | An organization and its suppliers are interdependent, and a mutually beneficial relationship increases their ability to create value. |

Quality management is about the future, about prevention, and management's role in improving the quality system. The components of the quality system are people, processes, and culture. People are the primary component of the quality system. It is important to have people with requisite knowledge and skills, with the potential for the task and importantly with the correct attitude. The second component of the quality system is processes that comprise the system. The management processes are critical to the quality system. The communication of standards, directions, objectives and priorities to the organization; periodic reviews of the health system's; creation of the quality culture; and ensuring the availability of adequate resources are of

prime importance. The supporting processes such as audits, change control, failure investigations, labeling, maintenance, materials management, complaint handling, product release, stability/expiration dating, training, trend analysis, and validation add value to the quality system. The most important process component is ownership, i.e. the process is actively managed. For internal processes, the supporting processes, the inputs, and outputs are generally information; information that is supplied by internal suppliers and used by internal customers. The activities in the process should be value added. The third component of the quality system is culture. Given capable processes and skilled employees, management needs to create a culture that fosters the desired behaviour and that supports the natural motivation of employees to take pride in a job well done. A good model of a quality system view quality as a total system, holds management responsible for quality, and emphasizes the importance of design and quality planning. An effective quality management system hierarchal structure consists of the following:

a. Documentation that accurately describes the organization's core competencies and provides the necessary policies, processes, procedures, and records to support the organization's quality management system.
b. Implementation of the documents, based on the operational use on a daily basis.
c. Demonstration of effectiveness based on the monitoring, testing, and analyzing of operational data and the corresponding corrective and preventive action programs.

## QUALITY APPROACHES

Quality and specifically quality improvement have become recognized as business differentiators. Over the years, most of the quality improvement tools that were developed included quality circles, zero defects, statistical process control, quality function deplopment, etc. These rarely succeeded in transforming companies since in most cases they were stand-alone techniques that were expected to resolve the problems of the company. The major change in the quality approach came about when it was realized that quality improvement should apply to every business activity and should be integrated into the business strategy—total quality. Table 29.2 evaluates three specific types of systems: a typical total quality management (TQM) example, a system based on ISO quality standard, and a system based on the Malcolm Baldrige National Quality Award (MBNQA). All the three systems can be designed to encompass all of the organization's core competencies. In addition, all three can employ action teams to measure cost of quality and to provide top management with a corrective and preventive action protocol. Although the MBNQA is a nationally recognized certification, of the three concepts, only the ISO 9000 approach to attain certification, either nationally, internationally, or both.

**Table 29.2:** Quality approaches for effective quality management

| Quality management system | | |
|---|---|---|
| **Custom TQM** | **ISO quality standard** | **MBNQA program** |
| • Worldwide application | • Worldwide application | • National application |
| • Covers core competencies | • Covers core competencies | • Covers core competencies |
| • Based on action team formation with top management oversight | • Action team formation is optional | • Based on action team formation with top management oversight |
| • Based on customized | • Based on international | • Based on national standard with |

| | | |
|---|---|---|
| standards | standard with accredited national and international certification | accredited national certification |

## ISO Quality Standard

ISO 9000 is a series of International Standards for quality management system. It is a written set of rules (a standard) published by International Organization for Standardization (ISO) and is administered by accreditation and certification bodies. The rules define universally recognized and accepted quality practices which, when well implemented, give customers confidence that suppliers can consistently meet their needs. ISO 9000 is not a product standard. It does not contain any requirement with which a product or service can comply. There are no product acceptance criteria in ISO 9000 so one cannot inspect a product against the standard. The purpose of these standards is to simplify things for organizations and to assist organizations to implement and operate effective quality management systems. These standards provide a vehicle for consolidating and communicating concepts in the field of quality management that have been approved by an international committee of representatives from national standards bodies. The rules, referred to as requirements of the ISO quality system apply to the management of the organization and hence affect the manner in which the product and services are designed, manufactured, installed or delivered. ISO 9000 requirements are generic and applicable to any organization providing any product or service, regardless of type and size. Within the organization, the impact of the requirements and the QMS are similarly broad. Every activity within the organization that impacts the process of creating customer satisfaction is affected by the requirements of the standard. The requirements fall roughly into the following types:

- Requirements that help to achieve customer satisfaction by assuring that the organization's output (whether product, service, or both) meets customer specifications.
- Requirements that assure that the quality system are consistently implemented and verifiable. We must actually do what we say, justify what we do, record and review what we did. This must be verifiable via independent, objective audit.
- Requirements for practices that measure the effectiveness of various aspects of the system.
- Requirements that support continuous improvement of the company's ability to meet customer needs.

ISO 9000 standards were prepared by technical committee ISO TC/176, quality management and quality assurance, subcommittee SC 2, quality systems. The standards in the family are:

| | |
|---|---|
| ISO 9000 | Quality Management and Quality Assurance Standards – guidelines for selection and use. |
| ISO 9001 | Quality systems – model for quality assurance in design, development, production, installation and servicing. |
| ISO 9002 | Quality systems-model for quality assurance in production, installation and servicing. |
| ISO 9003 | Quality systems-model for quality assurance in final inspec-tion and testing. |
| ISO 9004 | Quality Management and Quality Systems Elements – guidelines for quality management planning, implementation and performance improvements. |

The purpose of ISO 9000 is to provide an appreciation of the fundamental concepts of quality management systems and an explanation of the terminology used in the family of standards. Without an understanding of the terminology, the standards are prone to misinterpretation. Although they are not requirements, the context of the requirements will not be understood without an appreciation of the concepts that underpin the requirements. ISO 9001 deals with requirements for a quality management system that can be used for external quality assurance purposes. ISO 9001 is useful primarily for companies that design and develop their own products. It provides requirements which if met will enable organizations to demonstrate they have the capability to consistently provide product that meets customer, regulatory and the organization’s own requirements. Third parties can also use the standard to assess the capability of organizations to provide product that meets customer and regulatory requirements. Organizations can use ISO 9001 as a model in designing their management systems providing they also use ISO 9000 and ISO 9004.

ISO 9002 and ISO 9003 are complementary (not alternative) to the technical (product) specified requirements. ISO 9002 is applicable to manufacturers, distributors, and service vendors whose products have been designed and serviced by a subcontractor, while ISO 9003 is applicable for testing laboratories and equipment distributors. Such companies are exempt

from design control requirements. ISO 9004 should be used as guidance in designing, operating and improving a management system and overall performance of an organization. It is not intended for assessment purposes but there may be benefits in using the standard as a basis for assessing current capability. ISO 9001 and ISO 9004 have been developed as a consistent pair of standards that complement each other. They have a common structure but can be used independently. ISO 9004 is not intended as a guide to ISO 9001. Although ISO 9004 includes the requirements of ISO 9001 it does not contain an explanation of these requirements or guidance in meeting them.

## *ISO 9000 Requirements*

The process to produce an effective QMS requires the following key elements:

**Management responsibility:** The supplier's management shall define and document its policy for quality. The quality policy shall be relevant to the supplier's organizational goals and needs of its customers. The management responsibility is to ensure that this policy is understood, implemented and maintained at all levels of the organization.

**Quality system:** The supplier shall establish, document and maintain a quality system as a means of ensuring that product conforms to specified requirements. The supplier shall maintain a quality manual that includes or makes reference to the quality system procedures and outlines the structure of the documentation used in the quality system.

**Contract review:** The supplier shall establish and maintain documented procedures for contract review and for the co-ordination of these activities. Before submission of a contract, the supplier shall ensure that the contractor has the capability to meet the contract or order requirements and the requirements are adequately defined and documented.

**Design control:** The supplier shall prepare plans for each design and development activity and define responsibility for their implementation. The design and development activities shall be assigned to qualified personnel equipped with adequate resources. At appropriate stages of design, design verification shall be performed and the plans shall be updated as the design evolves. All changes and modifications shall be identified, documented, reviewed and approved by authorized personnel before their implementation. To ensure that product conforms to user needs, design validation shall be

performed.

**Document and data control:** The supplier shall establish and maintain all documents and data at all locations where operations essential to the effective functioning of the quality system are performed.

**Purchasing:** The supplier shall establish and maintain documented procedures to ensure that purchased product conforms to specified requirements. The supplier shall evaluate and select subcontractors on the basis of their ability to meet subcontract requirements including the quality system and any specific quality assurance requirements.

**Control of customer-supplied product:** The supplier shall establish and maintain documented procedures for the control of verification, storage and maintenance of customer-supplied product provided for incorporation into the supplies or for related activities. Any such product that is lost, damaged or is otherwise unsuitable for use shall be recorded and reported to the customer.

**Product identification and traceability:** The supplier shall establish and maintain documented procedures for identifying the product by suitable means from receipt, during the stages of production, delivery and installation. When traceability is a requirement, the supplier shall establish and maintain documented procedures for identification of individual products.

**Process control:** The supplier shall identify and plan the production, installation and servicing processes which directly affect quality and shall ensure that these processes are carried out under controlled conditions.

**Final inspection and testing:** The supplier shall carry out inspection and testing to complete the evidence of conformance of the finished product to the specified requirements. No product shall be dispatched until all the activities specified have been satisfactorily completed and the associated data and documentation are available and authorized. The supplier shall establish and maintain documented procedures for all inspection and testing activities.

**Control of inspection, measuring and test equipment:** Inspection, measuring and test equipment shall be used in a manner which ensures that the measurement uncertainty is known and is consistent with the required measurement capability. Documented procedures to control, calibrate and maintain inspection, measuring and test equipment (including test software) are required to be maintained by the supplier.

**Inspection and test status:** The inspection and test status of product shall be identified to indicate the conformance or nonconformance of product. The identification of inspection and test status shall be maintained, as defined in the quality plan, throughout production, installation and servicing of the product to ensure that only product that has passed the required inspections and tests is dispatched, used or installed.

**Control of nonconforming product:** The product that does not conform to specified requirements is prevented from unintended use or installation. This control shall provide for identification, documentation, evaluation, segregation, disposition of nonconforming product, and for notification to the functions concerned.

**Corrective and preventive action:** Any corrective or preventive action taken to eliminate the causes of nonconformities shall be to a degree appropriate to the magnitude of problems and commensurate with the risks encountered. The supplier shall implement and record any changes to the documented procedures resulting from corrective and preventive action.

**Handling, storage, packaging, preservation and delivery:** Procedures for handling, storage, packaging, preservation and delivery of the product shall be established, maintained and documented.

**Control of quality records:** Quality records demonstrate conformance to specified requirements and the effective operation of the quality system. Documented procedures for identification, collection, indexing, access, filing, storage, maintenance and disposition of quality records should be established and maintained by the supplier. All quality records shall be legible and shall be stored and retained in such a way that they are readily retrievable.

**Internal quality audits:** Internal quality audits to determine the effectiveness of the quality system and to verify whether quality activities and related results comply with the planned arrangements should be implemented. Internal quality audits shall be carried out by personnel independent of those having direct responsibility for the activity being audited. The results of the audits shall be recorded and the personnel responsible for the area shall take timely corrective action, if deficiencies were found during the audit.

**Training:** Personnel performing activities affecting quality shall be qualified on the basis of appropriate education, training and/or experience, as required. Appropriate records of training shall be maintained.

**Servicing:** Where servicing is required, the supplier shall establish and maintain documented procedures for performing, verifying and reporting that the servicing meets the specified requirements.

**Statistical techniques:** The supplier shall identify the need for statistical techniques required for establishing, controlling and verifying process capability and product characteristics. Appropriate documented procedures to implement and control the application of the statistical techniques shall be maintained.

Being international standards, the ISO 9000 family would be subject to review every five years. In ISO 9001:1994 there were roughly 323 requirements however, in ISO 9001:2000 the number of requirement reduces to 250. The 1994 version of ISO was primarily aimed at quality assurance. The 2000 version is aimed at customer satisfaction. The 1994 version required procedures to be established, documented and maintained. The 2000 version requires processes to achieve defined objectives. The 1994 version focused on correcting errors but the 2000 version focuses on continual improvement not only by better control but also by improving effectiveness. The 1994 version required management with executive responsibility to define its commitment to quality. The 2000 version requires top management to demonstrate its commitment to developing, implementing and improving a system of interrelated processes that will enable the organization to achieve its objectives. By far the most significant change in the ISO 9000:2000 is the change in structure from 20 elements described above to a model based on five elements. The ISO 9000:2000 family of standard is based on a process model—a model that is intended to represent a process-based quality management system. A version of this is illustrated in Fig. 29.1.

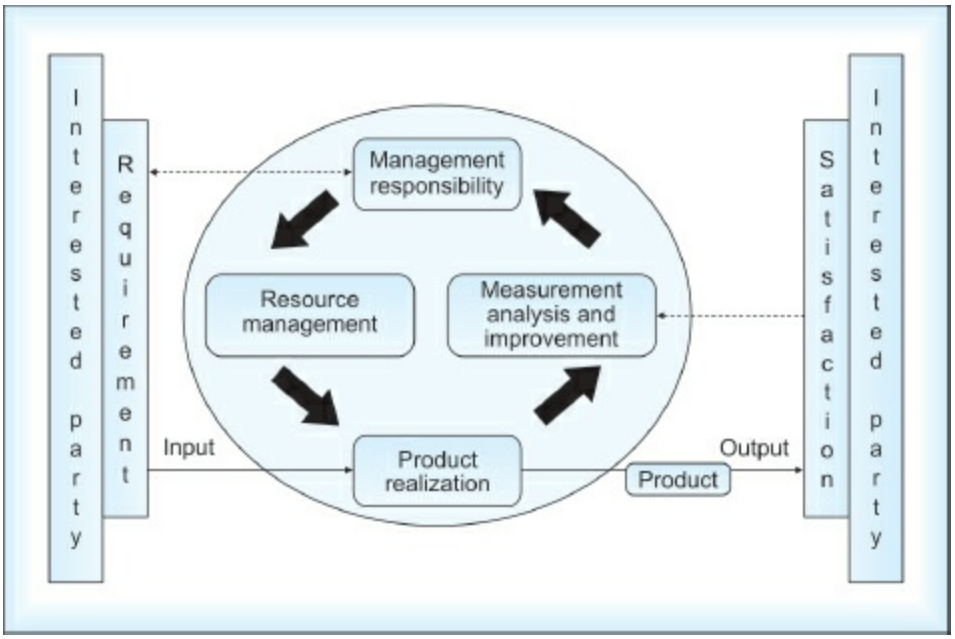



**Fig. 29.1:** Model of a process based quality management system

If the 250 or so requirements of ISO 9001 and the intentions of ISO 9004 were to be condensed into just five simple requirements they might read as follows:

*The organization shall*

- Determine the needs and expectations of customers.
- Establish policies, objectives and a work environment necessary to motivate the organization to satisfy these needs.
- Design and manage a system of interconnected processes necessary to attain the desired objectives.
- Measure and analyse the adequacy, efficiency and effectiveness of each process in fulfilling its purpose and objectives and;
- Pursue the continual improvement of the system from an objective evaluation of its performance.

Implementation of ISO 9000 affects the entire organization right from the start. If pursued with total dedication, it results in cultural transition to an atmosphere of continuous improvement. The process of implementing ISO 9000 depends on sophistication of existing quality program, size of the organization and complexity of process. The 14 essential steps, briefly described in Fig. 29.2, are to be followed through in order to implement ISO 

9000 quality management system successfully.

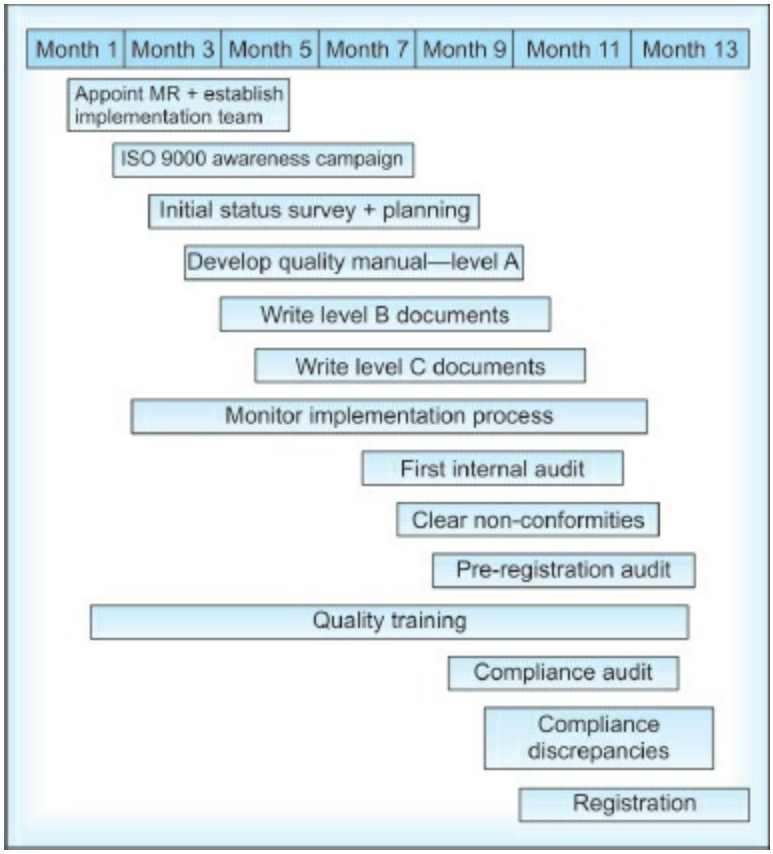



**Fig. 29.2:** Typical action plan to implement ISO 9000 quality management system

## *Impact*

ISO 9000 requirements will have considerable impact on organizations. If organizations wish to obtain or retain certification, they will need to look outwards towards their customers. The system can no longer be a bolt-on system; in existence only to serve the certification body and only to get the badge. The system has to cause results and for this to happen the system and the organization are synonymous. Organizations can no longer choose the parts of the structure or select functions or operations that will be subject to ISO 9000 certification. All parts, functions, operations and activities that serve the achievement of the organization's goal must be included. The performance of the organization relative to customer satisfaction will be monitored as a measure of the effectiveness of the quality system. Management must resource quality management system, ensure it fulfils its

purpose and strive for its continual improvement. Audits will need to focus on processes not procedures and will need to monitor performance against objectives.

## *Benefits*

A well-implemented quality system helps all functions within the process to understand their responsibility for meeting customer needs and appreciate their position in the overall process for doing so (1) It establishes and enforces consistent quality controls and working methods through out the organization and creates consistency throughout the organization, (2) Facilitates continuous improvement by admonitions to monitor, review, and improve the subprocesses of the quality system. The corrective and preventive actions required by the standard enlist all levels and functions in the effort to prevent quality problems and quickly mitigate those that do occur, (3) Provides tangible evidence to staff, customer and supplier that management is committed to quality practices and strengthens relationships between organization, its suppliers and customers, (4) A registered quality system improves customer focus and provides confidence to customers in the capability of your organization to meet quality commitments, (5) Internal audits, management reviews, analysis of organization-level data, and effective document and data control provide management with the intelligence it needs to make the right moves, (6) The levels of procedural development, documentation, record-keeping, and training required by an ISO 9000 quality system reduces dependence upon individuals, (7) ISO 9000 registration is a powerful marketing tool for its holder. Registered firms proudly display their certificate and logo, and their names appear in registries of approved firms, (8) ISO 9000 registration brings international credibility to the organization. It provides the route for a company to gain approval to use CE (communaute europeene) mark on its product which meets all the requirements of EU (European Union) product directives and (9) It warrants reduction in overall cost of quality. Although, initial prevention cost is high but this is offset by lower cost of internal and external failure thus resulting in overall value addition. ISO 9000 is deployed and practiced in nearly 100 countries around the world. In today’s evergrowing international economic climate, this is not a bad emblem to have, however narrow the scope of one’s market today.

# ASPECTS OF QUALITY MANAGEMENT

The concepts of QA, GMP and QC are interrelated aspects of quality management and their relationship can be viewed as a type of a cascade arrangement as shown in Fig. 29.3.

Quality management, with the overall policy of the organization towards quality, comes above everything else. Next comes QA, which is the unit that ensures the policy is achieved. Within an organization, QA serves as a management tool. QA checks and assures that the required product quality is achieved and the safety of the patient is guaranteed. It also covers a wide range of measures necessary to enable suitable standards of production to be maintained (premises, installations, documentation systems, self-inspections, pre- and post-audit preparations, etc.). GMP is a part of QA; it deals with the section that guarantees that the products are produced and examined within constant quality system. GMP guidelines build quality into the product. Quality control is a part of GMP: that is focused on testing of the facilities and environment, as well as the testing of the materials, components and product in accordance with the standard. The production (production department) is in the centre of the quality management system. The quality of the production process is monitored by selfinspections, release control and in-process control. Pharmaceutical production guarantees that a stable and repetitive product quality is maintained. Because of the existing connections between QA, QC and GMP, QA begins with raw material procurement and accompanies the process through development, retailing and use of the final product; therefore, also including Good Clinical Practice (GCP), GMP, Good Storage Practice (GSP), Good laboratory Practice (GLP) and further components.


**Fig. 29.3:** Model of a quality management system showing interrelationship between quality assurance, GMP, quality control and production

## QUALITY ASSURANCE

Quality assurance is part of quality management focused on providing confidence that quality requirements will be fulfilled. It is the totality of the arrangements made with the object of ensuring that pharmaceutical products are of the quality required for their intended use.

Quality assurance therefore incorporates GMP and other factors, such as product design and development. The system of quality assurance appropriate to the manufacture of pharmaceutical products should ensure that:

a. Pharmaceutical products are designed and developed in a way that takes account of the requirements of GMP and other associated codes such as good laboratory practice (GLP) and good clinical practice (GCP).
b. Managerial responsibilities are clearly specified in job descriptions.
c. All necessary controls on starting materials, intermediate products, and bulk products and other in-process controls, calibrations, and validations are carried out.
d. Production and control operations are clearly specified in a written form and GMP requirements are adopted.
e. The finished product is correctly tested and checked, according to the defined procedures and are sold or supplied after the authorized persons have certified that each production batch has been produced and controlled in accordance with the requirements.
f. Satisfactory arrangements exist to ensure, as far as possible, that the pharmaceutical products are stored by the manufacturer, distributed, and handled so that quality is maintained throughout their shelf-life.
g. There is a procedure for self-inspection and/or quality audit that regularly appraises the effectiveness and applicability of the quality assurance system.
h. Deviations are reported, investigated and recorded.
i. Regular evaluations of the quality of pharmaceutical products should be conducted with the objective of verifying the consistency of the process and ensuring its continuous improvement.

## Sources of Quality Variation

The assurance of product quality depends on more than just proper sampling and adequate testing of various components and the finished dosage form. Prime responsibility of maintaining product quality during production rests with the manufacturing department. Removal of responsibility from manufacturing department for producing a quality product can result in imperfect composition, such as missing ingredients, subpotent or superpotent addition of ingredients, or mix-up of ingredients; mistakes in packaging or filling, such as product contamination, mislabeling, or deficient package; and lack of conformance to product registration. Quality assurance personnel must establish control or checkpoints to monitor the quality of the product as it is processed and upon completion of manufacture. These begin with raw materials and component testing and include in-process, packaging, labeling, and finished product testing as well as batch auditing and stability monitoring.

Because of the increasing complexity of modern pharmaceutical manufacture arising from a variety of unique drugs and dosage forms, complex ethical, legal, and economic responsibilities have been placed on those concerned with the manufacture of modern pharmaceuticals. An awareness of these factors is the responsibility of all those involved in the development, manufacture, control, and marketing of quality products. A systematic effective quality assurance program takes into consideration potential raw material, in-process, packaging material, labeling and finished product variables.

## Control of Quality Variation

### *Raw Materials Control*

Good raw material specifications must be written in precise terminology, must be complete, must provide specific details of test methods, type of instruments, and manner of sampling, and must be properly identified. Table 29.3 lists general tests, limits, and other physical or chemical data for raw materials related to identity, purity, strength, and quality. Table 29.4 presents the quality assurance monograph for acetaminophen, USP, as an example of a specific raw material quality assurance monograph. The FDA Current Good Manufacturing Practices (cGMP) covering raw material handling procedures are found in the Code of Federal Regulations, Title 21, Section 211.42. It states that "components" be received, sampled, tested, and stored in a reasonable way, that rejected material be disposed of, that samples of tested components be retained, and that appropriate records of these steps be maintained. In practice, the manufacturer physically inspects and assigns lot numbers to all raw materials received and quarantines them until they are approved for use. Each raw material is sampled according to standard sampling procedures and is sent to the quality control laboratory for testing according to the written procedures (Table 29.4). If acceptable, it is moved to the release storage area, after being properly stickered to indicate the item number, name of material, lot number, date of release, reassay date, and signature of the quality assurance inspector. It is re-tested as necessary according to an established schedule to assure that it still conforms to specifications at time of use.

**Table 29.3:** Raw material quality assurance monograph

A. (Raw material name)

1. Structural formula, molecular weight
2. Chemical name(s)
3. Item number
4. Date of issue

5. Date of superseded, if any, or new material
6. Signature of writer
7. Signature of approval

B. Samples

1. Safety requirement
2. Sample plan and procedure
3. Sample size and sample container to be used
4. Preservation sample required

C. Retest program

1. Retesting schedule
2. Reanalysis to be performed to assure identity, strength, quality, and purity

D. Specifications (wherever applicable)

1. Description
2. Solubility
3. Identity
    a. Specific chemical tests such as related alkaloids, organic nitrogen basis, acid moiety, or inorganic salt tests; sulfate, chloride, phosphate, sodium, and potassium tests; or other spot organic and inorganic chemical tests as needed.
    b. Infrared absorption
    c. Ultraviolet absorption
    d. Melting range

e. Congealing point

f. Boiling point or range

g. Thin-layer, paper, liquid, or gas chromatography

4. Purity and quality

a. General completeness of solutions, pH, specific rotation, nonvolatile residue, ash, acid-insoluble ash, residue on ignition, loss on drying, water content, heavy metals, arsenic, lead, mercury, selenium, sulfate, chloride, carbonates, acid value, iodine value, saponification value.

b. Special quality tests, particle size, crystallinity characteristics, and polymorphic forms

c. Special purity tests, ferric in ferrous salts, peroxides and aldehydes in ether and related degradation products

5. Assay, calculated either on anhydrous or hydrous basis.

6. Microbial limits, especially for raw materials from natural sources

E. Test procedures

1. Compendial, USP, or NF references

2. Noncompendial, detailed analytical procedure, weights; dilutions; extractions; normality; reagents; instrumentation used and procedure, if any; calculations

F. Approved suppliers

1. List of prime suppliers and other approved alternative suppliers, if any.

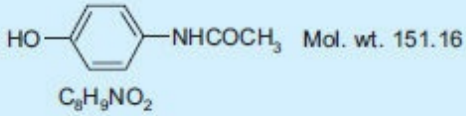
Quality assurance personnel should keep preservation samples of active raw materials that consist of at least twice the necessary quantity to perform

all tests required, to determine whether the material meets the established specifications. These preservation samples should be retained for at least 7 years. Approved materials should be rotated so that the oldest stock is used first. Any raw material not meeting specifications must be isolated from the acceptable materials, stickered as a rejection, and returned to the supplier or disposed of promptly. To verify the supplier's conformance to specifications, further supporting assurance by means of on-site periodic inspections are pertinent to the total quality of raw materials. This procedure assures that cross-contamination does not take place because of improperly cleaned equipment or poor housekeeping practices; otherwise, contaminants could go undetected because specifications generally are not designed to control the presence of unrelated materials. In general, raw materials may be classified into two groups (1) active or therapeutic and (2) inactive or inert.

**Table 29.4:** Acetaminophen, USP, quality assurance specifications

| Item No. | Date of issue | Superseded | Written by | Approved by |
|---|---|---|---|---|
| 001 | Jan 1, 1984 | New | J Doe | T Mullen |

| Sampling plan | | Retest program | |
|---|---|---|---|
| Containers | Containers to be sampled | Schedule | Tests |
| 1 | All | 2 yr | Identity IR |
| 2–5 | All | | p-aminophenol |
| 6–10 | 6 | | assay |
| 11–18 | 7 | | |
| 19–28 | 8 | | |
| 29–100 | 9 | | |
| 101 | 10 | | |




*Chemical formula*

4-Hydroxyacetanilide, p-hydroxyacetanilide, p-acetaminidophenol, p-acetaminophenol, p-acetylaminophenol, N-acetyl-p-aminophenol

*Specifications*

| | |
|---|---|
| Description | White, odorless, crystalline powder, possessing a slightly bitter taste |
| Solubility | Soluble in boiling water and in sodium hydroxide, freely soluble in alcohol |
| Identity A-IR | Scan conforms to reference standard |
| B-UV | Scan conforms to reference standard |
| C-$FeCl_3$ | Violet-blue color is produced |
| Melting range | 168–172°C |
| pH (saturate solution) | 5.3–6.5 |
| Water | NMT 0.5% |
| Residue on ignition | NMT0.1% |
| Chloride | NMT 0.014% |
| Sulfate | NMT 0.02% |
| Sulfide | No coloration or spotting of the test paper occurs |
| Heavy metals | NMT 0.001% |
| Readily carbonizable substances | No more color than matching fluid A |
| Free p-aminophenol | NMT 0.005% |
| p-Chloroacetanilide | NMT 0.001% |
| Assay | 98.0–101.0% on the anhydrous basis |
| Completeness of solution | 1 g/20 ml methanol is not less clear than an equal volume of methanol examined similarly |

*Test procedures*

For all tests, see USP XXI, United States Pharmacopeial Convention, Rockville, MD, 1980.

*Approved suppliers*

1. SB Penick and Co., Lyndhurst, NJ
2. Mallinckrodt, Raleigh, NC

## *Active Materials*

The current editions of the USP and NF contain monographs on most therapeutically active materials used in manufacturing. Since there is such a wide variance in the nature of the active ingredients used in manufacturing, it is impossible to summarize briefly the testing of those raw materials. One of the most important decisions to be made in raw material control is the degree of purity to be maintained for each material. It is not uncommon to find an appreciable variation in the degree of purity between samples of the same raw material purchased from different commercial sources. The selection must then result in the highest purity practical for each raw material, consistent with safety and efficacy of the final dosage form. In general, a typical raw material currently found in a compendium has a purity requirement of at least 97%. Its specifications normally include solubility, identification, melting range, loss on drying, residue on ignition, special metal testing, specific impurities that are pertinent to the method of synthesis of each individual raw material, and assay. The methods of assay are usually chemical in nature.

It should be understood that these compendial tests are intended as the minimum required from a legal point of view. For certain products, it may be necessary to require that the active ingredient specifications be far tighter than those of the comparable compendial standard. Raw materials cannot be adequately evaluated and controlled without special instrumentation such as spectrophotometry; potentiometric titrimetry; column, gas (GLC), paper, thin-layer, and high-pressure liquid chromatography (HPLC); polarography; X-ray diffraction; X-ray fluore-scense; spectrophotofluorimetry; calorimetry; and radioactive tracer techniques. No less demanding are the tests required for microbiologic assay, pharmacologic assay, and safety testing. For certain products, even when highly purified and well characterized raw materials are involved, specifications should include additional critical tests, such as particle size, crystal shape, and crystalline versus amorphous forms. Any of these characteristics could affect the safety or effectiveness of the final dosage form. It is a cGMP requirement that all raw materials, active or inactive, be assigned a meaningful reassay date that assures the purity and potency of the raw material at time of use. This confirms the continued stability of each raw material.

## Antibiotics

Antibiotics are one of the few drugs for which the official analytic method appears in the Code of Federal Regulations. The USP and NF refer to the Code of Federal Regulations for specifications and analytic methods given in the individual monographs for each antibiotic. The code of federal regulations, Title 21, Chapter 1, Parts 436 to 436.517 and Parts 442 and 455 contain the analytic method specifications for all antibiotics approved for human use in the United States. The number of tests required varies from one antibiotic to another. The data in Table 29.5 specify the tests required by the code of federal regulations for some antibiotics.

Testing of antibiotics is usually performed either chemically, microbiologically, biologically, or by all three methods. Caution must be exercised in testing antibiotic raw material to assure that it is not altered during the sampling procedure. The sample must be taken in a relatively dry atmosphere, relatively free from dust, and free of both chemical and microbial airborne contamination, and exposure must be reduced to minimal time of sampling. Special attention should be given to the assay for potency of antibiotic raw materials. Since the potency value in terms of micrograms per milligram obtained for this material is used in calculating the number of grams or kilograms required for the working formula procedures, it is recommended that at least two separate weighings of such antibiotic raw material powder be assayed on each of three different days (six different assays using six different weighings). If all the individual results are not within the normal distribution of the group or show too much variance, additional assays should be performed until a mean potency is obtained with confidence limits of ± 2.5% (or better) at $p = 0.05$.

**Table 29.5:** Tests of some antibiotics

| | LOD | Moisture | pH | Crystallinity | Iodometric assay* | Residue on ignition | Heavy metals | Special test | Microbial limits | ID | Specific rotation |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Ampicillin | NMT 2% | – | 3.5–6.0 | X | X | | | | Potency† Safety | X | |
| Cephalexin-monohydrate | 4–8% | – | 3.0–5.0 | X | X | | | | Potency† Safety | X | |
| Neomycin - sulfate | NMT 8% | – | 5.0–7.5 | | | | | | Potency† Safety | X | |
| Tetracycline | – | NMT 13% | 3.0-7.0 | X | | | | Absorption at 380 nm | Potency† Safety | X | |
| Zinc bacitracin | – | NMT 5% | 6.0–7.5 | | | | | | Potency† NMT 10% Zn Safety | X | |
| Griseofulvin | NMT 1% | – | | X | | NMT 0.2% | NMT 25 ppm | UV absorption specific surface area | Potency† Safety | X | +348° to +364° |
| Erythromycin | – | NMT 10% | 8.0–10.5 | X | | NMT 2% | NMT 50 ppm | | Potency† Safety | X | –50° to –58° |
| Sodium novobiocin | NMT 6% | – | 6.5–8.5 | | | 10.5–12% | | | Potency† Safety | X | |

* Chemical methods of assay as alternatives to microbiologic assay are allowed for special antibiotics: Iodometric and hydroxylamine colorimetric for most penicillin, cephalothin, and cephaloridine; GLC for lincomycin and clindamycin; ultraviolet spectrophotometry for chloramphenicol succinate and palmitate, and griseofulvin; colorimetric for cycloserine and troleandomycin.

† Potency is determined microbiologically using the turbidimetric assay and the diffusion plate assay.

## *Inactive or Inert Materials*

Inactive or inert materials usually make up the major portion of the final dosage form. Therefore, their physical characteristics, such as color, odor, and foreign matter, are as important as their chemical purity. Among other important specifications of inactive or inert materials are particle size, heavy metal content, arsenic, selenium, water content, microbial limit, foreign matter, residue on ignition, and pH.

Approved certified water-soluble Food, Drug and Cosmetics (FD and C) dyes, or mixtures thereof, or their corresponding lake, may be used to color oral dosage forms. The color in oral dosage forms is mainly a means of identification. The FDA determines and approves colorants for use in food and drugs with recommendation of limits, if any. A typical analysis of a color contains identity tests and tests of total volatile matter, heavy metals, water-insoluble matter, synthesis impurities, arsenic, lead, and total color. An FD and C color lake analysis contains additional tests for chloride, sulfate, and inorganic matter.

If a flavored oral dosage form is desired, flavors or volatile oils may be used. Flavors or volatile oils are usually tested for refractive index, specific gravity, solubility, and alcohol content, if any. A GLC chromatogram can be used as a “fingerprint” for each specific flavor to help in assuring the

supplier’s continuous compliance to specifications. Knowledge of the presence of any synthetic FD and C dyes in a flavor formula is important for the formulator and quality control personnel to assure compliance with FDA colorant regulations.

The most popular sweetening agents used are sucrose, glucose, mannitol, lactose, crystalline and liquid sorbitol, and such artificial sweetening agents as saccharin, sodium saccharin, calcium saccharin, and aspartame.

Testing for unwanted impurities resulting from synthesis side reaction in the manufacturing procedure is essential in the analysis of sweetening agents, for example, furfuraldehyde in lactose, and reducing sugars in mannitol. Sweetening agents are usually tested for water content, heavy metals, residue on ignition, arsenic, and special tests such as specific rotation, melting range, selenium, and readily carbonizable substances.

## *In-process Items Control*

Conformance to compendial standards as the sole basis for judging the quality of a final dosage form can be grossly misleading. Obviously, a compendial monograph could never cover all possibilities that might adversely affect the quality of a product. The difficulty lies in part in the fact that final dosage forms are frequently produced in batches of hundreds or thousands or even millions of units. The numbers of units assayed at the end of the process is not likely to be representative of more than a small fraction of the actual production.

There is a real and significant difference between a finished product compendial standard and the quality assurance of the manufacturing process. The FDA-cGMP regulations emphasize environmental factors to minimize cross-contamination of products and errors in labeling and packaging, and the integrity of production and quality control records; however, they do little to minimize within-batch and batch-to-batch variation in the output of production. Therefore, it is an important function of the in-process quality assurance program to ensure that finished dosage forms have uniform purity and quality within a batch and between batches. This is accomplished by identifying critical steps in the manufacturing process and controlling them within defined limits.

## *Quality Assurance Before Start-up*

## *Environmental and Microbiologic Control and Sanitation*

To assure that finished dosage forms meet high standards of quality and purity, an effective sanitation program is required at all facilities where such products are manufactured. A successful extermination program must be enforced within and outside the plant to control insects and rodents. People are the mainstay of any plant housekeeping and sanitation program. Consequently, personal cleanliness and proper haircovering and clothing should be required. Floors, walls, and ceilings should be resistant to external forces, capable of being easily cleaned, and in good repair. Adequate ventilation, proper temperature and proper humidity are other important factors. Ventilation in manufacturing departments is usually designed so that dust can be contained and removed. In such departmental operations, dust collectors, air filters, and scrubbers to clean the air are checked on a routine schedule. Air quality monitoring at the work station could indicate the adequacy of these elements.

The water supply may be potable, distilled, or deionized, and must be under adequate pressure to keep the water flowing. Deionization units should be monitored, and the resins changed or regenerated frequently, to deliver water of consistently high chemical and microbial quality as per written compendial or inhouse specifications.

Quality assurance should review and monitor the following programs, based on written procedures that specify the details of each:

| | |
|---|---|
| Sanitation: | Example is shown in Table 29.6 for one insecticide. |
| Cleaning: | Building and equipment. |
| Ventilation: | Filter conditions and changes; pressure gauge; humidity monitoring; temperature monitoring; microbial monitoring (Table 29.7); light intensity. |
| Water: | Release at point of use after checking by quality control; proper flushing period and/or volume before water use. |

**Table 29.6:** Quality assurance operating procedure

| | | |
|---|---|---|
| Page Date | No. Supersedes New | Sanitation Control—Pest Control |
| *Certox:* Insecticide | | |

*Type of action*

Kills on contact

| *Formula* | *Approximate %* |
|---|---|
| Petroleum distillates | 71.8% |
| Technical piperonyl butoxide* | 12.0 |
| Pyrethrine | 1.2 |
| Inert ingredients | 15.0 |

*Dilution*

Dilute 1 gallon of concentrate with 4 gallons of water

*Time interval*

To be used once weekly after working hours on Friday evenings

*Area designation*

Floor and drains

*Equipment*

Spray unit for Certox

Certox concentrate

Safety equipment

*Removal of waste materials*

Removal of waste materials remaining in the spray units after exterminating shall be the responsibility of the exterminator

*Effectiveness inspection*

It will be the responsibility of the quality assurance department to perform routine area checks to ascertain the effectiveness of the frequency of spraying

It will be the responsibility of the area supervisor, however, to take necessary action immediately upon seeing any infestation

*Special restrictions and cautions*

1. Foods should be removed or covered during treatment
2. Do not store or use near heat or open flame

3. Apply only as designated on area designation assignments

*Toxicity in humans*

Severe allergic dermatitis and systemic allergic reactions are possible

*Toxic symptoms*

Large amounts may cause nausea, vomiting, tinnitus, headache, and other CNS disturbances.

*Government status*

EPA Registration Number: 1748–110

Since Certox presents no significant toxicity problem, no tolerance data are available

* equivalent to 9.6% (butylcarbityl) (6-propyl-piperonyl) ether and 2.4% related compounds.

Written by Checked by

**Table 29.7:** Environmental control in manufacture of oral dosage forms

| | | |
|---|---|---|
| Product _____ | Lot. No. _____ | |
| Room _____ | Date exposed _____ | |
| Media _____ Time of exposure _____ | Incubation temperature ____°C | Date read _____ |
| | Duration _____ | |
| 1—Location of plate exposure | Plate No. | Colony count |
| | | |
| 2—Location of air sampler ($m^3$ air/hr) | Plate No. | Colony count |
| Comments | | |
| Microbiologist _____ | Supervisor _____ | |
| Date reported _____ | | |

## MANUFACTURING WORKING FORMULA PROCEDURES (MWFPs)

Documentation of the component materials and processing steps, together with production operation specifications and equipment to be used, make up the MWFPs.

A working formula procedure should be prepared for each batch size that is produced. To attempt expansion or reduction of a batch size by manual calculations at the time of production cannot be considered good manufacturing practice.

Quality assurance personnel must review and check the working formula procedures for each production batch before, during, and after production for the following details:

- Signature and date of issue given by a responsible production or quality assurance employee.
- Proper identification by name and dosage form, item number, lot number, effective date of document, reference to a superseded version (if any), amount, lot, and code numbers of each raw material utilized.
- Initializing of each step by two of the operators involved.
- Calculations of both active and inactive materials, especially if there have been any corrections for 100% potencies for active ingredients used.
- Starting and finishing times of each operation.
- Equipment to be used and specification of its set-up.
- Proper labeling of released components and equipment, indicating product name, strength, lot number, and item number.

## Raw Materials

Quality assurance should check the original containers of released raw materials for cleanliness before they are taken to the production department. Most raw materials, however, are weighed in an environmental control weighing area, where they are transferred to a secondary container that circulates only inside the production department. This secondary container should be properly labeled with a sticker that bears all the information on the original container label. Only released raw materials with proper reassay dates should be allowed to enter the production department. Raw materials intended for use in specific products should be stacked and stored together in an approximate staging area with proper identification (name, dosage form, item number, lot number, weight, and signatures).

## Manufacturing Equipment

Quality assurance personnel must ensure that manufacturing equipment is designed, located, and maintained so that it facilitates thorough cleaning, is suitable for its intended use, and minimizes potential for contamination during manufacture. Manufacturing equipment and utensils should be thoroughly cleansed and maintained in accordance with specific written directions. Whenever possible, equipment should be disassembled and thoroughly cleaned to preclude the carryover of drug residues from previous operations. Adequate records of such procedures and tests, if appropriate, should be monitored by quality assurance employees. It is good manufacturing practice to use laboratory checks whenever possible to detect trace quantities of drugs if products containing such drugs were produced on specific equipment prior to cleaning.

Prior to the start of any production operation, the quality assurance personnel should ascertain that the proper equipment and tooling for each manufacturing stage are being used. Equipment must be identified by labels bearing the name, dosage form, item number, and lot number of the product to be processed. Equipment used for special batch production should be completely separated in the production department, and all dust-producing operations should be provided with adequate exhaust systems to prevent cross-contamination and recirculation of contaminated air.

Weighing and measuring equipment used in production and quality assurance processes, such as thermometer and balances, should be calibrated and checked at suitable intervals by appropriate methods; records of such tests should be maintained by quality assurance and production personnel. Examples for such calibration methods are given in Table 29.8.

**Table 29.8:** Quality assurance calibration procedure

| Page Date | No. Supersedes | Calibration of thermometers |
|---|---|---|

Thermometers (to be checked every 6 months)

1. Employ suitable USP melting point standards for the range of the thermometer to be tested
2. Use USP method class I to determine the actual melting range of the standards

3. Tag the thermometer with calibration date, next calibration date, temperature correction, if any, and signature of the person conducting the test

Written by Checked by

## *Quality Assurance at Start-up*

### *Raw Materials Processing*

Only released, properly labeled raw materials are allowed in the in-processing area. Depending on the nature of the product, quality assurance personnel should check and verify that the temperature and humidity in the area are within the specified limits required for the product. If the temperature and/or humidity is beyond the specified limits, production must be informed and corrective actions taken.

The specified in-process procedure is to be checked, at each step in the process, according to written in-process quality assurance procedures.

Quality assurance personnel should verify and document the proper equipment, addition of ingredient, mixing time, drying time, filtering, and mesh size of sieves used in screening.

At certain points, samples are to be taken to the quality control laboratory for potency assay and any other testing that is necessary to ensure batch uniformity and purity.

Containers of in-process raw materials are labeled with product name, item number, lot number, and gross, tare, and net weights of the contents.

### *Compounding*

Quality assurance personnel are responsible for ascertaining that all containers of raw materials are properly labeled and staged in the compounding staging area, that they are clean, and that manufacturing equipment is properly identified as to product, strength, item number, and lot number.

The production process begins with the setup of the manufacturing equipment to prepare the finished dosage form within the specified limits for

the particular product. At each significant step in the procedure, quality assurance personnel verify that the process is being performed in accordance with the written directions and is conforming to required standards.

A variable group of tests that are widely used for in-process controls measure characteristics including physical appearance, color, odor, thickness, diameter, friability, hardness, weight variation, disintegration time, volume check, viscosity and pH. Such in-process tests are designed to ensure control of problems that can arise during finished dosage form manufacturing.

Current good manufacturing practices require that in-process quality assurance be adequately documented throughout all stages of manufacturing. Throughout the production run, in-process samples are removed and tested, and data are recorded on special forms as specified in the product's in-process monograph. The number of samples taken for testing and the type of testing obviously depend on the size of the batch and the type of product. If deviation from the specified limits occurs, the necessary corrective action is taken and recorded, and a resample is taken and tested to determine whether the quality attribute of the product is now within the limits. In some instances, as in the case of compendial weight variation or pH specifications, the deviation is such that units produced prior to the corrective action are isolated, accounted for, and rejected.

In addition to the foregoing, portions of the initial, final, and in-process samples are used for collecting average run samples for the quality control laboratory to perform final batch analysis and release.

## *Packaging Materials Control*

The USP defines the container closure system as the device that holds the drug and/or may be in direct contact with the drug. The immediate container is that which is in direct contact with the drug at all times. The closure is a part of the container.

Packaging materials should not interact physically or chemically with the finished product to alter the strength, quality, or purity beyond specified requirements. The compendium provides specifications and test procedures for light resistance: well-closed, tightly closed, and four different types of glass containers.

Specifications and test methods are designed for containers on the basis

of tests performed on the product in the container. The following features are to be considered in developing container specifications:

- Properties of container tightness.
- Moisture and vapor tightness regardless of container construction.
- Toxicity and chemical/physical characteristics of materials needed in container construction.
- Physical or chemical changes of container upon prolonged contact with product.
- Compatibility between container and product.

Good manufacturing practices require that stability data be submitted for the finished dosage form of the drug in the container closure system in which it is to be marketed.

## *Labels Control*

Production control issues a packaging form that carries the name of the product; item number; lot number; number of labels, inserts, and packaging materials to be used; operations to be performed, and the quantity to be packaged. A copy of this form is sent to the supervisor of label control, who in turn counts out the required number of labels. Since labels may be spoiled during the packaging operation, a definite number in excess of that actually required is usually issued; however, all labels must be accounted for at the end of the operation, and unused labels must be accounted for before their destruction. If the lot number and expiration date of the product are not going to be printed directly on the line, the labels are run through a printing machine, which imprints the lot number and expiration date. The labels are recounted and placed in a separate container with proper identification for future transfer to the packaging department. The packaging department then requests, according to the packaging form, the product to be packaged, along with all packaging components, such as labels, inserts, bottles, vials, ampoules, stoppers, caps, seals, cartons, and shipping cases. Quality assurance personnel inspects and verifies all packaging components and equipment to be used for the packaging operation to ensure that it has the proper identification, that the line has been thoroughly cleaned, and that all materials from the previous packaging operation have been completely removed. Proper reconciliation and disposition of the unused and wasted

labels should occur at the end of the packaging operation.

## *Finished Product Control*

**Specifications:** Final testing of finished product is made in the quality control laboratories. These tests are designed to determine compliance with specifications. Thus, the testing of the finished product for compliance with predetermined standards prior to release of the product for packaging and subsequent distribution is a critical factor for quality assurance. This testing, along with in-process testing, assures that each unit contains the amount of drug claimed on the label, that all of the drug in each unit is available for complete absorption, that the drug is stable in the formulation in its specific final container closure system for its expected shelf-life, and that dosage units themselves contain no toxic foreign substances.

Normally, the design of test parameters, procedures, and specifications is made during product development. It is a good manufacturing practice to base such parameters on experiences developed from several pilot and production batches. Furthermore, the results of these studies should be subjected to statistical analysis where appropriate, to appraise the precision and accuracy of each procedure correctly for each characteristic. In the long run, with additional production experience, specifications may possibly be modified to upgrade product specifications.

**Bulk product testing:** Each lot of bulk product should be tested to ensure identity, quality, potency, and purity. Quality assurance authorizes the release for further processing based on actual physical, chemical, and/or biologic laboratory testing.

Tests required by the official compendia on the ingredients and the dosage form applies to all manufacturers of a specific compendial product. The manufacturer frequently employs alternative methods that are more accurate, specific, or economical than those in the compendia.

The manufacturer is not required to employ the official analytic procedures as long as the quality of his product complies with the compendial requirements. In the case of a legal action, however, the compendial procedures become the referee procedures for determining compliance.

**Quality assurance during packaging operation:** If the quality control laboratory analysis confirms that the product complies with specifications,

and if the quality assurance audit indicates that manufacturing operations are satisfactory, the bulk product is released to the packaging department, and production control is notified. Packaging operations should be performed with adequate physical segregation from product to product. Quality assurance personnel should periodically inspect the packaging lines and should check filled and labeled containers for compliance with written specifications. Some packaging operations, especially those using high-speed equipment, are fitted with automated testing equipment to check each container for fill and label placement. Alternatively, an operator may visually inspect all packages fed into the final cartons. Quality assurance should perform an independent inspection and select finished preservation samples at random from each lot. The preservation samples should consist of at least twice the quantity necessary to perform all tests required to determine whether the product meets its established specifications. These preservation samples should be retained for at least 1 year after the expiration date and should be stored their original package under conditions consistent with product labeling.

Quality assurance personnel should also select an appropriate size sample of the finished package product and send it to the analytic control laboratory for final testing.

**Auditing:** Good manufacturing practice requires that the manufacturing process be adequately documented throughout all stages of the operation. The history of each task, including the starting materials, equipment used, personnel involved in production and control until completed packaging is complete, should be recorded. The areas of record keeping include:

- Individual components, raw materials, and packaging materials.
- Master formula and procedure.
- Batch production.
- Packaging and labeling operation.
- Laboratory in-process and finished control testing.
- Proper signing and dating by at least two individuals independently for each operation in the proper spaces.
- Reconciliation of materials supplied with amounts of tablets produced, taking into account allowable loss limits.

Before releasing the product for distribution, quality assurance should evaluate the batch records of all in-process tests and controls and of all tests of the final product to determine whether they conform to specifications.

## Concept of Statistical Quality Control

*Statistical quality control* (SQC) has been defined as the monitoring of quality by application of statistical methods in all stages of production. Statistical methods of investigation based on the theory of probability may be used for estimating parameters, for performing tests of significance, for determining the relationship between factors, and for making meaningful decisions on the basis of experimental evidence. The selection of an appropriate method essentially depends on the type of measurement, the sampling techniques, the design of the experiment, and the type of sample distribution (normal, binomial, or Poisson).

SQC has been used to serve (1) as a basis for improved evaluation of materials through more representative sampling techniques and (2) as a means of achieving sharper control in certain manufacturing processes. The procedures consist of properly sampling the product, determining the quality variation of the sample, and from this data, making inferences to the entire batch under investigation. Once the characteristic data pattern of a process has been determined, the pattern can be utilized to predict the limits within which future data can be expected to fall as a matter of chance, and to determine when significant variations in the process have taken place. The objective is to determine whether the major source of observed variation is due to “chance variation,” which is inevitable during the manufacturing process, or to “assignable causes,” which can usually be detected and corrected by appropriate methods.

Any program of production and inspection has its own unique chance causes of variation, which cannot be controlled or eliminated and often cannot be identified. This variation represents a pool of small individual variations for which limits can be established. On the other hand, the assignable causes can usually be identified by statistical techniques, for example, the detection of an outlier or a trend or pattern. Assignable causes, by definition, are associated with something special and assignable, such as excessive variations caused by a specific machine, a specific batch of material, or a container. Therefore, the use of SQC permits the evaluation of the magnitude of “chance variation” and the detection of “assignable variation” of product quality. This may be detected by means of quality control charts.

## *Normal Frequency Distribution*

There are many types of variation patterns in product quality. The most common pattern of data distribution is the normal curve, a symmetric, bell-shaped curve. By plotting the relative frequency of the data obtained from a large number of results along the vertical axis against the magnitude of the measured characteristic, such as tablet weight or chemical assay, on the horizontal axis, a normal frequency distribution curve is often obtained, as shown in Fig. 29.4. Most statistical techniques are based on the assumption of a normal frequency distribution curve.

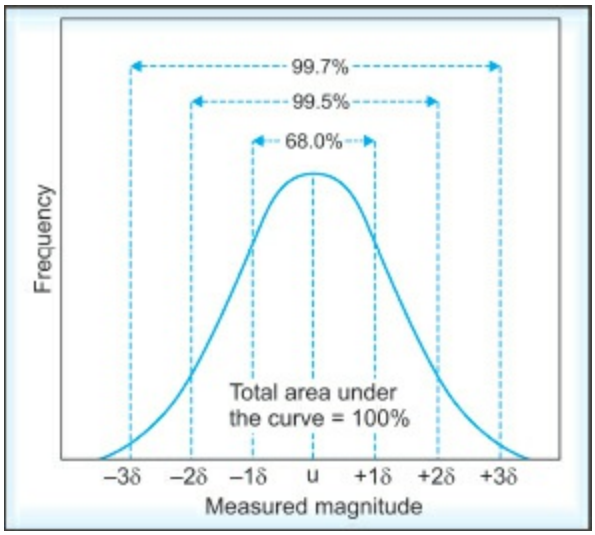



**Fig. 29.4:** A normal frequency distribution curve characterized by the mean (μ) and the standard deviation (δ)

The normal distribution is defined completely by two parameters: the mean and the standard deviation. The observed mean (X) is the arithmetic average of a series of values and is calculated by dividing the sum of such values by the number of values (N) in the series. It is expressed as:

$$\overline{X} = \frac{\Sigma X}{N} \qquad \ldots (1)$$

$\overline{X}$ is the best measure of the central value of a normal distribution and is an estimate of the theoretic mean ($\mu$). For quantitative expression of the dispersion or scatter about the central value for establishing an estimate of variation among the values, the range (R) and the standard deviation ($\delta$) are commonly employed. The range is calculated as the arithmetic difference between the largest and the smallest value in a group of data. The standard

deviation of the sample (s) is an unbiased estimate of the standard deviation of the population ($\delta$) and is calculated from the following formula:

$$s = \sqrt{\frac{\Sigma (\bar{X} - X)^2}{N - 1}} \qquad \ldots (2)$$

The standard deviation is a measure of the variation expressed in the same units of measurement as the original values. Furthermore, if the distribution of measurements is normal, it permits delineation of a zone or range within which a given portion of the original observations normally lie. Thus, as seen in Fig. 29.4, about 68% of all results fall within one standard deviation (± 1 s) on either side of the mean, 95.5% within two standard deviations (± 2 s), and 99.7% within three standard deviations (± 3 s).

When sampling from a normal distribution, $\bar{X}$ and s are independent of each other, i.e. the dispersion or scatter as measured by s is independent of the magnitude of the $\bar{X}$. Figure 29.5 shows a series of normal probability distributions whose mean values are the same, but whose standard deviations differ. Such distributions are so constructed that the total area under each curve is unity, or 100%; thus, segments of area under the curve represent probabilities.

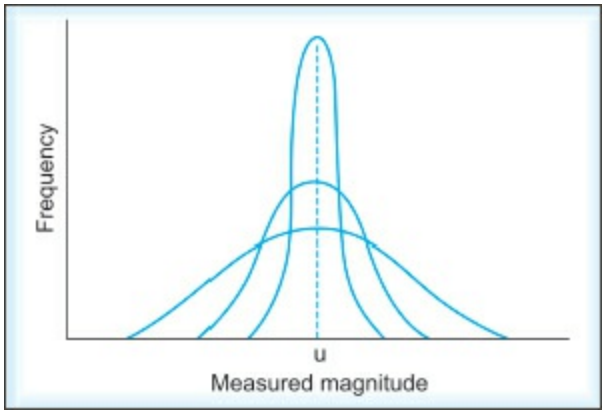



**Fig. 29.5:** Frequency distributions with the same mean but different standard deviation

Both the range (R) and standard deviation(s) are measures of the variability among individuals in a group. Although R is a less efficient measure of that variability, there is a definite relationship between the two values. In fact, it is possible to obtain an estimate of s from R by dividing R by an appropriate divisor, the numerical value of which depends on the

sample size (N). Divisors (D) for selected N are presented in Table 29.9. If it is known that the mean sample range calculated from a series of samples, each containing N = 4 items, has the value 6.3, then the standard deviation for the distribution of individual items can be estimated as s = 6.3/2.1 = 3.0. If s for the distribution of individual items in a certain population is 4.0, the mean sample range in samples of N = 7 could also be estimated from Table 29.9 as R = 2.7 × 4.0 = 10.8. Estimations of the standard deviation from the range are used quite extensively in industrial statistics.

**Table 29.9:** Divisors for estimating standard deviation from the range

| Sample size (N) | Divisor (D) |
|---|---|
| 2 | 1.1 |
| 3 | 1.7 |
| 4 | 2.1 |
| 5 | 2.3 |
| 6 | 2.5 |
| 7 | 2.7 |
| 8 | 2.8 |
| 9 | 3.0 |
| 10 | 3.1 |
| 15 | 3.5 |
| 30 | 4.0 |
| 50 | 4.5 |
| 100 | 5.0 |
| 500 | 6.0 |

## *Quality Control Charts*

Control charts have been employed for various pharmaceutical operations and may be used as aids in controlling and analyzing physical, chemical, analytic, or biologic parameters of a product, such as (1) weight variation of tablets and capsules, (2) thickness of tablets, (3) volume of liquid filling in a container, (4) the number or percentage of defects in parenteral products and

(5) the number or fraction of defects in a sample of packages emanating from a packaging operation. Any measurements that could form the basis of acceptance or rejection of the product would be amenable to surveillance via the control chart.

Control charts are useful in highlighting trends for intra- and inter-batch variation by following a moving mean value of a specification, as for tablet hardness or assay. There are two basic types of quality control charts: charts based on variable and charts based on attributes. Attribute charts refer to go or no-go situations in which each sample inspected is tested to determine whether it conforms to the requirements; variable charts are based on a continuous distribution of measurements that can, in a sense, measure degrees of unacceptability. A quality control chart by variables or by attributes developed on the basis of certain quality characteristics serves as an aid in keeping the product in control.

The application of control charts to manufacturing processes can best be described by means of the following examples:

## *Control Charts by Variables*

**Example 1:** During the automatic filling of a parenteral solution in vials, control of the volume filled during a production run should be established and maintained. One vial was taken at random from each of the four needles of the filling machine at designated times and the average and range of this subgroup of four was computed. The numerical data obtained for the process control record is summarized in Table 29.10 and graphically depicted in Fig. 29.6. The upper curves of Fig. 29.6 and $X$ (mean) charts and the lower curves are R (range) charts. The lines connecting the points serve only to aid in visualizing the trend. The upper and lower control limits for the average (U$CL_x$ and $LCL_x$) of the filled volume were calculated as shown on the lower part of Table 29.10, using the formulas:

$$\text{U}CL_x = \bar{X} + A_2\,\bar{R} \ldots (3)$$

$$LCI_x = \bar{X} - A_2\,\bar{R} \ldots (4)$$

**Table 29.10:** Process control record of automatic filling of a parenteral solution in vials from four filling needles of machine (Label fill is 10 ml; required fill is 10.5 ml)

| Time the vial is sampled | Volume of solution filled (ml) | | | | | |
|---|---|---|---|---|---|---|
| | Needle A | Needle B | Needle C | Needle D | $\bar{X}$ | R |
| Day 1 | | | | | | |
| 8:30 am | 10.7 | 10.5 | 10.6 | 10.5 | 10.58 | 0.2 |
| 9:20 | 10.5 | 10.5 | 10.8 | 10.7 | 10.63 | 0.3 |
| 10:15 | 10.5 | 10.9 | 10.5 | 10.5 | 10.60 | 0.4 |
| 11:25 | 10.8 | 10.5 | 10.5 | 10.5 | 10.58 | 0.3 |
| 1:00 pm | 10.5 | 10.7 | 10.5 | 10.5 | 10.55 | 0.2 |
| 2:00 | 10.5 | 10.8 | 10.5 | 10.5 | 10.58 | 0.3 |
| 3:10 | 10.6 | 10.5 | 10.9 | 10.6 | 10.65 | 0.4 |
| 4:00 | 10.5 | 10.5 | 10.7 | 10.7 | 10.60 | 0.2 |
| Day 2 | | | | | | |
| 8:20 am | 10.5 | 10.7 | 10.7 | 10.5 | 10.60 | 0.2 |
| 9:30 | 10.6 | 10.5 | 10.8 | 10.5 | 10.60 | 0.3 |
| 10:15 | 10.7 | 10.8 | 10.8 | 10.7 | 10.75 | 0.1 |
| 11:30 | 10.5 | 10.6 | 10.6 | 10.6 | 10.58 | 0.1 |
| 1:10 pm | 10.4 | 10.5 | 10.5 | 10.8 | 10.55 | 0.4 |
| 2:20 | 10.5 | 10.6 | 10.4 | 10.9 | 10.60 | 0.5 |
| 3:30 | 10.6 | 10.6 | 10.5 | 10.8 | 10.63 | 0.3 |

*Calculation*

$\bar{\bar{X}}$ = Grand average = 10.60

$\bar{R}$ = Average range = 0.28

$UCL_{\bar{X}} = \bar{\bar{X}} + A_2\bar{R} = 10.60 + (0.73)(0.28) = 10.80$

$LCL_{\bar{X}} = \bar{\bar{X}} - A_2\bar{R} = 10.60 - (0.73)(0.28) = 10.80$

$UCL_R = D_4\bar{R} = (2.88)(0.28) = 0.64$

$LCL_R = D_3\bar{R} = (0)(0.28) = 0$

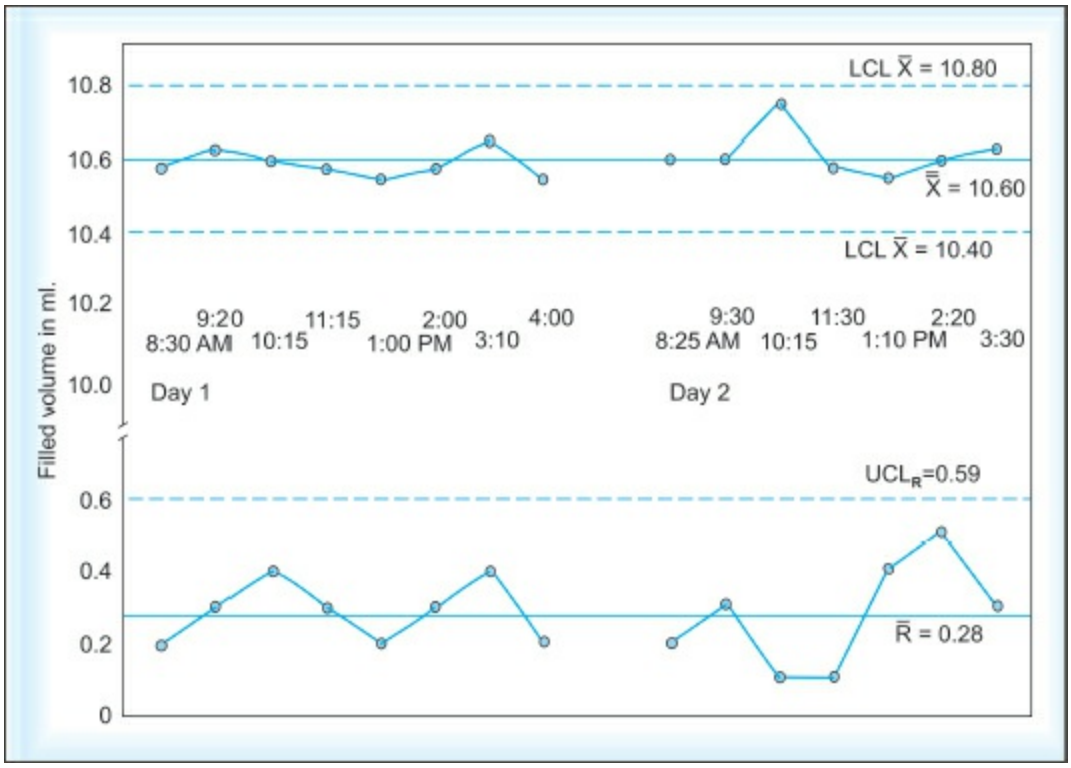



**Fig. 29.6:** Control chart for automatic filling of a parenteral solution in vials. The upper curves are x (mean) charts and the lower curves are R (range) charts.

where $\bar{X}$, the grand average, is the aritiimetic average of all the $\bar{X}$; $\bar{R}$, the average range, is calculated as the arithmetic average of the values of R; and A2, the factor for using the range to calculate three standard deviation limits for the average, is a constant that for subgroups of four has the value of 0.73. $A_2$ for selected sample sizes (N) are presented in the last column of Table 29.11.

The upper and lower control limits for the range ($UCL_R$ and $LCL_R$) were calculated from the formulas:

$$UCL_R = D_4 \cdot \bar{R} \ldots (5)$$

$$LCL_R = D_3 \cdot \bar{R} \ldots (6)$$

where $D_3$ and $D_4$, the factors to calculate three standard deviation limits for the range, are constants, which for subgroups of four in this case have the values of 0.00 and 2.28, respectively. $D_3$ and $D_4$ for selected sample sizes (N) are tabulated in Table 29.11.

**Table 29.11:** Factors for estimating the three standard deviation limits

| Sample Size (N) | Factors for $\bar{R}$ Chart | | Factors for X Chart |
|---|---|---|---|
| | $D_3$ | $D_4$ | $A_2$ |
| 2 | 0.00 | 3.27 | 1.88 |
| 3 | 0.00 | 2.57 | 1.02 |
| 4 | 0.00 | 2.28 | 0.73 |
| 5 | 0.00 | 2.11 | 0.58 |
| 6 | 0.00 | 2.00 | 0.48 |
| 7 | 0.08 | 1.92 | 0.42 |
| 8 | 0.14 | 1.86 | 0.37 |
| 9 | 0.18 | 1.82 | 0.34 |
| 10 | 0.22 | 1.78 | 0.31 |

Control charts are depicted in Fig. 29.6 are characterized by a vertical axis that has a scale of varying measurement, such as a mean (X), range (R), standard deviation (s), or fraction unacceptable (p), and a horizontal axis that is time-oriented. Each point in the control chart represents a value obtained from a selected number of items referred to as a subgroup. These values are plotted in sequence. The control charts as shown in Fig. 29.6 consist of a solid and two horizontally parallel lines on either side of the solid line. The central solid line is the target value or the historical process average and/or range. The two dotted parallel lines indicate the limits within which practically all the observed results should fall as long as the process is under normal variation (statistically controlled). The upper dotted line is the upper control limit (UCL), which is normally three standard deviations above the center line. Likewise, the lower line is the lower control limit (LCL) and is three standard deviations below the center line.

If the process is in control, the six standard deviations spread between the upper and lower control limits encompass 99.7% of the values in a normal distribution with its mean at the center line. This is the case for the example shown in Fig. 29.6. The control chart of Fig. 29.6 is said to show evidence of "control," since all points fall within the designated control limits.

The control chart for averages ($\bar{X}$ chart) is a measurement of the central tendency. Approximately half of the values are above the average while the other half are located below it. This would be expected if random variation is present and the process is under control. Control charts that are used for plotting sample averages are frequently used in conjunction with a range or standard deviation chart. The control chart for ranges (R chart) indicates the

variation present in a set of samples. It is advantageous to plot both the $\bar{X}$ and R charts together, because one may be in control and the other may show excessive variation.

**Example 2:** The performance of the filling machines for filling by weight a relatively viscous parenteral suspension in multiple dose containers can be followed with control charts. Fig. 29.7 shows a control chart on the performance of two machines in filling a batch of the material. Two samples were removed from the production flow and weighed approximately every 20 min, and the average and range measurements were used to construct the control chart in Fig. 29.7. It can be seen that Machine A is filling vials within the predicted limits, the averages being within the control limits. Machine B shows large variation, however, as can be seen in the range chart as well as by points indicated outside the control limits on the average chart. Since the measured values fall outside the limits, causes for the process being out of control, which may be assignable, must be investigated. Some possibilities of the assignable causes are filling machine connections, filters, filling heads, and automatic measuring equipment.

Assuming that the assignable causes for the out-of-control behavior in Machine B are found and corrected, it still may be desirable to reduce the variation as observed in the range charts of Machines A and B. This would allow the establishment of narrower control limits on the average chart, thus making it possible to reduce the overfill and still maintain the same degree of assurance of meeting the specifications. Upon intensive process investigation and improvement, it was possible to narrow the control limits on the average chart as shown in the control chart of Fig. 29.8. It is worth noting here that the control limits were narrowed on both X and R charts as compared to those of Fig. 29.7. Thus, the process was brought into closer statistical control, with the fill weight variation being minimized.

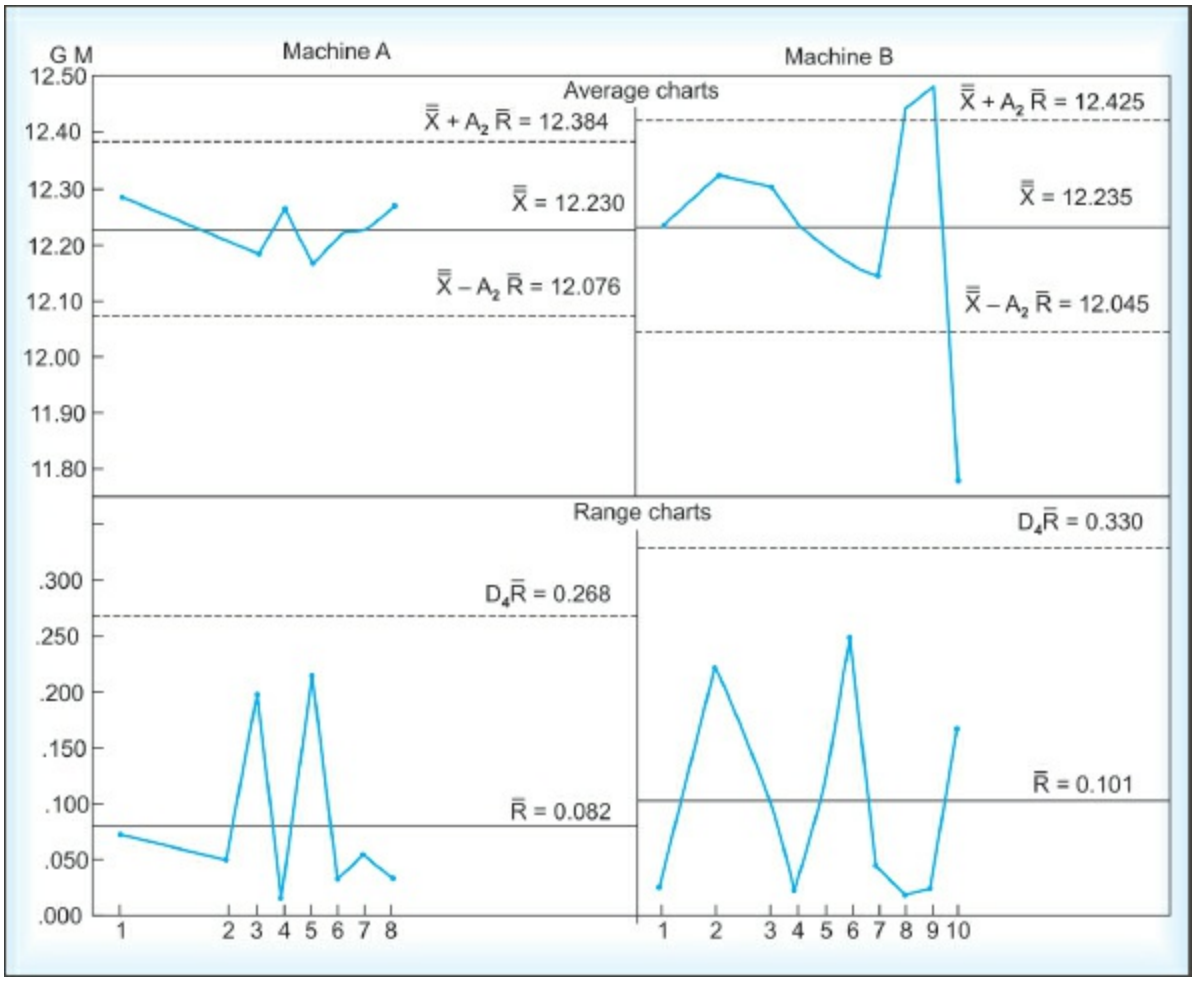



**Fig. 29.7:** Control chart containing the $\bar{X}$ and R charts on comparative performance of two filling machines for a parenteral suspension

## *Control Charts by Attributes*

As mentioned, when a record shows only the number of articles conforming and the number of articles failing to conform to any specified requirement (go or no-go) it is said to be a control record by attributes. It is obvious that most routine inspections of manufactured pharmaceutical products, such as the inspection of parenteral products or counting of broken tablets in a bottle, are inspections for attributes. The interest is in the number of flaws or fraction defective per batch. Thus, the weakness of attribute measurements is that gradation of quality cannot be measured. In general, variable charts are more sensitive than attribute charts, although the latter are usually easier to implement. To employ control charts for attributes, a plan with desirable properties would include the following:

1. Samples should be chosen at random.
2. For easier evaluation there should be a fixed number of samples (n) taken each time for inspection. Each tablet or ampoule inspected is considered a sample.
3. Each sample is evaluated so that the sample is accepted or rejected, i.e. each tablet or ampoule is either good or defective.
4. Each sample must be independent, i.e. each tablet or ampoule, good or defective, has nothing to do with another tablet or ampoule to cause it to be good or defective.

The essential steps in constructing the control charts for fraction defective are (1) recording the number inspected, n, and the number of defectives found, d, (2) computing the fraction defective, p, which is the ratio of the number of defectives found to the total number of units actually inspected in the batch, (3) computing the average fraction defective, P, obtained by dividing the total number of defectives found by the total number of units inspected in a series of batches:

$$\overline{P} = \frac{\Sigma d}{\Sigma n} \qquad \ldots (7)$$

(4) calculating the upper and lower control limits, UCL and LCL, through the following formulas:

$$UCL = \overline{P} + \frac{3\sqrt{\overline{P}\,(1-\overline{P})}}{\sqrt{n}} \qquad \ldots (8)$$

$$LCL = \overline{P} - \frac{3\sqrt{\overline{P}\,(1-\overline{P})}}{\sqrt{n}} \qquad \ldots (9)$$

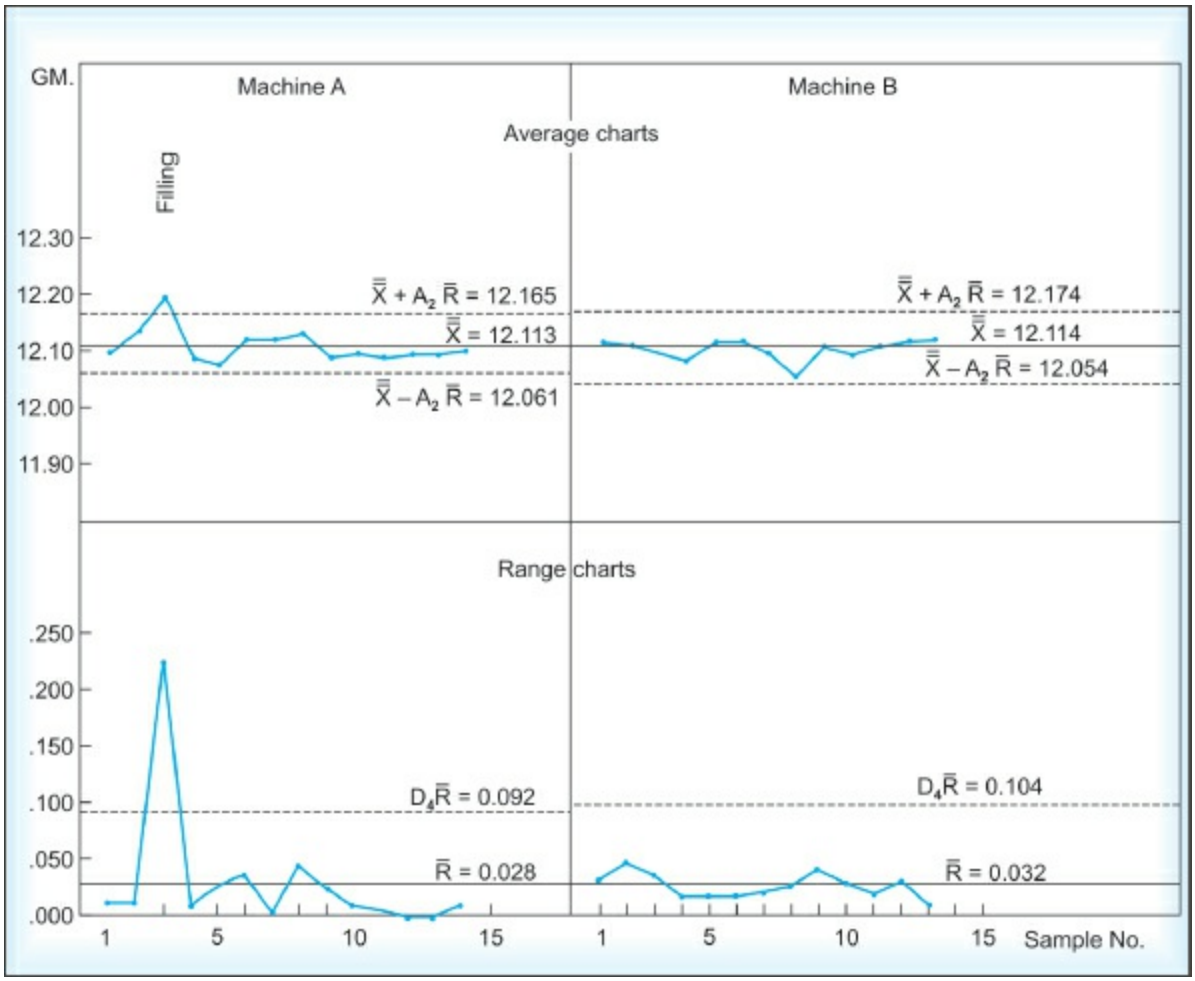



**Fig. 29.8:** Control chart containing the X and R charts on performance of two filling machines after process investigation and improvement. Note that the control limits were narrowed on × and R charts as compared to that of Figure 29.7

If n is different for different batches, upper and lower control limits will vary from batch to batch. To calculate the limits, first the value of $3\sqrt{\bar{P}(1-\bar{P})}$ is computed. The value of the square root of n is calculated for each batch and divided into $3\sqrt{\bar{P}(1-\bar{P})}$ to obtain the value of the three standard deviations for the batch, 3*s*, that is:

$$3s = \frac{3\sqrt{\bar{P}(1-\bar{P})}}{\sqrt{n}} \qquad \text{... (10)}$$

and (5) plotting of P's on a control chart with P as the average and control limits calculated as previously shown.

**Example 1:** The example used for illustrating the control chart (p chart) for percentage fractional defective for a parenteral product is shown in Fig. 29.9. The process control record accumulated from 10 batches of the product and the computation of p, P, 3s, UCL, and LCL employing equations (7) to (11) are summarized in Table 29.12 for constructing Fig. 29.9.

**Table 29.12:** Process control record and control limits calculated (for ten batches of a parenteral product)

| Batch # | Number inspected (n) | Number defective (d) | Fraction defective (P) | 3S | UCL | LCL |
|---|---|---|---|---|---|---|
| 1 | 4956 | 84 | 0.0170 | 0.0059 | 0.0254 | 0.0136 |
| 2 | 4924 | 71 | 0.0144 | 0.0059 | 0.0254 | 0.0136 |
| 3 | 4900 | 120 | 0.0245 | 0.0059 | 0.0254 | 0.0136 |
| 4 | 4883 | 98 | 0.0201 | 0.0059 | 0.0254 | 0.0136 |
| 5 | 4891 | 114 | 0.0233 | 0.0059 | 0.0254 | 0.0136 |
| 6 | 4952 | 88 | 0.0178 | 0.0059 | 0.0254 | 0.0136 |
| 7 | 4905 | 109 | 0.0222 | 0.0059 | 0.0254 | 0.0136 |
| 8 | 4897 | 72 | 0.0147 | 0.0059 | 0.0254 | 0.0136 |
| 9 | 4868 | 95 | 0.0197 | 0.0059[5] | 0.0255 | 0.0135 |
| 10 | 4845 | 103 | 0.0213 | 0.0059[6] | 0.0255 | 0.0135 |

*Calculation* : $\Sigma n = 49021$

$\Sigma d = 954$

$$p = \frac{84}{4956} = 0.0170 = 1.70\%$$

$$\overline{p} = \frac{954}{49021} = 0.0195 = 1.95\%$$

$$3S = \frac{3\sqrt{0.0195[1-0.0195]}}{\sqrt{4956}} = 0.0059 = 0.59\%$$

$$UCL = \overline{p} + 3S = 0.0195 + 0.0059 = 0.0254 = 2.54\%$$

$$LCL = \overline{p} - 3S = 0.0195 - 0.0059 = 0.0136 = 1.36\%$$

The inspectors are required to detect and remove the defective samples, which may be characterized by (1) broken ampoules or vials, (2) presence of particulate matters such as lint, dirt, glass fragment, and other foreign suspended matter, or (3) imperfections in the glass. The number of defectives found in the batch inspected was recorded, and the fraction defective of the batch was computed. Percentage defective (100 P), a more convenient value than fraction defective, has been plotted in Fig. 29.9. The numerical values of three standard deviations (3s) for each individual batch in this particular

example are considered to be constant, since n for each batch changes only slightly. The LCL may be set at zero if only high values of p are of concern. Because all points fall within the control limits, as seen in Fig. 29.9, the product is said to be in statistical control.

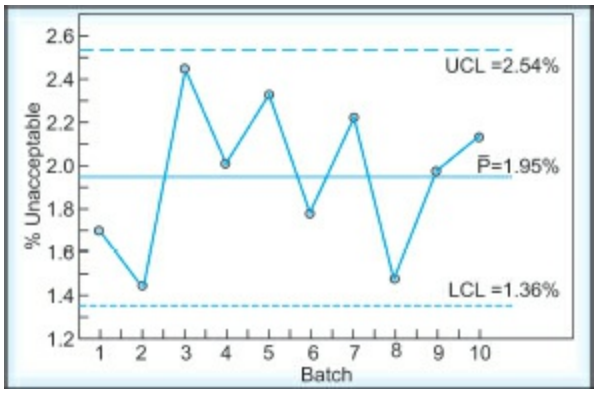



**Fig. 29.9:** Control chart (p chart) for percentage of defective units inspected for a parenteral product

**Example 2:** During the automatic packaging of a tablet product in containers, control of the number of broken tablets in a container released from packaging lines should be established and maintained. By sampling containers from each batch at random and inspecting and recording the percentage unacceptable for each batch, data may be accumulated. Occasionally, it may be observed that batches are out of control (P>UCL) during the accumulation of data, as shown in the p chart of Fig. 29.10. Batches 6, 13,16 and 18 are out of control. After all assignable causes have been corrected, these rejected batches should be eliminated, and a new set of P, UCL, and LCL values should be calculated for a new p chart.

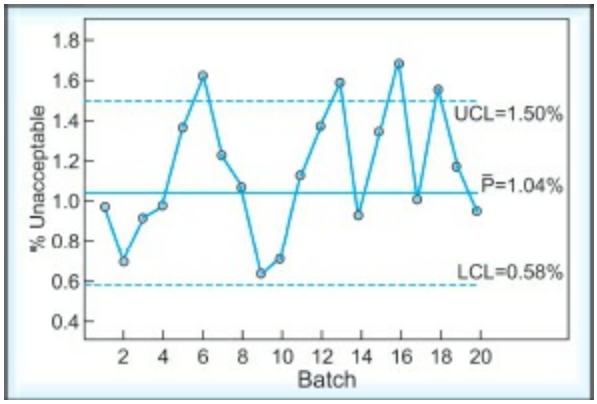



**Fig. 29.10:** Control chart (p chart) for percentage of broken tablets in plastic

containers

All methods and procedures set forth for quality control administration should be strictly followed, since changes may cause significant variation in the data being collected. When data are being recorded, any conditions that have changed since the last sample was taken should also be recorded. These include such items as changes in machine settings and operators. In general, the control chart is plotted initially without the benefit of control limits. When sufficient historical data from at least 10 to 12 batches have been collected, the control limits computed are reasonably reliable, and the control chart can be established for future batches. As the new batches of the product are produced, the inspection data are plotted on the control chart. Close attention to control charts indicates to manufacturing personnel and control inspectors whether the product is of poor quality.

## *Quality Level and Inherent Variability*

The ideal state of a statistically controlled process or situation is one in which both the quality level, as reflected by averages in chart, and the inherent variability, as indicated by ranges in R chart, remain within limits predicted from ordinary variation. This is illustrated in Fig. 29.11D, where the process is in the ideal state of statistical control. This situation is not always ideal in practice, however, and lack of control may be indicated in the following three different situations (1) the inherent variability remains essentially constant, but the quality level shifts from time to time as shown in Fig. 29.11A, (2) the level of quality may remain essentially constant, but the inherent variability changes from time to time, a situation illustrated in Fig. 29.11B and (3) both quality level and inherent variability shift together as depicted in Fig. 29.11C. Through the use of a control chart, basic variability of the quality parameter, average level of the quality characteristic, and the consistency of the performance of the product are easily visualized. Since the control chart provides a continuous monitoring of a process, it sounds a warning signal quickly when the property being measured falls outside the control limits. As a result, steps can be taken promptly to remedy any indication of trouble before future batches or the remainder of a current batch are manufactured. Thus, the control chart indicates when a process is out of control, but does not necessarily indicate the exact manner in which a shift in quality level or inherent variability occurs.

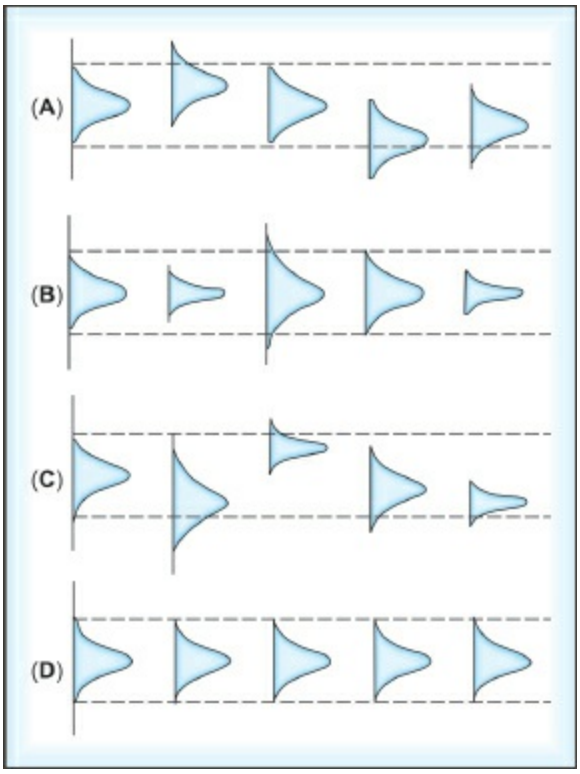



**Fig. 29.11:** Examples of product quality variation during manufacturing key: (A) Lack of control due to a shifting quality level; (B) Lack of control due to changes in inherent variability; (C) Lack of control due to changes in both quality level and inherent variability; (D) An ideal statistically controlled process

## Sampling and Sampling Plans

Sampling may be defined as *the process of removing an appropriate number of items from a population in order to make inferences to the entire population*. The object of sampling and subsequent testing, in the present context, is to provide an effective check on the quality of the product or substances being processed. Representative of materials to be sampled are drug, substances, raw materials, intermediate products, and final products before, during, and after manufacturing and packaging operations. The quality control inspector must be empowered to sample at any point or stage of manufacturing and packaging operations. Proper methods of sampling and adequate number and size of samples are needed for an effective quality assurance program, since the judgment “accept” or “reject” is made on the basis of the sample, irrespective of the conditions in the remainder of the batch.

Although controlled manufacturing and packaging systems provide the largest measure of quality assurance, the quality level of final dosage forms has to be tested and inspected. The degree of uniformity in content of any component in a dosage form is subject to an additive effect of the variabilities of the process steps. To obtain a value representative of the total population, the sample taken is most important. An improper sample may result in a value that can be biased and in error. Careful consideration of the design of the sampling plan enables these errors to be kept to a minimum. The only possible way to avoid such sampling errors is to do a 100% inspection. This normally cannot be attained practically and can also introduce errors due to personnel fatigue arid other related human factors. Even if it is possible to perform 100% inspection, good sampling plans are preferred for efficient inspection.

A good statistical sampling plan should be able to pass a high percentage of acceptable batches and reject the unacceptable ones. The number of unacceptable units are controlled to rigid standards by the stringency of the sampling plan. The variety of sampling plans, procedures, and tables that can be constructed is unlimited. The advantages and disadvantages of each of several possible choices should be carefully weighed from both theoretic and practical viewpoints. Frequently, the ease of implementation weighs more heavily than the statistical procedure because overall costs and ease of

application may be more important. Sometimes, trials of alternative procedures under actual operating conditions may bring to light unanticipated factors that result in the adoption of a plan that seems less efficient in theory. The choice of the most advantageous plan is determined after experiences have been accumulated.

Sampling plans based on data from measurements of attributes or variables may be constructed. As discussed previously, attribute data refer to go or no-go situations in which each piece inspected is examined or tested to determine whether it conforms to the requirements imposed by specifications; variable sampling is based on a continuous distribution of measurements, which can, in a sense, measure degrees of unacceptability covering the gray zone between accepted and rejected situations.

For practical purposes, the work involved in designing a sample plan may be greatly reduced or eliminated by use of a series of government-sponsored sampling plans such as MIL-STD-414 for variables sampling plans, and MIL-STD-105D for attribute sampling plans. In addition to providing a savings in time, these books have gained acceptance by industry throughout most of the United States. These publications provide sampling procedures and related reference tables for use in planning and conducting inspection. Acceptability of a batch is determined by the use of a sampling plan associated with the designated acceptable quality level or levels. A sampling plan indicates the number of units of product from each batch to be inspected (sample size or series of sample sizes) and the criteria for determining the acceptability of the batch. The inspection level determines the relationship between the batch size and the sample size and is to be determined by the responsible authority.

The most common and distinct methods of inspection are single and double sampling methods. In single sampling, only the specified sample size is inspected before a decision is reached regarding the disposition of the batch, and the acceptance criterion is expressed as an acceptance number. In double sampling, a second sample for inspection is permitted if the first fails, and two acceptance numbers are used-the first applying to the observed number of defectives for the first sample alone, and the second applying to the observed number of defectives for the first and second samples combined. Triple and multiple samplings are merely extensions of the foregoing.

The following example illustrates the double sampling method. A sample

of 50 tablets is taken. If it contains no more than two defectives, the batch is accepted. If it contains four or more defectives, it is rejected. If it contains more than two but less than four defectives, a second sample size of 50 is taken. If the two samples combined contain less than four defectives, the batch is accepted.

The construction of a statistical sampling plan normally requires that four basic quality standards be specified (1) an acceptable quality level (AQL) (e.g. a batch of tablets is considered to be accepted if it contains 2% or less unacceptable tablets), (2) an unacceptable quality level (UQL) (e.g. the same batch of tablets is said to be rejected if it contains 4% or more unacceptable tablets), (3) the risk or error, designated as $\alpha$ (Producers Risk), which is the probability of rejecting a good batch and (4) the risk or error, designated as $\beta$ (Consumers' Risk), which is the probability of accepting a bad batch. In general, the sampling scheme should be designed in such a way that the $\alpha$ and $\beta$ errors are appropriately shared by the producer and the consumer.

The usual approach is the determination of desirable AQL, UQL, $\alpha$ and $\beta$ and subsequent computation of the sample size and acceptance criteria. If the AQL and UQL are close together and $\alpha$ and $\beta$ are very small, as in the case of low-dose or potent drugs, a large sample is required for a suitable sampling plan. Conversely, the plan calls for few samples if the AQL and UQL are far apart and $\alpha$ and $\beta$ are large. A convenient graphic method of presenting these risks, $\alpha$ and $\beta$, is through the operating characteristic (OC) curves, which are graphs illustrating the ability of a sampling plan to discriminate between acceptable and unacceptable batches. For every sampling plan, there is an OC curve. The OC curve is prepared by plotting the percentage of batches of a given quality that is expected to be accepted versus the quality of submitted batches expressed as percentage unacceptable.

For characteristics that may be measured on a continuous scale and for which quality may be expressed in terms of percentage unacceptable, the MIL-STD-414 may be used as a source of statistical plans if the measurements are random and independent observations follow a normal distribution. To illustrate the application of MIL-STD-414, a hypothetical example for teaching purposes is employed. The tolerances of pure drug contained in the tablets in question were set at 93 to 107% of the labeled amount, and the desired quality characteristics of the tablets produced were specified at AQL = 10%, UQL = 40%, $\alpha$ = 10% and $\beta$ = 8%. By searching

through the OC curves of MIL-STD-414, Fig. 29.12A was found, which most nearly corresponds to these requirements. From this OC curve, one learns that batches of AQL = 10% will be accepted 90% ($\alpha$ = 10%) of the time, since the curve crosses the 10% abscissa at the 90% ordinate. By a similar operation, batches of UQL = 40% will be accepted only 8% (β) of the time. By referring to the appropriate table of MIL-STD-414, the sample size was found to be N = 10, and the acceptability criteria (or estimated percentage unacceptable) that balances the risks is 21.06%. Therefore, the final statistical plan takes the following form. Select 10 tablets at random from an inspection batch, and individually assay each tablet. From these ten determinations, estimate the percentage unacceptable in the batch. The batch will be accepted if this estimate does not exceed 21.06% unacceptable. Otherwise, the batch will be rejected.

Fig. 29.12B shows an OC curve, where one would pass that quality at least 95% of the time ($\alpha$ = 5%) if the submitted batch of material contains 10% or less unacceptable units (AQL = 10%). This same OC curve shows that a batch 20% or more unacceptable (UQL = 20%) would not pass more than 10% of the time (β = 10%). For this example, a sample size of 85 tablets is required. An ideal OC curve, illustrating perfect discrimination but unrealistic stringency, is shown in Fig. 29.12C. to obtain such a curve, the entire batch of products would have to be examined.

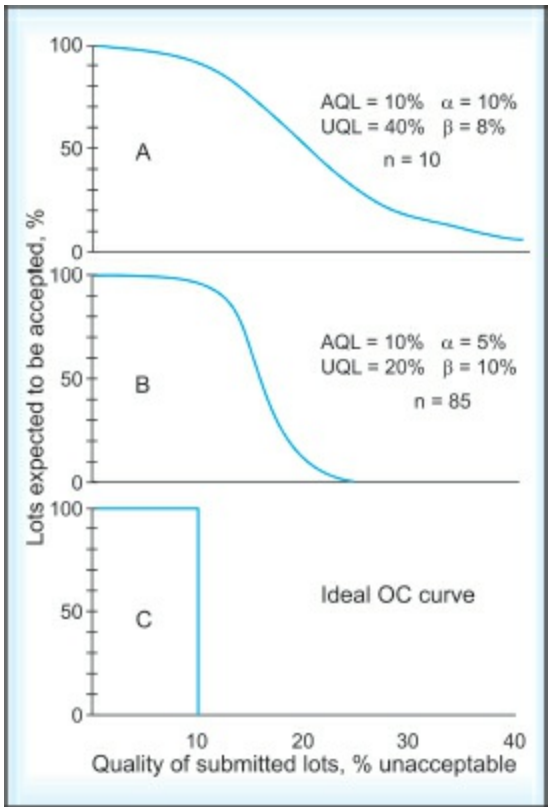



**Fig. 29.12:** Examples of operating characteristic curves

The five characteristics (AQL, UQL, $\alpha$, $\beta$ and N) operating within a variable sampling plan can be seen graphically from the operating characteristic curves for sampling plans at N = 10. As shown in Fig. 29.13, the horizontal scale running from 0 to 100% is the quality of submitted lots in the percentage defective, i.e. the percentage of the units in a lot that fall outside the established tolerances for the product. The vertical scale, which also runs from 0 to 100%, is the percentage of lots of a stated percentage unacceptable that are expected to be accepted under the sampling plan. For the OC curve of Fig. 29.13, labeled AQL = 4%, for example, it is shown that a lot that is 4% unacceptable will be accepted 90% of the time ($\alpha$ = 10%) at N = 10, and a lot that is 20% unacceptable (UQL = 20%) will be accepted 20% of the time ($\beta$ = 20%) at N = 10. At a given AQL, these plans are designed to greatly enlarge a probability of accepting a batch if the batch has AQL or a smaller percentage unacceptable. Figure 29.13 shows that when AQL of the product is raised, the percentage unacceptable that can be tolerated is also increased at the given level and sample size.

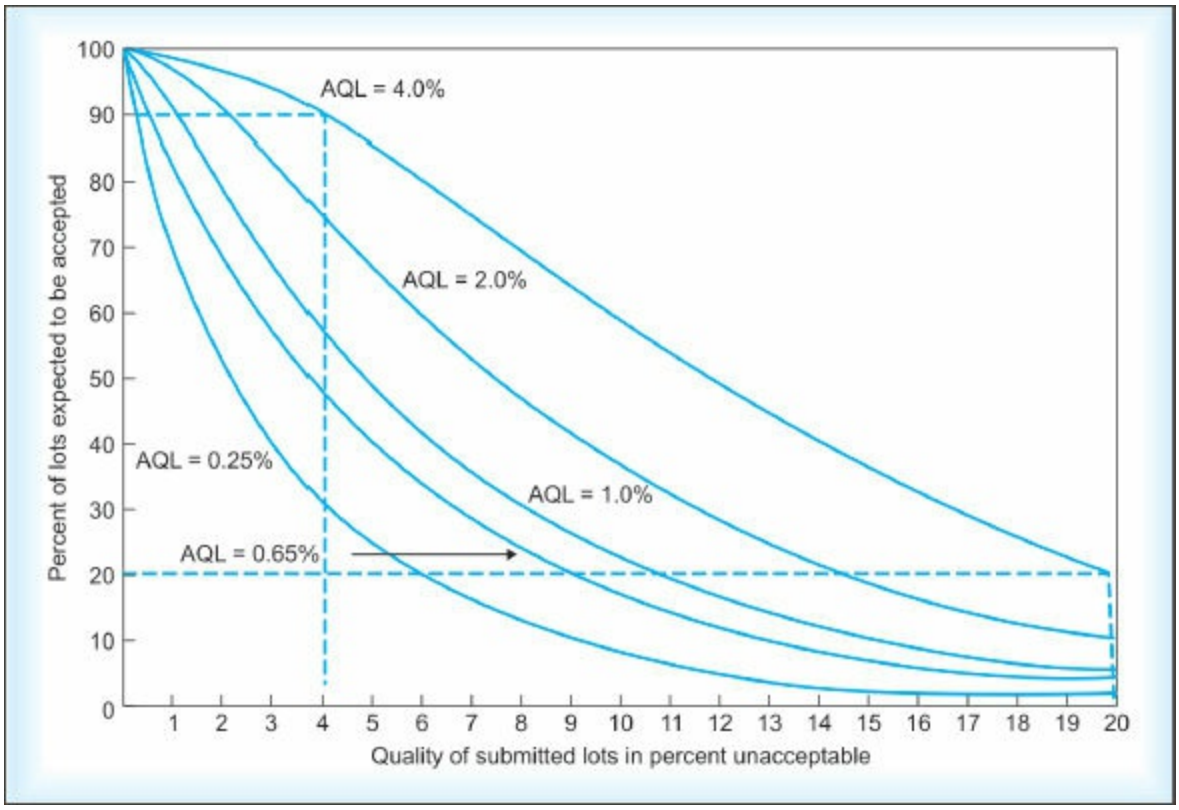



**Fig. 29.13:** Operating characteristic curves for sampling plans at N = 10

The operating characteristic curves for sampling plans at AQL = 1% are illustrated in Fig. 29.14. It should be noted that there are different curves for different sample sizes. For the OC curve labeled N = 10, it can be seen that a lot with 1% unacceptable (AQL =1%) will be accepted 90% of the time or conversely rejected 10% of the time ($\alpha$ = 10%). A lot that is 20% unacceptable (UQL = 20%) will be accepted 5% of the time ($\beta$ = 5%). As shown in Fig. 29.14, at a given level and AQL, the percentage unacceptable of the submitted lot that can be tolerated is decreased by increasing the sample size. It shows that plans at a given AQL are more discriminating with larger sample sizes; for example, it will accept good lots and reject bad lots a greater percentage of the time.

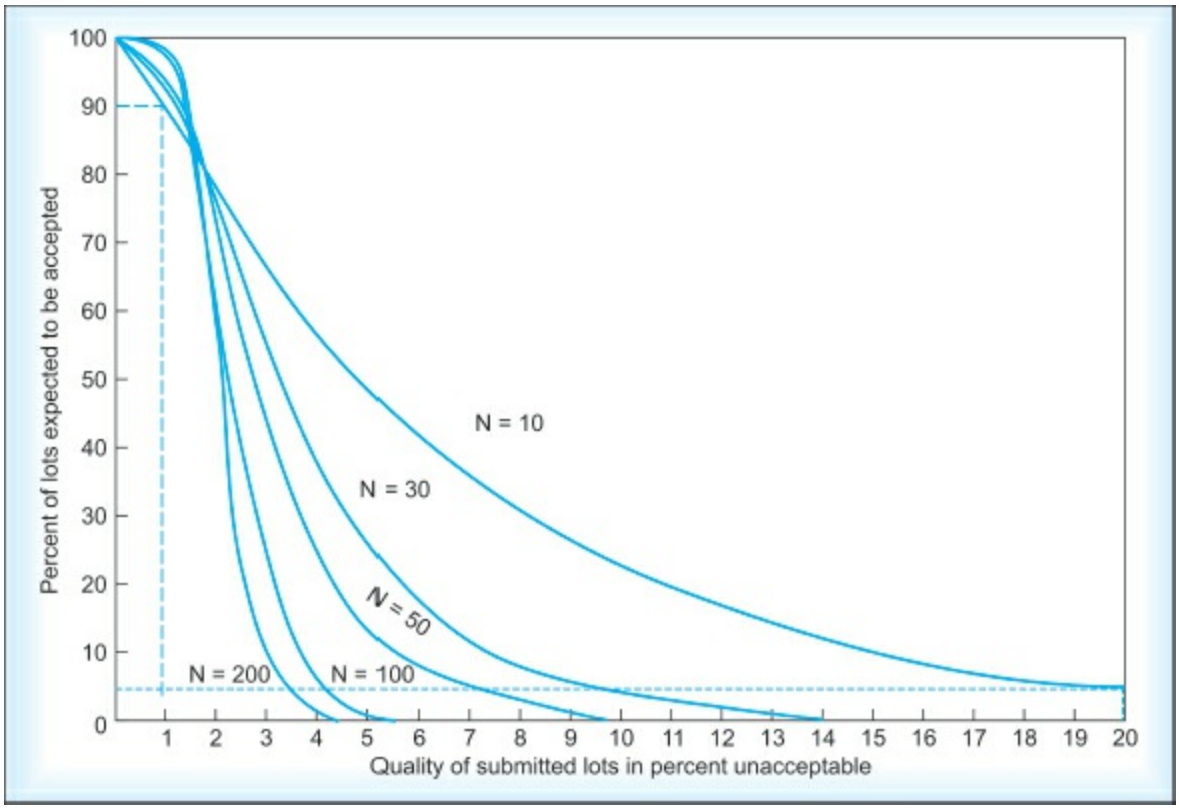



**Fig. 29.14:** Operating characteristic curves for sampling plans at AQL = 1%

When operating characteristic curves were developed for the compendial weight variation specifications, a batch of tablets containing a total of 5% underweight and overweight tablets were found to have 93% probability of being accepted when samples of 10 tablets were employed. The relations depicted in Table 29.13 show that a larger sample size increases acceptance probability. These data were obtained on the basis of a million-tablet batch.

The same curves showed that a batch with 20% unacceptable tablets had a probability of passing the weight variation test 40%, 23%, 4%, and 0% of the time as the sample size increased from 10–20 to 50–100 tablets.

Tablets are often manufactured in batch sizes consisting of a million units. For practical purposes, these batches represent infinite populations, and relatively large sample size is required to gather significant data. For such a batch, samples ranging up to 200 units are required according to the MIL-STD-414 variable sampling plan.

**Table 29.13:** Relation of sample size to acceptance probability

| Size of samples (No. of tablets) | Acceptance probability (%) |
|---|---|
| 10 | 93 |
| 20 | 95 |
| 50 | 98 |
| 100 | 99 |

The dodge sampling plan is distinctive, in that it is continuous in nature, requiring at the start that 100% of the tablets be consecutively tested until, p, tablets in succession are found to be without defect. Then 100% testing is discontinued, and a fraction, f, of the tablets is checked by random sampling. When an unacceptable sample is found, 100% testing is resumed until p tablets in succession are again found free of defect. For example, using a rotary tabletting press of 25 punches to manufacture a batch of a million tablets, let p be 50 (to contain two tablets from each punch) and f be one tablet in ten thousand. The dodge plan then calls for testing an additional 100 tablets, thus bringing the total to 150 tablets if no operation problem is found and if the process is under control. The two advantages offered by this sampling plan are (1) it reduces the amount of testing necessary when compressing proceeds as designed and predicted and (2) it increased the chances for detection and correction if the tablets fall out of specification.

In the application of sampling plans to pharmaceutical dosage forms, homogeneity of the samples must be emphasized. For homogeneous dosage forms such as oral or parenteral solutions, samplings are ordinarily taken with a sample size as small as one unit. If a drum of a powdered drug or excipient is known to be homogeneous, then for statistical purposes it can be thought of as a solution, and single or duplicate samples are considered sufficient to provide a reliable response. Semisolid dosage forms such as ointments, creams, and suspensions may be considered statistically as resembling oral or parenteral solutions when they are assumed to be homogeneous. To verify homogeneity, it is frequently necessary to take more than one sample from semisolid dosage forms. Knowledge and/or indication of the presence of heterogeneity allows the batch to be treated statistically as having an infinite population of variable components.

Statistical sampling has worked well in controlling the quality of printed material. Statistical sampling plans with a sensitivity in the range of 1 to 5%

unacceptable were considered adequate for inspection of packaging supplies several years ago; however, sensitivity of 0.1% for defects that interrupt high-speed packaging lines would be desirable. When problems such as these arise, sampling schemes must be modified, since the cost of sampling and inspecting 3000 to 5000 units would be prohibitive, and the cost of manufacturing certain products with such a low level of defectiveness would be high.

Although attribute sampling plans are the simplest to use and to enforce, variable sampling plans that yield more information are increasing in importance. This is especially true with automation of the in-process quality control functions during the manufacturing and packaging operations.

## Control and Assurance of Manufacturing Practices

The factory inspection provision of the 1962 Amendments to the Food, Drug, and Cosmetic Act empowered the FDA to inspect drug manufacturing sites (in which drugs are processed, manufactured, packaged, and stored) for compliance with accepted standards of operations, practices, and sanitation. The regulations promulgated as a result of the 1962 Amendments require a complete and full description of the controls the manufacturer employs in the preparation of dosage forms for clinical trials as well as of the marketed products routinely produced. The law recognizes that Current Good Manufacturing Practices (cGMP) with the attendant quality control procedures, are paramount to the production of quality products. As a result, the FDA has required pharmaceutical manufacturers to commit themselves to a method of manufacturing and quality assurance for each product. The manufacturer should have performed sufficient control tests to permit an evaluation of the adequacy of the manufacturing, processing, and packaging operations.

The first cGMP regulations were issued in 1963, one year after the enactment of the 1962 Kefauver-Harris Drug Amendments. Although it took about eight years (1971) to revise them, they are in no way stagnant. In fact, the word “current”, in referring to cGMP, suggests that they are dynamic, and the regulatory agency constantly updates and maintains them in relation to the current state of the art and science of drug manufacturing practices in the industry. The next revision of the cGMP took place in 1978, and these are in effect at present.

The successful application of cGMP is complex but possible if the systems governing the various phases of personnel, equipment, buildings, control records, and production procedures are at the state of proper planning and control. It should be kept in mind that cGMP is an aid and by no means a substitute for a good total quality assurance program.

## GOOD MANUFACTURING PRACTICE (GMP)

GMP is a term that is renowned worldwide for the control and management of manufacturing, testing and overall quality control of pharmaceutical products. GMP is also occasionally referred to as “cGMP”. The “current”, methodologies of manufacture, testing, design and control. The Manufacturers should employ technologies and systems, which are up-to-date. Systems and equipment used to avoid contamination, mix-ups, and errors are adequate by today’s standards. The cGMP regulations apply to all drug products, whether OTC or prescription. GMP is defined as, “that part of quality assurance which ensures that products are consistently produced and controlled to the quality standards appropriate to their intended use and as required by the marketing authorization”. GMP are aimed primarily at diminishing the risks inherent in any pharmaceutical production. Such risks are essentially of two types: cross-contamination (in particular of unexpected contaminants) and mix-ups (confusion) caused by, for example, false labels being put on containers. This in turn, protects the consumer from purchasing a product, which is not effective or even dangerous. Failure of firms to comply with GMP regulations can result in very serious consequences including recall, seizure, fines, and jail time.

## GMP Guidelines

The WHO guidelines are the basic minimum standard that should be attained by any country that is locally manufacturing pharmaceuticals. However, in practice it is only applicable to countries that do not intend to export their products widely. In many countries, it is not accepted as an appropriate standard for imports. Increasingly, there is reference to the EU or FDA requirements in this context.

The first version of the WHO GMP arose from the WHO Health Assembly in 1967. It was published in 1968 and subsequently revised in 1971 and 1975. There were no further revisions until the latest version, which was published in 1992. The latest version recognizes the existence of a number of major national and international reference texts to GMP. Additionally, it acknowledges the role of the ISO 9000 series of standards in the management of quality.

The layout of the WHO guidelines differs from that of the European and other national standards, although the overall spirit of the guidelines is the same. It is published in three parts.

**Part one** covers the general approach to quality assurance and the main elements of a GMP system.

**Part two** reviews the separate responsibilities for personnel within production and QC in order to comply with the requirements of GMP.

**Part three** contains the annexes or supplementary guidelines. At the time of publication, there were only two: manufacturing of sterile pharmaceutical products, and good manufacturing practices for active pharmaceutical ingredients (bulk drug substances). It is intended that further annexes would be added with time.

As summarized by Table 29.14, under GMP:

**Table 29.14:** The cGMPs regulations for finished pharmaceuticals, as set forth in 21 CFR part 211

| Subparts | Subpart title | Topics covered in sections | |
|---|---|---|---|
| A | General Provisions | §211.1 | Scope |
| | | §211.2 | Definitions |
| B | Organization and Personnel | §211.22 | Responsibilities of Quality Control Unit |
| | | §211.25 | Personnel Qualifications |
| | | §211.28 | Personnel Responsibilities |
| | | §211.34 | Consultants |
| C | Buildings and Facilities | §211.42 | Design and Construction Features |
| | | §211.44 | Lighting |
| | | §211.46 | Ventilation, Air Filtration, Air Heating, and Cooling |
| | | §211.48 | Plumbing |
| | | §211.50 | Sewage and Refuse |
| | | §211.52 | Washing and Toilet Facilities |
| | | §211.56 | Sanitation |
| | | §211.58 | Maintenance |
| D | Equipment | | |
| | | §211.63 | Equipment Design, Size, and Location |
| | | §211.65 | Equipment Construction |
| | | §211.67 | Equipment Cleaning and Maintenance |
| | | §211.68 | Automatic, Mechanical, and Electronic Equipment |
| | | §211.72 | Filters |
| E | Control of Components and Drug Product Containers and Closures | §211.80 | General Requirements |
| | | §211.82 | Receipt and Storage of Untested Components, Drug Product Containers, and Closures |
| | | §211.84 | Testing and Approval or Rejection of Components, Drug Product Containers, and Closures |
| | | §211.86 | Use of Approved Components, Drug Product Containers, and Closures |
| | | §211.87 | Retesting of Approved Components, Drug Product Containers, and Closures |
| | | §211.89 | Rejected Components, Drug Product Containers, and Closures |
| | | §211.94 | Drug Product Containers and Closures |
| F | Production and Process Controls | §211.100 | Written Procedures; Deviations |
| | | §211.101 | Charge-In of Components |
| | | §211.103 | Calculation of Yield |
| | | §211.105 | Equipment Identification |
| | | §211.110 | Sampling and Testing of In-Process Materials and Drug Products |
| | | §211.111 | Time Limitations on Production |
| | | §211.113 | Control of Microbiological Contamination |
| | | §211.115 | Reprocessing |
| G | Packaging and Labeling Control | §211.122 | Materials Examination and Usage Criteria |
| | | §211.125 | Labeling Issuance |
| | | §211.130 | Packaging and Labeling Operations |
| | | §211.132 | Tamper-Resistant Packaging Requirements for Over-The-Counter Human Drug Products |
| | | §211.134 | Drug Product Inspection |
| | | §211.137 | Expiration Dating |
| H | Holding and Distribution | §211.142 | Warehousing Procedures |
| | | §211.150 | Distribution Procedures |
| I | Laboratory Controls | §211.160 | General Requirements |
| | | §211.165 | Testing and Release for Distribution |
| | | §211.166 | Stability Testing |
| | | §211.167 | Special Testing Requirements |
| | | §211.170 | Reserve Samples |
| | | §211.173 | Laboratory Animals |
| | | §211.176 | Penicillin Contamination |
| J | Records and Reports | §211.180 | General Requirements |
| | | §211.182 | Equipment Cleaning and Use Log |
| | | §211.184 | Component, Drug Product Container, Closure, and Labeling Records |
| | | §211.186 | Master Production and Control Records |
| | | §211.188 | Batch Production and Control Records |
| | | §211.192 | Production Record Review |
| | | §211.194 | Laboratory Records |
| | | §211.196 | Distribution Records |
| | | §211.198 | Complaint Files |
| K | Returned and Salvaged Drug Products | §211.204 | Returned Drug Products |
| | | §211.208 | Drug Product Salvaging |

a. All manufacturing processes are clearly defined, systematically reviewed in the light of experience, and shown to be capable of consistently manufacturing pharmaceutical products of the required quality that comply with their specifications.
b. Qualification and validation are performed.
c. All necessary resources are provided, including:
    i. Appropriately qualified and trained personnel.
    ii. Adequate premises and space.
    iii. Suitable equipment and services.
    iv. Appropriate materials, containers and labels.
    v. Approved procedures and instructions.
    vi. Suitable storage and transport.
    vii. Adequate personnel, laboratories and equipment for in-process controls.
d. Instructions and procedures are written in clear and unambiguous language, specifically applicable to the facilities provided.
e. Operators are trained to carry out procedures correctly.
f. Records are made (manually and/or by recording instruments) during manufacture to show that all the steps required by the defined procedures and instructions have in fact been taken and that the quantity and quality of the product are as expected; any significant deviations are fully recorded and investigated.
g. Records covering manufacture and distribution, which enable the complete history of a batch to be traced, are retained in a comprehensible and accessible form.
h. The proper storage and distribution of the products minimizes any risk to their quality.
i. A system is available to recall any batch of product from sale or supply.
j. Complaints about marketed products are examined, the causes of quality defects investigated, and appropriate measures taken in respect of the defective products to prevent recurrence.

## Different Versions of GMP

There are many different versions of GMP in use around the world (Table 29.15). As more and more countries improve the standards achieved within their pharmaceutical industries, they all go through the stage of wanting to have their own guidelines in place. There are, however, a small number of versions in place that were developed some 30 years ago, when the concept of quality manufacture was being put in place. Although the wording of the various versions of GMP may differ to a greater or lesser extent, the spirit of what is to be achieved is the same in all cases. It is often a question of degree and interpretation.

**Table 29.15:** Different versions of GMP in use around the world

| Countries | Title of GMP guidelines | Topics covered |
|---|---|---|
| EU Member State: The UK | 'The Orange Guide' or Rules and guidance for pharmaceutical manufacturers and distributors | A number of documents relating to manufacturing licences and qualified persons in addition to GMP guidelines. |
| European Union | Directive 91/356/EEC | Quality management, personnel, premises and equipment, documentation, production, quality control, work contracted out, complaints and product recall, self-inspection |
| The USA | Code of Federal Regulations, Title 21, Parts 210–226 | Enlisted in following table |
| Australia | The Australian Code of Good Manufacturing Practice | General provisions of GMP, the use of computers, the manufacture of sterile products, appendices on radiation sterilization, supplementary |

| | | |
|---|---|---|
| | | notes for hospital pharmacists, guidelines on tests for stability, on laboratory instrumentation, on industrial ethylene oxide sterilization of therapeutic goods, and guidelines for the estimation of microbial count in process water |
| India | (SCHEDULE M) Rules 71, 74, 76 and 78 | GMP Requirements of factory premises for manufacture of:<br>1. Oral solid dosage form, sterile products, oral liquids, topical products and MDIs (M)<br>2. Homoeopathic preparations (M-I),<br>3. Cosmetics (M-II),<br>4. Medical devices (M-III) |

## GMP Certification

The steps for GMP implementation are shown in Fig. 29.15. Once certified, to an organization, it is an authorization and certification of company's product and process, that quality standards are adequate, up-to-date and controlled for the intended use by the consumer.

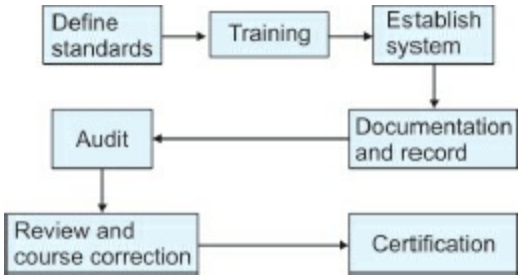



**Fig. 29.15:** Steps for GMP implementation and certification

## Benefits of GMP

1. Most GMP requirements are very general, open-ended, and provide much flexibility, allowing each manufacturer to decide individually how to best implement the necessary controls.
2. GMP is a good business tool, which will help to refine both compliance and performance of the Company.
3. GMP requirements are largely common sense practices, which will help companies better itself as it moves toward a quality approach using continuous improvement.

## *Personnel*

One criterion for a successful quality assurance program is the encouragement of quality consciousness in the personnel of the entire company. Proper selection, training, and motivation of production, packaging, and control personnel are vital to produce quality pharmaceuticals consistently. The degree to which the desired quality for the product is attainable is proportional to the attitudes or desires of the individuals working in production, packaging, and control. By building a sense of pride in performance and showing the importance of the contributions of these individuals in producing products that could be lifesaving, the risk of errors can be minimized. In reality, this is the basis for the so-called control of quality by the zero-defect concept, which operates through prevention rather than detection of the mistakes by properly directing and motivating personnel. Quality work and products result from this approach.

It is essential that qualified personnel be employed to supervise the formulation, processing, sampling, testing, packaging, and labeling of the drug product, and that competent staff be placed in charge of the maintenance of machinery, equipment, and sanitation. The qualified personnel are persons who by virtue of education, training, and experience have the knowledge and ability to execute the technologic assignments. The key personnel involved in the manufacture and control of the drug should assume the responsibility of assuring that the drug and the dosage form they are handling have the desired characteristics. The responsibility for keeping the manufacturing process within the cGMP regulations of the FDC Act must be delegated to the people directly involved with the various aspects of the process and their immediate

supervisors. The supervisors are required to provide the necessary direction and control of the operation and be available at all times in case a question or a problem arises. The operating personnel should have the necessary authority to sign the manufacturing documents for each process for which they are responsible. The document is then countersigned by the supervisor. The signature and endorsement should appear on the proper worksheet at the completion of a production operation.

## *Equipment and Buildings*

Equipment and buildings used in the manufacture, processing, packaging, labeling, storage, or control of drugs should be of suitable design, size, construction, and location and should be maintained in a clean and orderly manner.

The building should provide adequate space for the orderly placement of materials and equipment to minimize any risks of mixups or cross-contamination between the drugs, excipients, packaging, and labeling from the time the materials are received to the time the products are released. Adequate lighting, ventilation, dust control, temperature, and humidity should also be provided. To avoid conditions unfavorable to the integrity and safety of the product, other considerations may be required for particular operations and products, such as bacteriologic controls for preventing microbial contamination, for sanitizing work areas for parenteral preparations, and for preventing the dissemination of microorganisms from one area to another or from previous manufacturing operations.

The desired characteristics of equipment for producing quality products are numerous, and they differ from machine to machine; however, the equipment should be of suitable size, accuracy, and reproducibility. Their surfaces should be inert, nonreactive, nonabsorptive, and nonadditive so that the identity, purity, and quality of the drug substance and other components are not affected to any significant extent. The equipment should be constructed to facilitate adjustment, cleaning, and maintenance to assure the precision, reliability, and uniformity of the process and product, and to assure the exclusion of contaminants from previous and current manufacturing and packaging operations.

## *Control of Records*

The records, such as master formula and batch production records, should be prepared and maintained in accordance with established procedures.

## *Master Formula Record*

Master formula records for each product should be prepared, endorsed, and dated by a competent and responsible individual and should be independently checked, endorsed, and dated by another competent and responsible individual. The information contained in the records should be provided in a format and language that will not be misinterpreted by the operating personnel and the supervisor, to assure that each batch of a product can be identically reproduced.

Although the content and format may differ from product to product, the master formula record shall include the following information:

1. The name of the product, a description of the dosage form, and its strength.
2. The complete list of ingredients, designated by whole names and codes sufficiently specific to indicate any special characteristic.
3. The quantity by weight or volume of each ingredient, regardless of whether it appears in the finished product. If variations in the quantity of a particular ingredient are permitted, as is sometimes necessary in the preparation of a dosage form, an adequate statement should be provided in the record.
4. The standards or specifications of each ingredient used in the product.
5. An appropriate statement concerning any calculated excess of an ingredient.
6. Appropriate statements of theoretic yield at various stages and the termination of processing.
7. Manufacturing and control instructions, specifications, precautions, and special notations to be followed.
8. A detailed description of the closures, containers, labeling, packaging, and other finishing materials.

## *Batch Production Record*

Batch production records should be prepared, maintained, and controlled for

each batch of product. In general, they should be retained for a period of approximately five years after distribution has been completed. The batch production record shall contain an accurate reproduction of the manufacturing formula, procedure, and product specifications from the correct master formula procedure to be used in the production of a batch of product. These batch records are then sent to each of the departments involved in the production, packaging, and control of the product. The records include dates, specific code or identification numbers of each ingredient employed, weights or measures of components and products in the course of processing, results of in-process and control testing, and the endorsements of the individual performing and supervising each step of the operation.

In addition, a lot number is assigned that permits the identification of all procedures performed on the lot and their results. This lot number appears on the label of the product. This procedure facilitates a search for the details of manufacture and control history of any particular product.

## *Control of Production Procedures*

To ensure that products have the intended characteristics of identity, strength, quality, and purity, production and the related inprocess quality control (IPQC) procedures should be rigidly followed as required by the master formula record or the batch production record. To a large extent, IPQC is concerned with providing accurate, specific, and definite descriptions of procedures to be employed from the receipt of raw materials to the release of the finished dosage forms. It is a planned system to identify the materials, equipment, processes, and operators; to enforce the flow of manufacturing and packaging operations according to the established rules and practices; to minimize human error or to detect the error if and when it does occur; and to pinpoint the responsibility to the personnel involved in each unit operation of the entire process. In general, the in-process control procedures are usually rapid and simple tests or inspections that are performed when the manufacturing of a product batch is in progress. They are used to detect variations from the tolerance limits of the product so that prompt and corrective action can be taken. The in-process control procedures and tests should be openly discussed, experimentally justified, written in detail, properly explained, and in particular, rigidly enforced once they are

established.

For the convenience of discussion, the actual production procedures are subdivided and discussed as manufacturing control and packaging control.

## *Manufacturing Control*

Although the scope and structure of the manufacturing control operation differs appreciably from company to company, it is possible to highlight their common elements. The production planning department issues the formula and manufacturing worksheets bearing an identification number, tide of product, and the names and quantities for all ingredients, together with a complete description of the procedures to be followed, the precautions to be taken, the equipment to be used, the lot sizes to be processed, and the suitable in-process controls to be undertaken.

Material tickets for each raw material are written and issued by the production department to the department of material stores, where the orders are filled and verified. After the materials have been sent to the production department and checked similarly, the pattern of control takes shape as production proceeds. The addition of raw materials to the batch is verified and countersigned by a qualified person. Notation is made on the manufacturing worksheet of the identifying number of each ingredient as it is used, and each unit operation is checked off as it is performed. An appropriate label is attached to each container or piece of equipment in use to identify its contents and to ensure that the in-process stage is properly designated. Any deviation from standard operating conditions, no matter how small, should be reported to both production and control personnel responsible for the product.

The in-process checking during manufacturing plays an important role in the auditing of the quality of the product at various stages of production. Duties of the auditor or the control inspector consist of checking, enforcing, and reviewing procedures and suggesting the change for upgrading the procedures when necessary.

The primary objective of an IPQC system is to monitor all the features of a product that may affect its quality and to prevent errors during processing. Only the most commonly practiced methods for parenterals, solid dosage forms (tablets and capsules), and semisolid preparations (ointments, creams, and lotions) are briefly described.

For parenteral products, the in-process quality controls are (1) checking the bulk solution, before filling, for drug content, pH, color, clarity, and completeness of solution, (2) checking the filled volume of liquids or the filled weight of sterile powders for injection in the final containers at predetermined intervals during filling, (3) testing for leakage of flame-sealed ampoules, (4) subjecting the product to physical examination (visually or mechanically) for appearance, clarity, and particulate contamination, (5) examining the sterility indicator placed in various areas of the sterilizer for each sterilization operation and (6) submitting the product for sterility testing or other predetermined biologic tests to establish the safety and other parameters of the product.

The in-process quality controls for solid dosage forms are (1) determining the drug content of the formulation, (2) checking the weight variation for tablets and capsules at predetermined intervals during manufacturing, (3) checking the disintegration and/or dissolution time, hardness, and friability of the tablets at least during the beginning, middle, and end of production or at prescribed intervals during manufacturing, (4) testing soluble tablets for compliance with solution time requirements and (5) examining products by line inspection or other equally suitable means and removing the defective units prior to packaging.

For semisolid preparations, the following in-process controls are available (1) checking for uniformity and homogeneity of drug content prior to the filling operations, (2) determining the particle size of the preparation when appropriate, (3) checking the appearance, viscosity, specific gravity, sediment volume, and other physical parameters at prescribed intervals, (4) testing for filling weight during the filling operation and (5) testing for leakage on the finished jars or tubes.

At the completion of the manufacturing process as well as in-process stages, actual yields are checked against theoretic value, and the representative samples are withdrawn for laboratory testing by the control inspector according to the predetermined sampling plan. The operators actively performing the process, their supervisors, and the control inspector must all verify that the entire operation was accomplished in the prescribed manner. Occasionally, materials in bulk storage are sampled at random and are examined to determine that no detectable change has taken place, and that the batch is satisfactory for final packaging.

The batch production records and other needed documents are then delivered to the quality control office, together with the withdrawn samples of the product. These records and test results are reviewed for conformance to specifications and CGMP. The bulk finished products are held in quarantine until they are released for packaging by quality control personnel.

## *Packaging Control*

At some time before the manufacture of a product is completed, a packaging record bearing an identification number is issued to the packaging section. This record specifies the packaging materials to be used, operations to be performed, and the quantity to be packaged. Simultaneously, requisitions are issued for the products to be packaged and for the packaging and printed materials, such as labels, containers, inserts, brochures, cartons, and shipping cases.

The packaging operation unites the product, container, and label to form a single finished unit. Not only must individual package components be correct, but they must also be assembled correctly. Prior to the start of a packaging operation, the control inspector and packaging supervisor must check and verify the line to ensure that it has been thoroughly cleaned and that all materials from the previous packaging order have been completely removed. In addition, various pieces of packaging equipment are stripped down, cleaned, and examined when an order is finished. The bulk of the product and each of the packaging components should be checked, endorsed, and dated by qualified packaging personnel with the cooperation of a control inspector. In practice, only the exact number of labels required for a batch, including a small excess, should be delivered to the labeling area after careful and meticulous inspection of each label.

In the packaging area, a large group of people work on several different products simultaneously, using high-speed packaging equipment. Therefore, the operation should be performed with adequate physical segregation of products. Dosage forms of similar color, size, and shape should not be scheduled consecutively on a line. Tablets of similar shape should not be scheduled on the neighboring packaging lines at the same time, even though the size or color may be dissimilar. Proper on-line inspection should be made during the packaging operation to ensure absence of foreign drugs and labels, adequacy of the containers and closures, and accuracy of labeling. Proper

reconciliation and disposition of the unused and wasted labels should occur at the end of the packaging operation. The yield must be justified against the theory represented by the batch size of the starting materials. Suitable and reasonable procedures should be established for action to be taken when an unexplained discrepancy exists between the number of labels issued and the number accounted for on the finished product. The common approach is for key personnel to prevent distribution of the batch in question, and of other batches of products that were packaged during the same period of time, until a satisfactory explanation can be obtained for the discrepancy.

## *Validation*

The terms validation and qualification have, in recent years, become familiar in connection with pharmaceutical processes. Qualification is generally related to equipment and is used to determine whether the equipment operates as it was designed to in a reproducible manner. *Validation* of a process is the demonstration that controlling the critical steps of a process results in products of repeatable attributes (e.g. content uniformity) or causes a reproducible event (e.g. sterilization).

The concept of applying a systems approach to pharmaceutical manufacture and control, requiring validation of the process and qualification of equipment, facilities, personnel, and so forth, received considerable impetus when it was recognized by both the FDA and industry that sampling and testing of finished products alone cannot provide the necessary assurance of drug product quality within and between batches. The customary sample size in end product testing does not provide sufficient statistical validity for high product quality; it alone cannot verify that the various factors in the system intended to assure quality within and between batches of product are functioning as they were designed to function. Such verification can only be accomplished through identifying the critical components of the system and implementing control tests for these components, which when taken as a whole, demonstrate that their characteristics are repeatable from batch to batch of product. Consequently, a validated process is a systematic, documented program that provides a high degree of assurance that a specific process will consistently produce a product meeting its predetermined specifications and quality attributes.

The FDA, in its program guidance manual to FDA investigators, defines a

validated manufacturing process as follows:

> "*A* validated manufacturing process is one that has been proved to do what it purports or is represented to do. The proof of validation is obtained through collection and evaluation of data, preferably, beginning from the process development phase and continuing through into the production phase. Validation necessarily includes process qualification (the qualification of materials, equipment, systems, building, personnel), but it also includes the control of the entire processes for repeated batches or runs".

The FDA interprets Subpart F Section 211.100(a) and 211.110(a) of the Current Good Manufacturing Regulations for drug products to mean that pharmaceutical processes must be validated. These two sections of the regulations are presented:

Section 211.100(a)—"There shall be written procedures for production and process control designed to assure that the drug products have the identity, strength, quality and purity they purport or are represented to possess. Such procedures shall include all requirements in this subpart. These written procedures including any changes, shall be drafted, reviewed and approved by the appropriate organizational units and reviewed and approved by the quality control unit".

Section 211.110(a)—"To assure batch uniformity and integrity of drug products, written procedures shall be established and followed that describe the in-process controls, and tests or examinations to be conducted on appropriate samples of in-process materials of each batch. Such control procedures shall be established to monitor the output and to validate the performance of these manufacturing processes that may be responsible for causing variability in the characteristics of in-process material and drug product".

The documentation, including protocols, data, and results obtained from process validation and equipment qualification are important, since the validation performed should be auditable by an appropriate responsible individual who, after reviewing the records, should be able to certify the validation of the process to produce products to defined attributes consistently provided that the system validated is not altered.

## Control and Assurance of Finished Products

Good manufacturing practices, well-defined specifications, sound sampling procedures, and efficient process controls do not in themselves constitute an overall quality control program. Unless the testing procedures by which the product quality is finally measured are established on an equally sound basis, the entire system may be deficient. Product failures causing rejections or recalls after market introduction are serious and can be easily detected and minimized by an effectively administered quality testing program. Therefore, the testing of the finished products for compliance with the established standards prior to release of the material for distribution is a critical factor for quality control and assurance. The testing indicates the possible deviations from perfection that occur in the batch. Product quality assurance is not complete with the release of the batch, however. The stability of the product in the marketed package should be repeatedly reconfirmed by actual physical, chemical, and biologic tests performed on several representative batches of the product over the period of its expected shelf-life. The activities associated with the control of the dosage form after manufacture are discussed briefly in this section.

### *Types of Specifications*

The element of potential hazard to public health and the risk of violating stringent and exacting statutes render specification writing for the pharmaceutical industry quite unique. The main purpose of establishing specifications is to ensure that the characteristics of the finished dosage forms conform to appropriate standards of identity, quality, purity, potency, safety, and efficacy.

The first four types of specifications (i.e. identity, quality, purity, and potency) are distinctively analytic in nature and are embodied in specifications known as drug standards. Examples of tests for standard of identity, purity, and quality are shown in Table 29.16. These drug standards may be further delineated and illustrated by citing the typical examples in the following sections.

**Table 29.16:** Examples of tests for standards of identity, purity, and quality

| Standard of identity | Standard of quality | Standard of purity |
|---|---|---|

| | | |
|---|---|---|
| Color formation | Absorbance | Color and/or odor |
| Precipitation | Refractive index | Clarity and/or color of solution |
| Decomposition | Optical rotation | Acidity or alkalinity |
| Derivative formation | Specific gravity | Acid of alkali |
| Infrared spectra | pH | Inorganic salt |
| Ultraviolet spectra | Viscosity | Heavy metals |
| Visible spectra | Melting point or range | Foreign matter |
| Specific reactions | Saponification value | Residue on evaporation |
| Cations or anions determination | Acid value | Readily carbonizable substances |

## *Standard of Identity*

Identity tests are usually the distinctive qualitative chemical methods used to confirm the actual presence of the compound. For some drugs, microbiologic and pharmacologic tests also may be employed. Examples are:

1. The infrared absorption spectrum for chlorothiazide dispersed in mineral oil exhibits maxima only at the same wavelengths as that of a similar preparation of chlorothiazide reference standard.
2. Ouabain, when dissolved in sulfuric acid, produces a dark red color by transmitted light and shows a greenish fluorescence in reflected light.
3. The ultraviolet absorption spectrum of propranolol hydrochloride in methanol exhibits maxima and minima at the same wavelengths as that of a similar preparation of reference standard. The respective absorptive values at the maximum wavelength do not differ by more than 2.5%.
4. A solution of quinidine gluconate in dilute sulfuric acid exhibits a vivid blue fluorescence. On the addition of a few drops of hydrochloric acid, the fluorescence disappears.

## *Standards of Quality*

Quality tests are usually the physical methods used to measure accurately the

characteristic properties of the drug. Results are expressed as a permissible range of values for a measured property of the drug. Examples are:

1. The specific notation of propranolol hydrochloride is between −1.0° and +1.0° calculated on the dried basis.
2. A solution of quinidine gluconate is dextrorotary.
3. The refractive index of clofibrate is between 1500 and 1505 at 20°.
4. The specific gravity of ethanol is between 0.812 and 0.816 at 15.56°.

## *Standards of Purity*

Purity tests are generally designed to estimate the levels of all known and significant impurities and contaminants in the drug substance under evaluation. These standards are numerically expressed as maximum tolerable limit or the absence of an impurity or contaminant based on the prescribed methods. Examples are:

1. The maximum tolerable limit of diazotizable substances of chlorothiazide is 1.0%.
2. A solution of ouabain yields no precipitate with tannic acid or with iodine to indicate the absence of alkaloids.
3. The residue on ignition of propranolol hydrochloride is not more than 0.1%.
4. After drying quinidine gluconate at 105° for 1 hour, it loses not more than 0.5% of its weight.

## *Standards of Potency*

Potency tests are assays that estimate the quantity of active ingredient in the drug. Employing physical, chemical, biologic, pharmacologic, or microbiologic means, these quantitative tests yield the strength or potency of the drug substance. Examples are:

1. Chlorothiazide contains not less than 98.0% and not more than 100.5% of $C_7H_6Cl\ N_3O_4S_2$, calculated on the dried basis.
2. Ouabain contains not less than 95.0% and not more than 100.5% of $C_{29}H_{44}O_{12}.8H_2O$.
3. Propranolol hydrochloride contains not less than 98.0% and not more

than 101.5% of $C_{16}H_{21}NO_2$.HCl calculated on the dried basis.

4. Quinidine gluconate contains not less than 99.0% and not more than 100.5% of total alkaloid salt, calculated on the dried basis as:

$$C_{20}H_{24}N_2O_2.C_6H_{12}O_7$$

It is content of dihyroquinidine gluconate is not more than 20.0% by weight of its content of total alkaloid salt, calculated as:

$$C_{20}H_{24}N_2O_2.C_6H_{12}O_7$$

Examples of standards of potency for several drug substances and their dosage forms are listed in USP XXI and are demonstrated in Table 29.17. For drug substance, most of the chemically synthesized raw materials have a lower specification limit of at least 97% potency with the exception of antibiotics such as chloramphenicol, ampicillin, and tetracycline hydrochloride, which in general have a lower potency limit of 90% or more. As shown in Table 29.17, different dosage forms of the same drug substance are assigned the same specification limits, such as. 90.0 to 130.0% for chloramphenicol capsules, ophthalmic solutions, and ophthalmic ointments, and 95.0 to 110.0% for all dosage forms containing promazine hydrochloride. As shown in Table 29.15, however, different dosage forms of the same drug substance may be assigned different specification limits as well. Using ephedrine sulfate as an example, limits are 90.0 to 110.0% for syrups, 92.0 to 108.0% for capsules, 95.0 to 105.0% for injections, and 93.0 to 107.0% for tablets and nasal solutions. It should be noted that the positive and negative specification limits are generally symmetric around the potency for most of the dosage forms. Therefore, in many cases, the potency range is expressed in the form of 100 ± 10% or 100 ± 7% than 90 to 110% or 93 to 107%.

**Table 29.17:** Standard of potency of different dosage forms having same limits

| | **Standard of potency expressed as percentage of claim** | |
|---|---|---|
| **Drug substance** | **Raw material** | **Dosage forms** |
| Amitriptyline HCl | 99.9–100.5 | 90.0–110.0 (Injection, |

| | | |
|---|---|---|
| | | tablet) |
| Chloramphenicol | ≥90.0 | 90.0–130.0 (Capsule, ophthalmic solution, ointment) |
| Chlorpromazine HCl | 98.0–101.5 | 95.0–105.0 (Injection, syrup, tablet) |
| Digoxin | 97.0–103.0 | 95.0–110.0 (Injection, elixir, tablet) |
| Fluphenazine HCl | 97.0–103.0 | 95.0–110.0 (Injection, oral solution, elixir, tablet) |
| Griseofulvin | 90.0–105.0 | 90.0–115.0 (Oral suspension, tablet, capsule) |
| Meperidine HCl | 98.0–101.0 | 95.0–105.0 (Injection, syrup, tablet) |
| Reserpine | 97.0–101.0 | 90.0–110.0 (Injection, elixir, tablet) |
| Promazine HCl | 97.0–101.5 | 95.0–110.0 (Injection, oral solution, syrup, tablet) |

Data compiled from USP XXI, United States Pharmacopeial Convention, Rockville, MD, 1980.

A drug substance is suitable for use if the totality of data derived from these four attributes shows that it meets all of the specifications.

In general, specifications should provide for the maintenance of reasonable standards, while allowing a sensible degree of latitude for manufacturing variations or tolerances, and acceptable latitude for analytic errors, particularly in control techniques. To guide in the appropriate development of specifications, historical data accumulated from batch analyses of the drug substance or the drug product are usually the best source (Table 29.18).

| **Table 29.18:** Utilization of historical batch data to guide the specification for standard of purity | |
|---|---|
| Drug substance: | An experimental antihypertensive |
| Proposed limit of purity: | 1% of an impurity |
| Method: | High-pressure liquid chromatography for an impurity |
| Detection limit: | 0.1% |
| Batch record: | 0.2 0.2 0.3 0.1 < 0.1 0.3 < 0.1 0.1 0.1 0.3 |
| Acceptable limit: | 0.6% |
| “House” limit: | 0.5% |

Specifications are also used in the procurement of drug substance, excipients, reagents, packaging, and printed materials. The incoming materials and the finished products are checked against the specifications to an extent sufficient to determine the compliance and the acceptability of the materials and products. Therefore, the specifications set forth as the standard should be discriminating enough to differentiate good or acceptable material from inferior or rejectable material. The specifications should be practical and realistic, however, and they should reflect those parameters necessary to define the product as well as to permit manufacture of the product to a defined quality level. In developing specifications for purchased material, the vendor’s capability to supply the material should also be considered.

In the development of specifications, the following objectives should be carefully considered (1) to ascertain which physical, chemical, and biologic characteristics of dosage form are critical, which are important, which are helpful, and which are not particularly important but are useful, (2) to decide which dosage form characteristics shall be established as the criteria for evaluating routine production batches, (3) to establish the appropriate test methods for evaluating the selected criteria and (4) to determine the acceptable tolerances and limits for each of the dosage form characteristics.

In deciding the elements that will contribute to the making of a satisfactory specification for a drug, the first question to be asked must always be “For what purpose is the specification needed?” No single set of

requirements meets all possible situations.

## *Specifications for New Products*

During the development of any new drug, there must be a rigorous investigation of both the chemical and physical properties of the material. A knowledge of the synthetic and appropriate testing by nuclear magnetic resonance, infrared spectroscopy, and mass spectrometry route usually provides strong evidence of structure, together with hints as to the identity of possible by-products that might be present. Specific chemical reactions may give further confirmation of structure. As frequently happens, possible contaminants bear a close structural relationship to the drug substance. Two or three percent of such substances may produce no readily detectable effect even by infrared spectroscopy. It is necessary to recognize the presence of impurities present at such low levels, however, because the contaminant might have an undesirable or even harmful effect. Thus, it is reasonable that its presence should be defined and controlled.

Although, ideally, a drug should be absolutely pure, this is generally not feasible in practice. A more practical approach is to require that the material to be subjected to pharmacologic and clinical trial is prepared to be as pure as is economically reasonable and then characterized as fully as possible with respect to its impurity content. If the trials prove satisfactory and acceptable, then the aim should be to ensure that subsequent batches of material should be at least as “clean” as the trial material. It is essential that the total identity of the material being used, including both its major and minor constituents, be known. Only in this way can the results of clinical trials be meaningful for a new drug.

## *Specifications for Well-established Products*

It is important to bear in mind that material synthesized on a pilot scale for clinical trial work may not be the same as material synthesized in a full-scale manufacturing process, especially in respect to its pattern of minor constituents.

What is a tolerable level of impurity in a drug? This question, of course, has no answer. Every impurity in each drug must be considered as a separate and individual case. If the impurity has been shown to have, or is suspected to have, undesirable properties, it should be limited as rigorously as possible.

## *Official Specifications*

The manufacturer's specifications are designed to be applied to the drug at the time of manufacture or quite shortly thereafter. An official specification, on the other hand, should be designed to apply to the material at any time during its period of possible use. An official specification is likely to be somewhat less stringent, especially with regard to changes due to slight decomposition.

Specifications for a product are developed not only to assure product quality but also to detect and identify impurities. There are two major kinds of impurities: product-nonspecific impurities, which are introduced externally into the product during processing, and product-specific impurities, which appear as by-products or degraded products of the drug substances as well as excipients used in the product. Therefore, the test parameters should be designed, and the testing procedures for the specifications should be established, so that the tests are capable of detecting product-specific as well as product-nonspecific impurities.

A description of the tests and reagents for various official drugs and their dosage forms are available from the official compendia. From content uniformity to dissolution tests, and from identity tests to assay procedures, each section outlines specific procedures that are mandatory for the proper control of the product.

All drug specifications listed in pharmacopeias have been verified in collaborating laboratories of pharmaceutical companies and academic institutions. The pharmacopoeia issued for a country is the legal standard of that nation. The specifications in a pharmacopoeia are so designed that the tests and the requirements for acceptance are applicable to all manufacturers products. This means that anyone who manufactures' a product of that type should conform with those specifications. It should be recognized that the specifications set in the compendium are minimum standards to which the product is expected to conform at any time during its expiry date period. A partial list of official compendia from several countries and WHO is presented in Table 29.19. A compendium is a collection of monographs of drug substances and drug products. "Official" as used in "official compendium" signifies governmental authorization. As defined in the FD and C Act of 1938, the phrase "official compendium" means the Official United

States Pharmacopeia, Official Homeopathic Pharmacopeia of the United States, Official National Formulary, or any of their supplements.

**Table 29.19:** Partial list of pharmacopeias and other books of standards

| Country | Pharmacopeia or other standards |
|---|---|
| Federal Republic | Deutsches Arzneibuch and of Germany supplements |
| France | Pharmacopee Francaise |
| Japan | The Pharmacopeia of Japan |
| Sweden | Pharmacopeia Nordica |
| Switzerland | Pharmacopeia Helvetica and its supplements |
| Union of Soviet Socialist Republics | State Pharmacopeia of the Union of Soviet Socialist Republics |
| United Kingdom | British Pharmacopeia; The Pharmaceutical Codex |
| United States of America | The United States Pharmacopeia/National Formulary; Code of Federal Regulations, 21 Foods and Drugs |
| World Health Organization | Pharmacopeia Internationales |

The complexity of quality control testing can be more clearly understood from the quality control profiles of a hypothetical parenteral dosage form containing sulfadimethoxide as an active ingredient (Table 29.20), and a hypothetical tablet dosage form containing glutethimide as the active ingredient (Table 29.21). For manufacturing the parenteral sulfadimethoxide injection, there are seven items or raw materials to be tested and cleared by quality control laboratory according to the specifications designed for each item. There are approximately 70 tests required prior to the release of these seven items for compounding. Similarly, a large number of tests are required to accept eight items for fabricating glutethimide tablets. There are approximately 80 tests that need to be performed to determine the acceptability of these eight ingredients. Based on the examples shown in

Tables 29.20 and 29.21, the average number of quality control tests required to accept or reject a raw material is approximately ten. It is evident that the total number of tests required by the official compendia on the ingredients in the dosage form are numerous. In addition to these tests are added the quality control tests on the finished dosage forms of sulfadimethoxide injection and glutethimide tablets according to the product specifications.

**Table 29.20:** Quality control profile of various raw materials used in a proposed parenteral sulfonamide injection

| Item | Function | Tests required |
|---|---|---|
| Sulfadimethoxine, NF XVI | Antibacterial | Appearance; solubility; identity; melting range; ultraviolet absorption; loss on drying; residue on ignition; content of heavy metals; assay |
| Glycerine, USP XXI | Cosolvent | Appearance; solubility; identity; color; specific gravity; residue on ignition; contents of chloride, sulfate, arsenic, and heavy metals; limit of chlorinated compound; readily carbonizable substances; limit of fatty acids and esters |
| Benzyl alcohol, NF XVI | Preservative | Appearance; solubility; identity; specific gravity; distilling range; refractive index; residue on ignition; limits for aldehyde and chlorinated compounds |
| Sodium bisulfite, NF XVI | Antioxidant | Appearance; solubility; identity; contents of arsenic, iron, and heavy metals; assay |
| Disodium edetate, USP XXI | Chelator | Appearance; solubility; identity; pH; loss on drying; contents of calcium and heavy metals; assay |
| Sodium hydroxide, NF XVI | pH adjuster | Appearance; solubility; identity; insoluble substances and organic |

| | | |
|---|---|---|
| | | matters; contents of potassium and heavy metals; assay |
| Water-for-injection, USP XXI | Solvent | Appearance; reaction; chloride, sulfate, ammonia, calcium, carbon dioxide and heavy metals; total solids and oxidizable substances; pyrogen test |

Data compiled from USP XXI/NF XVI, United States Pharmacopeial Convention, Rockville, MD, 1980.

**Table 29.21:** Quality control profile of various raw materials used in a proposed glutethimide tablet

| Item | Function | Tests required |
|---|---|---|
| Glutethimide, USP XXI | Depressant | Appearance; solubility; identity; melting range; ultraviolet absorption; loss on drying; residue on ignition; assay |
| Tragacanth, NF XVI | Binder | Appearance; identity; karaya gum; microbial limits, arsenic, lead, heavy metals |
| Lactose, USP XXI | Diluent | Appearance; solubility; identity; specific rotation; residue on ignition; heavy metals; clarity and color of solution, microbial limits, pH, water, alcohol soluble residue |
| Talc, USP XXI | Glidant | Appearance; identity; loss on ignition; acid-soluble substances; water-soluble iron; reaction and soluble substances |
| Magnesium stearate, NF XVI | Lubricant | Appearance; solubility; identity; loss on drying; lead content; assay |
| Polyethyleneglycol 4000, NF XVI | Binder and lubricant | Appearance; solubility; viscosity; completeness and color of solution; pH, acidity; average molecular |

| | | |
|---|---|---|
| | | weight; residue on ignition; arsenic; heavy metals, limit of ethylene glycol and diethylene glycol, ethylene oxide average molecular weight |
| Alcohol, USP XXI | Solvent | Appearance; solubility; specific gravity; nonvolatile residue; water-insoluble substance; aldehydes and other foreign organic substances; methyl ketones, methanol and other alcohol content; fusel -oil constituents, amyl alcohol and nonvolatile, carbonizable substances |
| Purified water, USP XXI | Solvent | Appearance; reaction; heavy metals content, total solids; bacteriologic purity, oxidizable substance, carbon dioxide, calcium ammonia, sulfate, chloride, pH |

Procedures in the compendium apply to all manufacturers of a particular product, whereas the quality control procedures of the manufacturer are intended to apply specifically to his own product. The pharmaceutical industry frequently employs alternative methods, which may be more accurate, specific, sensitive, and economical than those in the compendium. The manufacturer is not required to employ procedures in the official compendium as long as the quality of his product ultimately meets the requirements in the compendium for identity, quality, potency, and purity. In the case of a legal action, the test methods in the compendium are the basis for determining compliance.

Container components should not interact physically or chemically with the product to alter its identity, purity, quality, or potency beyond the official allowances. Usually, the major component of the container is glass or plastic. The official compendium provides for several types of glass and several classes of plastic to be used with various pharmaceutical products. Examples of quality control profiles of three major packaging materials are shown in

Table 29.22. Specifications designed for containers are meaningful only if they have been selected on the basis of tests performed on the product in the container. The following features are to be considered before container specifications are set (1) physical changes of container upon prolonged contact with the product, (2) moisture and gaseous permeability of the container, (3) compatibility between the container and the product and (4) toxicity and safety.

**Table 29.22:** Quality control profile of various packaging materials USP XXI/NF XVI

| | Item | Function | Test required |
|---|---|---|---|
| 1. | High-density polyethylene | Containers for capsules and tablets | Multiple internal reflectance, thermalanalysis, light transmission, water vapor permeation, extractable substances, nonvolatile residue, heavy metals |
| 2. | Glass Highly resistant, borosilicate glass (Type I) | Parenteral use | Light transmission, chemical resistance—powdered glass test |
| | Treated soda-lime glass (Type II) | Acidic and neutral parenteral preparation | Light transmission, chemical resistance—water attack at 121° test |
| | Soda-lime glass (Type III) | Parenteral and nonparenteral preparation | Light transmission, chemical resistance—powdered glass test |
| | General purpose soda-lime glass | Nonparenteral articles, i.e. oral or topical use | Light transmission, chemical resistance—powdered glass test |
| 3. | Elastomeric closure for injection | Parenteral use | Biologic test-acute systemic toxicity, intracutaneous reactivity; physicochemical test—turbidity, |

| | |
|---|---|
| | reducing agents, heavy metals, pH change, total extractables |

Data compiled from USP XXI/NF XVI, United States Pharmacopeial Convention, Rockville, MD, 1980.

## *Testing Program and Method*

Total quality assurance certifies that each received lot of raw material or each manufactured batch of product meets the established quality standards. It provides for the authorization of the release of the approved raw materials for manufacturing, and the release of the manufactured product to the market, based on actual laboratory testing—physical, chemical, microbiologic, and at times, biologic.

Outlines of various quality control tests for different properties of the product are presented here to illustrate the scope of various laboratory testing.

1. *Physical and chemical tests:* Tests for appearance, color, odor, identity, optical rotation, specific gravity, pH, solubility, viscosity, disintegration time hardness, friability, average weight or volume per unit, weight or volume variation, content uniformity, dissolution profile, polymorphic form, particle size, moisture content, and assay for active ingredient(s), impurities, contaminants, or degradation products.
2. *Biologic and microbiologic tests:* Macrobiologic or microbiologic assays, and tests for potency, safety, toxicity, pyrogenicity, sterility, histamine, phenol coefficient, antiseptic activity, and antimicrobial preservative effectiveness tests.

Most therapeutic agents are substances of known chemical structure or composition and can be assayed by quantitative physicochemical means. The standard purity statement for the active ingredient in the dosage form usually permits a wider variation than that for the active ingredient itself. Through purity and identity tests, the quality of the drug alone is established, and its level of impurities restricted, as in the limiting tests for chloride, sulfate, and heavy metals. The assay measures the concentration of this previously accepted drug in the dosage form.

When a physicochemical assay method is not possible, a macrobiologic or microbiologic procedure is employed. Biologic testing of drugs may be

quantitative or qualitative in nature; it utilizes intact animals, animal preparations, isolated living tissues, or microorganisms. Biologic methods are employed in the following situations (1) when adequate chemical assay has not been devised for the drug substances, although its chemical structure has been established (e.g. insulin), (2) when chemical structure of the drug substance has not been fully elucidated (e.g., parathyroid hormone), (3) when the drug is composed of a complex mixture of substances of varying structure and activity (e.g., digitalis and posterior pituitary), (4) when it is impossible or impractical to isolate the drug from its interfering substances, although the drug itself can be analyzed chemically (e.g. isolation of vitamin D from certain irradiated oils), (5) when the biologic activity of the drug substance is not defined by the chemical assay (as when active and inactive isomers of methylphenidate cannot be differentiated by the chemical method) and (6) when specificity, sensitivity, or practicality dictates the use of biologic rather than chemical assay procedures.

The accuracy of biologic tests does not approach that which is expected with good chemical methods. Accuracy within ± 20% of the true value is good, and within ± 10% is excellent, for most bioassays. Consequently, it is useful and advisable to supplement the biologic tests with select physicochemical tests when possible. Some official quantitative and qualitative biologic tests are summarized in Tables 29.23 and 29.24 respectively. Some biologic tests require as long as a month for completion; others may take only a few hours. Therefore, bioassays are often expensive and inconvenient.

**Table 29.23:** Partial list of official quantitative biologic and microbiologic tests

| **Biologic tests** | |
|---|---|
| ***Drug and dosage form*** | ***Test animal(s)*** |
| Insulin | Rabbit |
| Digitalis and the related cardiac glycosides | Pigeon |
| Parathyroid | Dog |
| Posterior pituitary | Rat |
| Tubocurarine chloride | Rabbit |

| **Microbiologic tests** | |
|---|---|
| ***Drug and dosage form*** | ***Test organism(s)*** |
| | Microbiologic assay |
| Calcium pantothenate | *Lactobacillus plantarum* |
| Cyanocobalamin | *Lactobacillus leichmannii* |
| Penicillin | *Staphylococcus aureus* |
| Other antibiotics | (Varied according to the antibiotics) |
| Antimicrobial preservatives | Antimicrobial preservations—effectiveness *Candida albicans, Aspergillus niger, Escherichia coli, Pseudomonas aeruginosa, Staphylococcus aureus* |

**Table 29.24:** Partial list of official qualitative biologic tests

| **Products to be tested** | **Test** |
|---|---|
| Preparations of liver or stomach | Antianemia tests |
| Antiseptics, disinfectants, fungicides, germicides | Antibacterial tests |
| Preparations containing toxoids | Antigenic test |
| Water, USP | Bacteriologic purity |
| Gelatin | Bacterial content |
| Protein hydrolysate injection | Biologic adequacy test |
| Protein hydrolysate injection | Nonantigenicity test |
| Diagnostic diphtheria toxin, influenza virus vaccine, and smallpox vaccine | Potency tests |
| Parenteral products, radioactive pharmaceutical, transfusion and infusion | Pyrogen test/bacterial assemblies endotoxin test |
| Most antibiotics, transfusion and infusion assemblies | Safety tests |
| Parenteral and ophthalmic products, antibiotics, sutures, all surgical dressings, transfusion and infusion assemblies | Sterility tests |
| Chloramphenicol, streptomycin, | Test for depressor |

| | |
|---|---|
| chlorotetracycline, tetracycline, liver injection | substances |
| Adrenal cortex injection | Test for pressor substances |
| Suramin sodium and preparations containing toxoids | Toxicity tests |
| Pharmaceutical articles, from raw materials to finished products | Microbial limit tests |

To minimize the source of error resulting from animal variation during biologic tests, reference standards or pure drug substances and standard reference preparations are used, where possible, for comparison of their potency with the potency of unknown preparations to be tested. A reference standard is the specific active principle of the drug in its purest obtainable form. The principle of using tests in which reference materials are employed is based upon successive testing of the unknown and standard preparations on two groups of similar animals, as in the case of digitalis and tubocurarine, or on the same animal or organ, as in the case of posterior pituitary and epinephrine. The potency of the unknown can therefore be expressed as a percentage of the standard, although other methods of computation have been devised to improve the reliability of the results.

Increasing emphasis on microbiologic attributes of nonsterile products has generated additional responsibility in the quality control of raw materials, especially those derived from animal or botanical origin. In recognition of this increased concern regarding the possible contamination of nonsterile products with pathogenic or otherwise objectionable microorganisms, the USP XXI has included microbiologic quality control test procedures. These are designed to monitor nonsterile drug products for possible adulteration with microorganisms such as *Salmonella* species, *Escherichia coli*, certain species of *Pseudomonas*, and *Staphylococcus aureus*. Medicinal substances of natural and mineral origin are likely to be contaminated with bacteria, while synthetic medicinal substances tend to be bacteriologically clean in comparison. Nevertheless, solutions, suspensions, and semisolid dosage forms of these synthetic medicines tend to acquire bacterial contamination from excipients, manufacturing processes and environment, and containers. At the present time, there are many monographs in the USP XXI requiring freedom from one or more of the aforementioned organisms. Typical 

examples are given in Table 29.25.

**Table 29.25:** Typical pharmaceutical materials required by official compendia to ensure freedom from specific objectionable microorganisms

| Absence of *E. coli* | Absence of *Salmonella* | Absence of *E. coli* and *Salmonella* | Absence of *Salmonella* and *Pseudomonas* |
|---|---|---|---|
| **USP XXI** | | | |
| Alumina and magnesia oral suspension | Dehydrocholic acid | Gelatin | Chymotrypsin |
| Aluminum hydroxide gel | Digitalis capsules | | |
| | Pectin | Pancrelipase | |
| | Thyroid | Activated charcoal | |
| | Milk of magnesia | | |
| | Trypsin | | |
| | Pancreatin | | |
| **NFXVI** | | | |
| Absence of *E. coli* | Absence of *Salmonella* | Absence of *E. coli* and *Salmonella* | Absence of *Salmonella*, *E. coli*, *S. aureus*, and *P. aeruginosa* |
| | Acacia | Starch | Caramel |
| | Agar | Alginic acid | |

Data compiled from USP XXI/NF XVI, United States Pharmacopoeial Convention, Rockville, MD, 1980.

The microbial flora associated with various raw materials can vary considerably. Standard plate-counting procedures are familiar and popular methods. In addition, broth enrichment procedures are used to detect low levels of microorganisms. The following categories of raw materials are often contaminated with various microbial flora and should be thoroughly investigated: processed water, colors, dyes, pigments, talcs, starches, clays, fillers, natural gums, and thickening agents. It is important to consider the pharmaceutical process when a limit is established for the number of microorganisms per gram of raw material. Samples should be taken throughout the production cycle on a random basis to evaluate the microbiologic spectrum of the process. This serves as an indicator of sanitation and good manufacturing practices.

The final dosage form is often statistically sampled and tested for microbiologic attributes. Here again, the absence of pathogens is important, and the total count gives a measure of microbiologic normality. The hazard of microbial contamination is related to the intended area of use for the product. For parenteral products, the dictum is that any viable microorganism is a

health hazard. Even for orally administered products, the presence of *E. coli* and *Salmonella* species are always objectionable, and the presence of *Pseudomonas* species, *C. albicans*, Enterobacter species and mycotoxin-producing fungi are usually objectionable since these organisms have been implicated as agents of diseases in foods. Further examples of objectionable microorganism contamination are tabulated in Table 29.26.

**Table 29.26:** Partial list of objectionable organisms in pharmaceuticals

| | **Examples of objectionable organisms** | |
|---|---|---|
| **Intended area of product application** | **Always objectionable** | **Usually objectionable** |
| Oral products | *Escherichia coli* | *Pseudomonas* sp. |
| | *Salmonella* sp. | *Enterobacter* sp. |
| | | Enterotoxigenic *Staphylococcus aureus* Myocotoxin-producing fungi *Candida albicans* |
| Parenteral products | Any viable microorganism | |
| Ophthalmic preparations | *Pseudomonas aeruginosa* | Other *Pseudomonas* sp. *Staphylococcus aureus* |
| Genitourinary tract products | *Escherichia coli* | *Klebsiella* sp. |
| | *Proteus mirabilis, Serratia marcescens, Pseudomonas aeruginosa, Pseudomonas multivorans* | *Acinetobacter anitratus* |
| Products for surface wounds and damaged epithelium | *Pseudomonas aeruginosa Staphylococcus aureus Klebsiella* sp. *Serratia* | |

| | | |
|---|---|---|
| | *marcescens*, *Pseudomonas multivorans*, *Pseudomonas putida*, *Clostridia perfringens* | |
| Topical products | *Pseudomonas aeruginosa* *Staphylococcus aureus* *Klebsiella* sp. *Serratia marcescens* | *Pseudomonas multivorans* *Pseudomonas putida* *Clostridia perfringens* *Clostridia tetani*, *Clostridia novyi* |

Microbiologic testing of sterile products (injection and ophthalmic) for the antimicrobial efficacy of added antimicrobial agents, for sterility, and for pyrogenicity are discussed in detail in Chapters 22, “Sterilization” and 23, “Sterile Products”.

Traditionally, pyrogen testing is performed on rabbits, and observations are made for febrile response. A recent innovation in pyrogen testing is the use of an in vivo limulus amebocyte lysate (LAL) test, which is capable of detecting the more potent endotoxin pyrogens. Manufacturers of nonantibiotic injectable products may substitute the LAL test for the official rabbit pyrogen assay as an end product endotoxin test immediately upon a supplemental submission to the FDA, provided that the firm has validated the test for the particular drug product. On the other hand, antibiotic drug makers and manufacturers of biologicals who change manufacturing controls are required to submit to the agency for advance approval a full statement describing the proposed change supported by the validated data.

An important aspect of dosage form control is the safety or toxicity test that is performed on the finished dosage form to guard against adventitious adulterations of pharmaceuticals. Analytic assays of drugs may not detect impurities or errors unforeseen in formulation, whereas in vivo testing may qualitatively show a change in a predetermined response to a specific drug dosage.

The use of reference standards has been extended beyond biologic assays, so that today they are required for many pharmacopeial assays and tests. This reflects the extensive use of modern chromatography and spectrophotometry,

which require measurements relative to a reference standard to attain accurate and reproducible results. For example, a set of standards is provided for checking the reliability of apparatus used for melting point determinations. There are also pure specimens of steroids suitably diluted with an inert diluent for use in the chromatographic identification of steroids.

Reference standards are prepared and distributed by the USP. A similar program providing international standards is maintained by the World Health Organization, which is concerned mainly with standards for serums, vaccines, toxins, vitamins, and endocrine extracts. In general, the critical characteristics of the specimens selected as standards are determined independently in three or more laboratories.

The traditional and conventional gravimetric and volumetric assay methods are being supplemented by newer instrumental methods. Pharmaceutical analysis now includes such techniques as partition or absorption chromatography, gas-liquid chromatography, high pressure liquid chromatography, ultraviolet and infrared spectrophotometry, complexometry, chelatometry, nonaqueous titrimetry, fluorometry, polarography, differential scanning calorimetry, X-ray spectroscopy, nuclear magnetic resonance, autoradiography, thermogravimetric analysis, and mass spectroscopy. Thus, the analytic work required may range from the standardized elemental analysis of the compound to the highly specialized and sophisticated chemical or instrumental functional group determination.

Just as there are differences between samples and between replicate determinations on the same sample, there are also variations between technicians, instruments, and laboratories. These variations for the most part can be measured, and their significance determined. Such knowledge provides for the most accurate appraisal of the data and optimal confidence in the results. If the error in the analytic methods employed for product testing is not understood, the possibility of rejecting a finished product may be due to the test methods and not to inferior products.

Occasionally, outside laboratories may be used to augment the testing capacity and capability of the manufacturer. It is advisable that the facilities and personnel of the outside laboratories be evaluated before they are engaged. The quality control personnel of the manufacturer should evaluate and endorse the results submitted by the outside laboratories.

## *Quality of Analytic Methodologies*

The quality of a method has to be characterized, monitored, measured, and validated. The nature of the analytic methods may be physical, chemical, microbiologic, biologic, or a combination of these types. The quality of analysis is built in during its design stage, validated in its development stage, and confirmed in its utilization stage.

## *Parameters of Quality Analysis*

The selection of an analytic method may be based on one or more of the following quality criteria or parameters, which serve as the foundations of a quality analysis:

1. Specificity
2. Sensitivity (limit of detection)
3. Linearity
4. Precision
5. Accuracy
6. Ruggedness
7. System Suitability

These important quality parameters of an analytic method are briefly described in this section.

*Specificity* is an important quality criterion. Analysis of a component of a mixture may interfere with other components of the mixture. If this occurs, the analytic method is nonspecific for the component under investigation. With a specific method, the concentration of the component can be completely measured regardless of what other compounds are present in the sample. Ideally, for chromatographic analysis, a specific method should be capable of resolving from the peak of interest all other components, including impurities, contaminants, excipients, and degradation products. In general, baseline resolution from other peaks or spots is accepted as adequate resolution for good specificity. It is not necessary, however, to have clean baseline separation among components that are not to be quantified. Furthermore, the integrity of the peak should be verified by collecting the peak fraction and chromatographing it by another solvent system or chromatographic method. It can also be further verified by wavelength-ratio

techniques for direct comparisons to a standard.

*Sensitivity* pertains to the ease of detection. Analysts usually try to develop methods in which sensitivity has a constant value in a range as large as possible. Limit of detection gives the minimum concentration of a component that can be detected by the analytic method. As a rule, whenever a sample contains a compound in a very low concentration, the signal from the instrument will be small. Therefore, uncertainty exists as to whether the signal comes from noise produced by the instrument, from the method, or from the actual component to be measured. This gives rise to the term limit of detection. The analyst should determine the lowest detectable quantity of major component of interest at the most sensitive instrument settings. This detection limit is usually taken to be twice the signal-to-noise ratio. For determining the detection limit of minor components, the analyst must determine the smallest amount of contaminants, impurities, precursors, or degradation products quantifiable as weight percent of major component in the presence of this major component. This limit of detection is usually taken as ten times the signal-to-noise ratio; however, the minimum quantifiable levels of these substances should be no more than 50% of the limit set in the specifications for these substances.

*Linearity* of method gives the characteristic trend of such parameters as absorbance, peak height, peak area, or response ratio as a function of concentration of component to be measured. At least five different concentrations of a standard solution should be employed and spanned 80 to 120% or even 50 to 150% of the expected working concentration range. By plotting the concentration versus response, the linearity of the observed data points can be visualized. The test of linearity can be accomplished by fitting the data points to a curve of the form:

$$y = mx^n + b$$

For perfect linearity, $n = 1$. Once the concentration—response equation is known, one can calculate the maximum error to be expected from deviations in linearity and the possible error to be anticipated from use of a single point standard within the acceptable assay range of the method.

*Precision* is a quality criterion referred to the reproducibility of measurement within a set number of independent replicate measurements of the same property. Thus, it refers to the dispersion of a set about its central value and is generally expressed as the standard deviation of a series of

measurements obtained from one sample. Usually, the precision of a method is established during the development stage by the multiple analyses of samples judged to be typical of the material that is to be analyzed. These analyses usually do not account for any additional sources of variation such as day-to-day fluctuation, laboratory-to-laboratory variation, small modification in technique, varying skill of analysts, undetected operational or instrumental factors, and other unexpected systematic errors.

Methods used to estimate the precision within a batch suffer from certain drawbacks. The most straightforward consists in the replicate analysis of a few selected samples and the subsequent calculation of the standard deviation. Unless random sampling is used, the samples selected may not be representative of the batch. In any analytic system, the precision of determination varies over the concentration range of the samples. This fact must be taken into account if the concentration range is wide. In a system in which the substance determined occurs in the same matrix in all the samples, the standard deviation of measurement increases with concentration, usually as a linear function. If all the duplicate determinations were made within a batch of analyses, the method would give an estimate only of the within-batch precision rather than of the overall characteristics of the analytic system. If the same procedure were used on many successive batches, a valid estimate of the overall variance would be obtained, consisting of the sum of the within—batch variances and the variance due to any systematic differences between batches (Table 29.27).

In general, the quality control laboratory determines the precision of the method on six replicates of a representative composite sample containing 18 to 24 times the amount of drug needed for one assay. A separate determination of the precision of the system is made, considering only the error attributable to the operating system and not the error attributable to sample handling and preparation. The measure of system precision is performed by repeatedly analyzing aliquots of a single standard solution, recording responses, and calculating the relative standard deviation of the response.

*Accuracy* is defined as the closeness of a measured value to its true value. It normally refers to the difference between the mean of the set of results and the value accepted as the true or correct value for the quantity measured. As a rule, results of analysis of the unknown are compared with the results

obtained from the analysis of standards or reference materials. The analyst should prepare six samples of drug in matrix spanning 80 to 120% or even 50 to 150% of the expected content, assaying each of those synthetic samples. The acceptance criterion in the accuracy test is expressed in terms of the standard deviation of the method as determined in the precision test, since the deviation from theory depends on the error inherent in the method itself. As acceptance criteria, recovery of drug expressed as percentage of theory must be ± 4s of the theoretic value at all levels where s is the relative standard deviation obtained in the precision test. The range of 4s units is intended, and preferred, to cover the additional error possibly introduced in preparing the synthetic samples for accuracy testing. Table 29.27 represents accuracy data of moxifloxacin in rabbit aqueous humor by a ultra high-pressure liquid chromatography.

**Table 29.27:** Precision and accuracy data of moxifloxacin in rabbit aqueous humor by a UHPLC method

| Nominal amount (ng/ml) | Amount found (ng/ml) | Precision | | | Accuracy (%) |
|---|---|---|---|---|---|
| | | SD | SE | RSD (%) | |
| *Intra-day* | | | | | |
| 10 | 9.73 | 0.09 | 0.05 | 0.92 | 97.30 |
| 100 | 98.57 | 1.08 | 0.62 | 1.09 | 98.57 |
| 1000 | 998.29 | 0.68 | 0.39 | 0.07 | 99.83 |
| *Inter-day* | | | | | |
| 10 | 9.83 | 0.10 | 0.06 | 1.02 | 98.30 |
| 100 | 98.02 | 1.69 | 0.98 | 1.72 | 98.02 |
| 1000 | 999.86 | 1.66 | 0.96 | 0.17 | 99.99 |

From Jain GK, et al., J Pharm Biomed Anal. 2010 May 1;52(1):110–3. Reproduced with permission of the copyright owner.

*Ruggedness* tests describes the influence of small but reasonable alterations in the procedures of the quality of analysis.

Examples of these minor variations are source and age of reagents, concentration and stability of solutions and reagents, heating rate, thermometer errors, column temperature, humidity, voltage, fluctuation, variations of column to column, plate to plate, analyst to analyst and instrument to instrument, and many others. Eight measurements suffice to investigate seven variables when the appropriate experimental design is employed. The various types of interlaboratory checks should be carried out to ensure that the analyst who developed the method is not the only one who

can obtain satisfactory results from the procedure, and that all details are written into the testing directions and are not inadvertently omitted. With the widespread use of automatic injectors for HPLC systems, it is necessary to check and validate the length of time for which prepared solutions or reagents are stable. The stability of these solutions should be checked over a period of 12 to 24 hours if an automated method of running samples overnight is practiced in the laboratory.

*System-suitability* tests help to answer the question, "How good and reliable is the performance of a given analytic system on a given day?" FDA and the compendia have recommended system suitability tests for inclusion in all HPLC procedures. System, in this context, means all components of the analysts, hardware, solvents, and electronics considered together. System-suitability tests are composed of a system's precision measurement and a system's powers of resolution measurement to check the performance of the analytic system on a given day. System-suitability testing differs from method validation testing. Validation of analytic method is generally initiated at the method development stage and is finalized by demonstrating that the method is scientifically sound and technically adequate for a particular drug. Therefore, validation is often done only once. System-suitability testing should be on a continuing basis, however.

The measurement of system precision is most easily made by employing replicate aliquots of the same solution. Signal responses such as peak height, peak area, and response ration derived from these aliquots are determined, and the relative standard deviation is calculated to indicate the system's precision as compared with the historical data. The second part of each system-suitability test is the system's powers of resolution. The most useful measurement of such power is the calculation of the resolution between two closely eluting chromatographic peaks for which the resolution is considered critical. Normally, the separation between the peak of interest and a second peak eluting close to it is chosen as the resolution factor to be calculated.

Although system-suitability tests are recommended for all HPLC methods, it is suggested that they can be used in some manner for all analytic methods adopted for drug products.

With the background knowledge of specificity, sensitivity, linearity, precision, accuracy, and ruggedness of an analytic method, it is relatively easy to derive the confidence and reliability of the analytic data obtained with

the method. With the use of a formal validation procedure and a system-suitability test, each new method is sure to meet the same performance standards, minimizing the problems encountered in daily routine analysis and in interlaboratory method transfer or method change.

To select an appropriate method, the analyst should have a thorough knowledge of the physicochemical properties of a drug, degradation products, degradation mechanisms, and degradation reaction rates. One can then develop a specific method suitable for monitoring the stability of an active ingredient or formulation. The methodology used for kinetic studies (solid state or solution) can generally be considered scientifically suitable for monitoring decomposition.

One of the major characteristics for a quality analytic method for pharmaceuticals is its ability to determine distinctively the parent compound from the degradation products. The current trends in stability-indicating methods are classified as electrometric methods, solvent extraction methods, spectro-photometric methods, direct chromatographic methods and derivatization chromatographic methods.

## *Automated Continuous System for Assay Procedures*

Automation usually enhances the quality, quantity, and efficiency of an operation. Its introduction into the analytic laboratory has dramatically changed the traditional look, capability, precision, and acceptability of most of our conventional analytic disciplines. The use of automated instrumentation of pharmaceutical analysis, data handling, and data storage is certainly on the rise. Within the last decade, and especially during the last few years, great strides have been made in a diversity of instrumental approaches to automated chemical, microbiologic, enzymatic, and other assays. Since automated continuous testing has great potential use for routine testing, it is not surprising that some pharmaceutical manufacturers are initiating automated control methods that are capable of sampling, analyzing, and accepting or rejecting. The use of these methods permits the analysts to cope with the increasing number of samples required to ensure product quality. Automated continuous assay procedures enhance the reliability of data and provide immediate feedback on process control with tremendous savings in time.

In the quality control laboratory, where a sufficiently large number of

similar dosage forms or dosage units must be subjected routinely to the same type of examination, automated methods of analysis provide for far more efficient and precise testing than manual methods. Such auto mated methods have been found especially useful in testing the content uniformity of tablets and capsules and in facilitating methods requiring precisely controlled experimental conditions. Many manufacturing establishments, as well as the laboratories of regulatory agencies, have found it convenient to utilize automated methods as alternatives to pharmacopoeial methods.

As a general practice, before an automated method for testing an article is adopted as an alternative, it is advisable to ascertain that the results obtained by the automated method are equivalent in accuracy and precision to those obtained by the prescribed pharmacopoeial method. It is necessary to monitor the performance of the automated analytic system continually, by assaying standard preparations of known composition that have been frequently interspersed among the test preparations.

Many of the manual methods given in the pharmacopoeia can be adapted for use in semiautomated or fully automated equipment, incorporating either discrete analyzers or continuous flow systems and operating under a variety of conditions. On the other hand, an analytic scheme devised for a particular automated system may not be readily transposable for use either with a manual procedure or with other types of automated equipment.

Compendial methods for testing drugs are based largely on ultraviolet absorption measurements. After extraction of dilution, the sample must exhibit maxima and minima at the same wavelengths as those of a similar solution of the reference standard, and the respective wavelength of maximum absorption should not differ significantly. Based on this principle, the Auto-analyzer, an automated continuous system for repetitive spectral scanning, was developed. This equipment is capable of accepting tablet, capsule, powder, solution, and suspension samples for analysis as required. The Autoanalyzer is, in fact, a mechanical chemist that can automatically perform designated tests at a much greater speed than a trained analyst. It has been utilized to perform many chemical microbiologic, biologic, and clinical tests, as well as to implement production in-process control.

Besides the autoanalyzer produced by technicon company, there are other automated systems for drug analysis, such as Titralyzer from Fischer Scientific Company. Recent advances in automation and computerization

have improved the efficiency and economics of laboratory performance. For example, numerous companies have developed dedicated gas-liquid or high efficiency liquid chromatographic systems that permit on-line sample analysis and data handling. One such system utilizes an automatic sample injector, which is time-sequenced and connected to a gas-liquid chromatograph that has been preset for a specific analysis. The signal from the gas-liquid chromatographic detector is digitized, and the computer automatically prints out such parameters as retention time, area under the curve, and concentration of the drug. Such automated data processing provides multiple analysis with minimum supervision and with a savings in time. Dedicated gas-liquid chromatographic systems are offered by such companies as Hewlett-Packard and Waters Associates, Inc. These and other suitable automated procedures have been developed for the analysis of phenothiazine derivatives, erythromycin, tolbutamide, vitamins A and B6, amitriptyline hydrochloride, and various steroid dosage forms.

As previously discussed, HPLC is one of the most important analytic techniques to be given various degrees of automation in recent years. Its proven versatility and fast-growing ability to supply fundamental information and quality control data have stimulated the development of a completely automated HPLC system. In addition, the great advances in minicomputer science and microprocessor technology, such as Technicon Solid Prep Sample II and Peristaltic Pump HI, have provided the extra dimension needed to automate fully most of the hardware components and flow-of-operation and data-management processes. The system has proved to be successful for the routine analyses of fat-soluble vitamins in more than nine different multivitamin formulations. The productivity has been at least three times greater than the conventional manual procedures. The continuous flow analyses with automated HPLC systems are especially useful in content uniformity analyses, which require large volumes of samples. When the microprocessor is properly programmed, the system can be easily adapted for reliable, around-the-clock, unattended operation in each specific application.

Microbiologic assays have always been time-consuming and tedious. Following the successful assay in fermentation media of streptomycin and penicillin (the first two antibiotics whose assay was automated), a variety of automated systems have been devised and applied successfully in turbidimetric and respirometric assays. For example, plates used in the microbiologic assay of several antibiotics are being prepared and read

automatically, with the data recorded magnetically and the results analyzed and recorded by computer.

Among the requirements in the current USP/NF is a content uniformity test for tablets and capsules containing 50 mg or less of active ingredient. In the event that general strength dosage forms of the active ingredient are available below and above 50 mg per tablet or capsule, then all strengths require a content uniformity test. It is only through the use of automated procedures that the analytic burden is manageable.

One may expound any number of good reasons to automate laboratory procedures. Automated systems may enable a laboratory to provide new services more quickly, with increased reliability, in greater abundance, or at lower cost. Any of these benefits may justify consideration of an automated instrument. In any case, the total economic effect must be considered.

## *Associated Activities*

## *Product Identification Systems*

To ensure the quality of a product, the unit doses are (1) individually packaged and labeled and (2) imprinted with an assigned code for the purpose of identification. The first procedure is usually termed the unit dose packaging concept. Single unit doses for medication—tablets, capsules, or parenterals—are individually sealed in packages that have been imprinted with complete identifying data. Thus, the danger of error is substantially reduced.

The second procedure in identifying the products is called the unit code system. Various techniques have been employed for this method. The codes vary in complexity, for example, a system adopts alphabetical and numerical combinations. A drug product code composed of nine characters to be used by pharmaceutical manufacturers for inclusion in the National Drug Code Directory has been developed and has been used by the FDA to establish a uniform code system for dosage forms. The nine-character code would identify the labeler, the dosage form, the strength of the product, and the number of product units in the package. In the case of tablets or capsules, those parts of the code identifying the drug and its strength would be used on each unit of the dosage form. Interestingly, any product not identified would be termed misbranded unless compliance were impractical and the FDA

granted an exemption.

## *Adulteration, Misbranding and Counterfeiting*

An adulterated, misbranded, or counterfeited drug is a fraud to the public. Such products can seriously endanger the health and safety of the person taking the medication, since there is no guarantee that the ingredients are safe, of highest quality, or of labeled potency. A situation of this nature may mislead the physician because his patient's response may differ from the response expected.

A drug product may be deemed adulterated because of various defects or shortcomings, as discussed in Section 501 of the Federal FDC Act. According to the law, a drug is considered adulterated if it does not meet the quality and purity characteristics it is represented to possess.

According to Section 502 of the Federal FDC Act, the major reason for considering a drug to be misbranded is the mislabeling of the product. To safeguard the manufacturer from marketing adulterated or misbranded products, the establishment of a total quality control system concerned with the dosage form, package, and labeling prior to release for distribution is essential.

## *Maintenance, Storage, and Retrieval of Records*

The proper control of records (master formula, batch production, and packaging records) in manufacturing operations has been discussed in a previous section of this chapter. Suitable maintenance storage and retrieval of records for dosage form control are mandatory in assuring product quality.

Product distribution records should be organized systematically for each manufactured product. Complete records of the distribution of each batch of product must be maintained in a manner that would facilitate its recall if necessary. Such records should include the batch or control numbers identifying the batch of product distributed, the date and quantity shipped, and the name and address of the consignee. Such records should be retained for at least two years after the batch of product has been distributed or at least one year after the drug's expiration date, whichever is longer.

For the vast amounts of records maintained in a quality assurance program, the storage and retrieval of these records when needed can be time-

consuming if done manually. The application of computer systems in the storage and retrieval of data in this area is inescapable.

In general, records of data converted to digitized form are stored on computer tapes, cards, or discs, or in case storage. The choice of storage medium depends on the volume, pattern of use, type of data, and cost factors. Case memory is the fastest, most expensive and most accessible type of memory. Disc storage is available for large volumes of data, and its almost random access features are particularly useful. Magnetic tape is compact and provides relatively cheap storage. The main disadvantage of tape is that if the data to be retrieved does not occur in the same sequence as that in which it is stored on the tape, the whole tape must be searched for the requested information.

The importance of batch or control numbers must be re-emphasized. The establishment of these identification numbers for the products manufactured is the direct key by which the entire quality assurance program in production, control, storage, and distribution can be readily unlocked. As an indirect assurance of the product quality to the consumer, this identification number not only opens the correct file without delay should the need arise, but also facilitates recall of that particular product from the market should it become necessary.

***Complaints:*** Suitable systems should be provided to investigate and follow up the complaints regarding deteriorated, adulterated, misbranded, or counterfeited products. Often, useful information can be generated from detailed investigation of the reason for complaint; the investigation may be physical, chemical, or biologic in nature. Complaints should be recorded and carefully evaluated by competent and responsible personnel, and the complaint files should indicate the action taken.

***Return of goods:*** Inevitably, a certain fraction of distributed product is returned for one reason or another. Returned products should be properly handled by capable individuals and correctly recorded in a manner that prevents confusion and the possibility of errors. On receipt, each should be listed on an appropriate form, giving the name and address of the sender, the batch or control numbers of the returned product, the reason for returning the product, and the estimated condition of the product. The product should then be analyzed by physical, chemical, or biologic methods whenever possible and stored in an orderly manner pending issuance of credit and a decision for

disposal.

***Recall procedures:*** The need to establish an effective drug recall system by the pharmaceutical manufacturers is evident in as much as hundreds of recalls that are made annually. In general, product recalls may be initiated by the manufacturer, by the regulatory agency, by government seizure through the courts, or by the manufacturer at the request of the regulatory agency. The last category applies to the majority of recalls. If the products to be recalled have been distributed only to the warehouses of the manufacturer or the distributor, the hazard to the public health may not be serious. The recall can be effected rapidly and easily. When recall involves the local pharmacist, the physician, or the public, however, the problem becomes complicated, and complete recall may be hampered. It then becomes necessary to enlist the aid of public communication systems to reach the consumer. An effective system for product recall requires, in addition to the efforts of pharmaceutical manufacturers, a careful system of recordkeeping at all levels of product distribution, from manufacturer to ultimate consumer.

***Adverse effects:*** Quality of marketed products can be further ensured if competent personnel are assigned to evaluate the significance and severity of each side reaction or adverse effect reported by the physician who prescribes the medicinals, the drug surveillance group of the hospital, published scientific results and even the regulatory agency. An effective product information system should be established to evaluate and accumulate these reports on the safety, potency, adverse reactions, clinical side effects, and drug abuse problems of the marketed product and to prepare appropriate replies, and reports in this area.

## Summary

The professional, social, and legal responsibilities that rest with the pharmaceutical manufacturers for the assurance of product quality are tremendous. It is only through well-organized, adequately staffed, and accurately performed process and dosage form control before, during, and after production that adequate quality assurance of the product can be achieved. It should be realized that no amount of dosage form testing and control can maintain and assure product quality unless good manufacturing practices are implemented systematically and process control is practiced vigorously. Product quality must be built into, and not merely tested in, the product.

The pharmaceutical manufacturer assumes the major responsibility for the quality of his products. The manufacturer is in a position (1) to control the sources of product quality variation, namely materials, methods, machines, and men, (2) to ensure the correct and most appropriate manufacturing and packaging practices, (3) to assure that the testing results are in compliance with the standards or specifications and (4) to assure product stability and to perform other activities related to product quality through a well-organized total quality assurance system.

For the total quality assurance system to function effectively, certain basic operational rules should be established and should always prevail. First, control decisions must be based solely on considerations of product quality. Second, the operation must adhere rigidly to the established standards or specifications as determined by systematic inspection, sampling, and testing, and should constantly strive for improving the levels of the current standards or specifications. Third, the facilities, funds for personnel, and environment necessary for personnel to perform their responsibilities effectively should be adequately provided. Last but not least, the control decisions should be independent administratively, and they must not yield to, or be overruled by, production or marketing personnel under any circumstances. Because the control decision can involve the health of the consumer and the reputation of the manufacturer, the climate necessary for making judicious decisions is essential. In times of major disagreement, the control decisions should be subjected to review only at the highest level of management.

# 30:
# Drug Regulatory Affairs

Humans have always made efforts for staying fit and healthy. But the complex physiological mechanisms of the human body, negative environmental factors (could be radiations, microorganisms, temperature and humidity, etc.) and modern lifestyle, inevitably manage to defeat the efforts. Drugs which are used for diagnosis, cure, mitigation, treatment or prevention of diseases are amongst the most significant way to get the human body back to a fit and healthy state.

Drugs and medical devices are among the most stringently regulated products in the developed world. The fundamental purpose of regulation is the protection of public health. Regulatory authorities across the world have been working their way to develop regulations which can ensure that only safe and effective products, of high quality and purity, reach into the markets for commercial distribution. Most of the landmark advances in the regulatory development were triggered by adverse events. A keen look into the history of different regulatory authorities will clarify the fact that the current regulations are actually based on the distilled wisdom of bitter past experiences.

US and EU are two parts of the developed world where drugs are the most stringently regulated. The following text describes the history, organization and responsibilities of the regulatory authorities of US and EU region. The chapter will discuss the drug approval and development process, mechanisms used to regulate manufacturing and marketing, as well as various violation and enforcement schemes adopted by US and EU regulatory authorities. Methods and tools for protecting industrial property and product liability are also discussed at the end of the chapter.

## DEFINITIONS

The terms most commonly used in this chapter are defined as they are in the Act:

## Drug

Sec. 201(g) (1): The term “drug” means (A) articles recognized in the official United States Pharmacopoeia, official Homeopathic Pharmacopoeia of the United States, or official National Formulary, or any supplement to any of them, (B) articles intended for use in the diagnosis, cure, mitigation, treatment, or prevention of disease in man or other animals and (C) articles (other than food) intended to affect the structure or any function of the body of man or other animals; and (D) articles intended for use as a component of any article specified in clause (A), (B) or (C); but does not include devices or their components, parts, or accessories.

## Label

Sec. 201(k): The term “label” means a display of written, printed, or graphic matter upon the immediate container of any article; and a requirement made by or under authority of this Act that any word, statement, or other information appearing on the label shall not be considered to be complied with unless such word, statement, or other information also appears on the outside container or wrapper, if any there be, of the retail package of such article, or is easily legible through the outside container or wrapper.

## Labeling

Sec. 201(m): The term “labeling” means all labels and other written, printed, or graphic matter (1) upon any article or any of its containers or wrappers or (2) accompanying such article.

## New Drug

Sec. 201(p): The term “new drug” means (1) any drug (except a new animal drug or an animal feed bearing or containing a new animal drug) the composition of which is such among experts qualified by scientific training and experience to evaluate the safety and effectiveness of drugs, as safe and effective for use under the conditions prescribed, recommended, or suggested in the labeling thereof, except that such a drug not so recognized shall not be deemed to be a “new drug” if at any time prior to the enactment of this Act it was subject to the Food and Drugs Act of June 30, 1906, as amended, and if at such time its labeling contained the same representations concerning the conditions of its use or (2) any drug (except a new animal drug or an animal feed bearing or containing a new animal drug) the composition of which is such that such drug, as a result of investigations to determine its safety and effectiveness for use under such conditions, has become so recognized, but which has not, otherwise than in such investigations, been used to a material extent or for a material time under such conditions.

## New Animal Drug

Sec. 201(w): The term “new animal drug” means any drug intended for use for animals other than man, including any drug intended for use in animal feed but not including such animal feed (1) the composition of which is such that such drug is not generally recognized, among experts qualified by scientific training and experience to evaluate the safety and effectiveness of animal drugs, as safe and effective for use under the conditions prescribed, recommended, or suggested in the labeling thereof; except that such a drug not so recognized shall not be deemed to be a “new animal drug” if at any time prior to June 25, 1938, it was subject to the Food and Drug Act of June 30, 1906, as amended, and if at such time its labeling contained the same representations concerning the conditions of its use or (2) the composition of which is such that such drug, as a result of investigations to determine its safety and effectiveness for use under such conditions, has become so recognized but which has not, otherwise than in such investigations, been used to a material extent or for a material time under such conditions or (3) which drug is composed wholly or partly of any kind of penicillin, streptomycin, chlor-tetracycline, chloramphenicol, or bacitracin, or any derivation thereof, except when there is in effect a published order of the Secretary declaring such drug not to be a new animal drug on the grounds that (A) the requirement of certification of batches of such drug, as provided for in section 512(n), is not necessary to insure that the objectives specified in paragraph (3) thereof are achieved and (B) that neither subparagraph (1) nor (2) of this paragraph (w) applies to such drug.

## Animal Feed

Sec. 201(x): The term “animal feed,” as used in paragraph (w) of this section, in section 512, and in provisions of this Act referring to such paragraph or section, means an article which is intended for use for food for animals other than man and which is intended for use as a substantial source of nutrients in the diet of the animal, and is not limited to a mixture intended to be the sole ration of the animal.

## Orphan Drugs

There are many (around 6000) rare disease which affect a small number of patient population. The small patient population of these rare diseases often does not present a sufficiently viable market for drug sponsors to recover the high costs of therapeutic research and development, much less to expect profit. Food and Drug Administration (FDA) amended the incentives established in the Orphan Drug Act (ODA) (1983) to encourage the development of potentially promising drugs. The main incentives include (i) regulatory assistance from the FDA in the form of written recommendations for non-clinical and clinical investigations for marketing approval purposes, (ii) a tax credit covering 50% of the clinical drug testing expenses, (iii) federal funding of grants and contracts to help defray such expenses and (iv) a seven-year exclusive marketing rights to the sponsor of the innovator drug after approval. To qualify for these benefits, the sponsor would need to request an orphan drug designation from the FDA. As originally enacted, the FDA would grant such a request if the disease or condition in question occurred so infrequently (less than 200,000 persons) in the United States that there was no reasonable expectation that the costs of developing and making the drug available would be recovered from its sales in the United States. As such, all sponsors were required to provide to the FDA detailed financial information to demonstrate the anticipated lack of profitability, regardless of how rare a drug's target population would be.

The Act initially allowed only designated orphan drugs "for which a United States Letter of Patent may not be issued" to qualify for the seven-year marketing exclusivity. In 1985, the Act was amended to make all drugs eligible for this incentive, regardless of their patentability. In addition, the amendment clarified that antibiotics would also be eligible for orphan drug incentives.

# US LEGISLATION

In the United States, all foods, drugs, cosmetics, and medical devices for both humans and animals are regulated by the Federal Food, Drug, and Cosmetic Act (FDCA), which in turn established the FDA. The FDA and all of its regulations were created by the government in response to the pressing need to address the health-related safety of public. FDA is the oldest comprehensive consumer protection agency in the US federal government. The USFDA is a scientific, regulatory, and public health agency that oversees items accounting for 25 cents of every dollar spent by the consumers in US. Its jurisdiction encompasses most food products (other than meat and poultry), human and animal drugs, therapeutic agents of biological origin, medical devices, radiation-emitting products for consumer, medical, and occupational use, cosmetics and animal feed. FDA monitors the manufacture, import, transport, storage, and sale of about $1 trillion worth of products annually at a cost to taxpayers of about $3 per person.

## FDA ROLE AND ORGANIZATION

FDA is organized into a number of offices and centers headed by a commissioner who is appointed by the President with consent of the senate. These offices and centers are assigned with different responsibilities to accomplish the objective of FDA. There are five product centers:

- Center for Drug Evaluation and Research (CDER)
- Center for Biologics Evaluation and Research (CBER)
- Center for Devices and Radiological Health (CDRH)
- Center for Food Safety and Applied Nutrition (CFSAN)
- Center for Veterinary Medicine (CVM)

These centers are responsible for approving and monitoring of market drugs/biologics/devices/foods/nutraceuticals/cosmetics, and inspecting the manufacturing facilities which fall under their respective authorities.

Following is the list of offices along with their main functions and responsibilities:

## Office of Regulatory Affairs (ORA)

It is the lead office for all regulatory activities as well as providing FDA leadership on imports, inspections, and enforcement policies. ORA supports the five FDA Product Centers by inspecting regulated products and manufacturers, conducting sample analysis on regulated products, and reviewing imported products offered for entry into the United States. ORA also develops FDA-wide policy on compliance and enforcement, and executes FDA’s Import Strategy and Food Protection Plans.

## Office of Generic Drugs (OGD)

OGD is a part of FDA’s Center for Drug Evaluation and Research. It is dedicated to approving safe, effective, high-quality and bioequivalent generic drug products for use by consumers.

## Office of Orphan Products (OOPD)

OOPD is dedicated to promoting the development of products that demonstrate promise for the diagnosis and/or treatment of rare diseases or conditions since 1982. OOPD interacts with the medical and research communities, professional organizations, academia, and the pharmaceutical industry, as well as rare disease groups. The OOPD administers the major provisions of the ODA which provide incentives for sponsors to develop products for rare diseases. The ODA has been very successful—more than 200 drugs and biological products for rare diseases have been brought to the market since 1983. In contrast, the decade prior to 1983 saw fewer than ten such products coming into the market. In addition, the OOPD administers the Orphan Products Grants Program which provides funding for clinical research in rare diseases.

## Office of Combination Products (OCP)

### *Combination Products*

Combination products are defined in 21 CFR 3.2(e). The term combination product includes:

1. A product comprising of two or more regulated components, i.e. drug/device, biologic/device, drug/biologic, or drug/device/biologic, that are physically, chemically, or otherwise combined or mixed and produced as a single entity.
2. Two or more separate products packaged together in a single package or as a unit and comprising of drug and device products, device and biological products, or biological and drug products.
3. A drug, device, or biological product packaged separately that according to its investigational plan or proposed labeling is intended for use only with an approved individually specified drug, device, or biological product where both are required to achieve the intended use, indication, or effect and wherein upon approval of the proposed product, the labeling of the approved product would need to be changed, e.g. to reflect a change in the intended use, dosage form, strength, route of administration, or a significant change in the dose.
4. Any investigational drug, device, or biological product packaged separately that according to its proposed labeling is for use only with another individually specified investigational drug, device, or biological product and where both are required to achieve the intended use, indication, or effect.

   The FDA OCP was established on December 24, 2002.

*The specific roles of OCP are*

- To serve as a focal point for the issue of combination product for agency reviewers and industry.
- To develop guidance and regulations to clarify the regulation of combination products.
- To assign an FDA center to have primary jurisdiction for review of both combination and single entity (i.e. non-combination) products where the

jurisdiction is unclear or in dispute.

- To ensure timely and effective pre-market review of combination products by overseeing the timeliness of and coordinating reviews involving more than one agency center.
- To ensure consistency and appropriateness of post-market regulation of combination products.
- To resolve disputes regarding the timeliness of pre-market review of combination products.
- To update agreements, guidance documents, or practices specific to the assignment of combination products.
- To submit annual reports to Congress regarding the office’s activities and impacts.

## Office of Ombudsman

The FDA Office of Ombudsman is the agency's focal point for addressing complaints and assisting in resolving disputes between companies or individuals and FDA offices concerning fair and even-handed application of FDA policy and procedures. It serves as a neutral and independent resource for members of FDA-regulated industries when they experience problems with the regulatory processes that have not been resolved at the center or district level. It works to resolve externally and internally generated problems for which there are no legal or established means of redressing by finding approaches that are acceptable to both the affected party and to the agency. The organizational chart of the FDA is presented in Fig. 30.1.

Along with the FDA, other agencies which have responsibilities for regulating pharmaceutical products are:

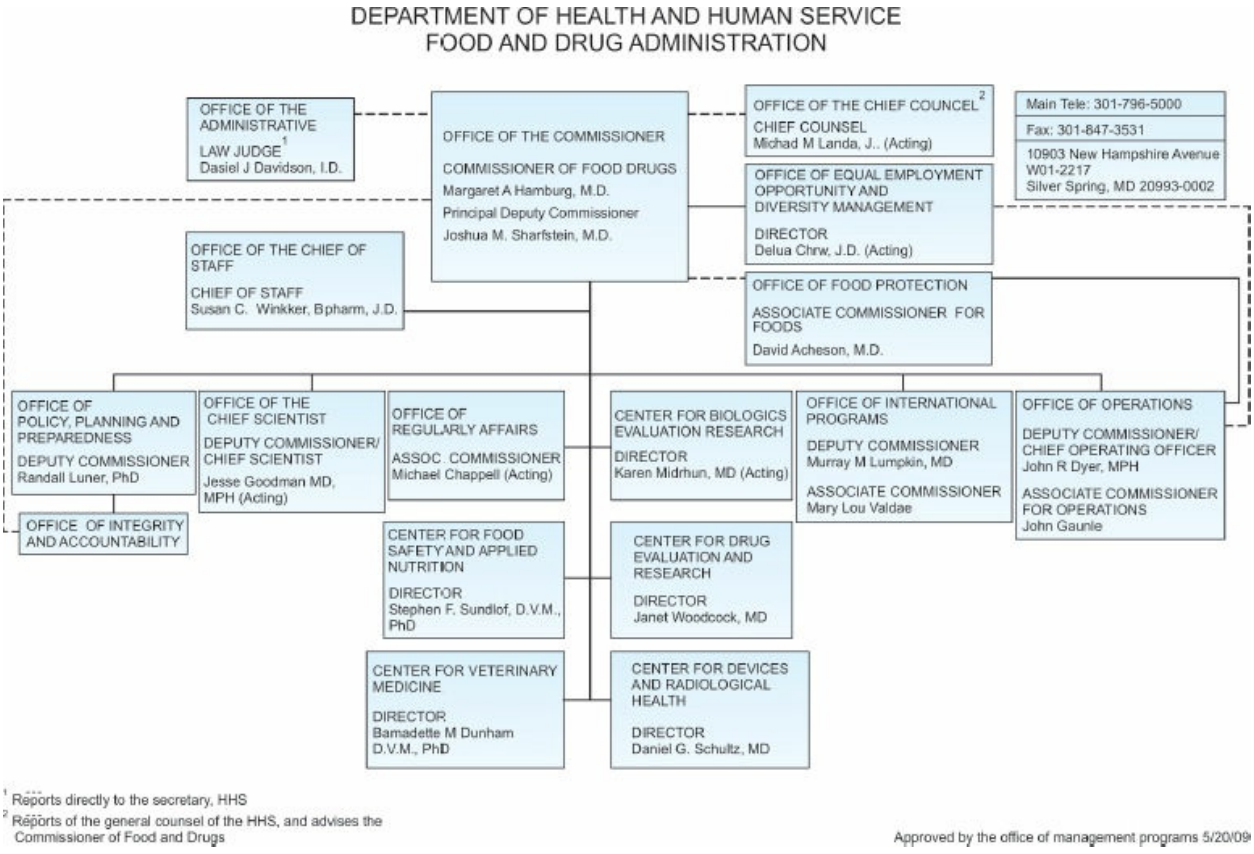



**Fig. 30.1:** Organizational chart of FDA

## Federal Trade Commission (FTC)

It has authority over business practices in general, such as deceptive and anti-competitive practices, i.e. false advertising of drugs. In addition, FTC regulates the advertising of OTC drugs, medical devices and cosmetics.

Consumer Product Safety Commission (CPSC). It regulates hazardous substances and containers of poisons and other harmful agents.

## Environmental Protection Agency (EPA)

It regulates pesticides used in agriculture and FDA-regulated products. It basically monitors and ensures that the use of pharmaceuticals does not create unnecessary and excessive hazardous burdens on the environment such as water, air and land.

## Occupational Safety and Health Administration (OSHA)

It regulates the working environment of employees who may use FDA-regulated commodities, i.e. syringes, chemotherapeutics, and chemical reagents.

## Drug Enforcement Administration (DEA)

It enforces the federal Controlled Substances Act (CSA) and is charged with controlling and monitoring the flow of licit and illicit controlled substances.

There are various state and local drug control agencies also which establish their own regulations and procedures for manufacturing, research, and development of pharmaceuticals.

## TYPES OF APPLICATIONS

Following are the types of drug approval applications that are submitted to the FDA:

- Investigational New Drug Application
- New Drug Application
- Abbreviated New Drug Application
- Biologic License Application
- Over-the-counter (OTC) drug application.

## Investigational New Drug (IND)

Current Federal law requires that a drug be the subject of an approved marketing application before it is transported or distributed across state lines. *An IND, or investigational new drug application, is a submission to the US Food and Drug Administration (FDA) requesting permission to initiate a clinical study of a new drug product in the United States*. Because a sponsor will probably want to ship the investigational drug to clinical investigators in many states, it must seek an exemption from that legal requirement. The IND is the means through which the sponsor technically obtains this exemption from the FDA.

During a new drug's early preclinical development, the sponsor's primary goal is to determine if the product is reasonably safe for initial use in humans and if the compound exhibits pharmacological activity that justifies commercial development. When a product is identified as a viable candidate for further development, the sponsor then focuses on collecting the data and information necessary to establish that the product will not expose humans to unreasonable risks when used in limited, early-stage clinical studies.

FDA's role in the development of a new drug begins when a drug's sponsor (usually the manufacturer or potential marketer) having screened the new molecule for pharmacological activity and acute toxicity potential in animals, wants to test its diagnostic or therapeutic potential in humans. At that point, the molecule changes in legal status under the Federal Food, Drug, and Cosmetic Act and becomes a new drug subject to specific requirements of the drug regulatory system.

*There are three IND types*

- **Investigator IND** is submitted by a physician who both initiates and conducts an investigation, and under whose immediate direction the investigational drug is administered or dispensed. A physician might submit a research IND to propose studying an unapproved drug, or an approved product for a new indication or in a new patient population.
- **Emergency use IND** allows the FDA to authorize the use of an experimental drug in an emergency situation that does not allow time for submission of an IND in accordance with 21CFR, Section 312.23 or Section 312.34. It is also used for patients who do not meet the criteria of

an existing study protocol, or if an approved study protocol does not exist.

- **Treatment IND** is submitted for experimental drugs showing promise in clinical testing for serious or immediately life-threatening conditions while the final clinical work is conducted and the FDA review takes place.

*There are two IND categories*

- Commercial
- Research (non-commercial)

*The IND application must contain information in three broad areas:*

- *Animal pharmacology and toxicology studies:* It contains preclinical data to permit an assessment as to whether the product is reasonably safe for initial testing in humans. Also included are any previous experiences with the drug in humans (often foreign use).
- *Manufacturing information:* It comprises of information pertaining to the composition, manufacturer, stability, and controls used for manufacturing the drug substances and drug products. This information is assessed to ensure that the company can adequately produce and supply consistent batches of the drug.
- *Clinical protocols and investigator information:* Detailed protocols for the proposed clinical studies to assess whether the initial-phase trials will expose subjects to unnecessary risks. Also, information on the qualifications of clinical investigators-professionals (generally physicians) who oversee the administration of the experimental compound to assess whether they are qualified to fulfill their clinical trial duties. Finally, it contains commitments to obtain informed consent from the research subjects, to obtain review of the study by an institutional review board (IRB), and to adhere to the investigational new drug regulations.

Once the IND is submitted, the sponsor must wait 30 calendar days before initiating any clinical trials. During this time, FDA has an opportunity to review the IND for safety to assure that research subjects will not be subjected to unreasonable risks. The act directs the FDA to place investigations on clinical hold if the drug involved presents unreasonable 

risks to the safety of the subjects. ***Clinical hold*** is an order issued by the FDA to the sponsor to delay a proposed clinical investigation or to suspend an ongoing investigation. Subjects may not be given the investigational drug or the hold may require that no new subjects be enrolled into an ongoing study. The clinical hold can be issued before the end of the 30-day IND review period to prevent a sponsor from initiating a proposed protocol or at any time during the life of an IND.

FDA's primary objectives in reviewing an IND are, in all phases of the investigation, to assure the safety and rights of subjects, and, in Phase 2 and 3 to help assure that the quality of the scientific evaluation of drugs is adequate to permit an evaluation of the drug's effectiveness and safety. Therefore, although FDA's review of Phase 1 submissions will focus on assessing the safety of Phase 1 investigations, FDA's review of Phases 2 and 3 submissions will also include an assessment of the scientific quality of the clinical investigations and the likelihood that the investigations will yield data capable of meeting statutory standards for marketing approval.

## *IND Application Process*

INDs can be withdrawn by the sponsor or terminated by the FDA. If no subjects are entered into clinical studies for a period of two years or all investigations are on clinical hold for one year or more, the IND may be placed on inactive status, either at the sponsor's request or at the FDA's initiative. If the IND remains on inactive status for five years or if FDA is unable to resolve deficiencies through a clinical hold or other alternative, it is terminated.

## *Content and Format of IND Application*

A sponsor who intends to conduct a clinical investigation subject to this part shall submit an "Investigational New Drug Application" including, in the following order:

(1) Cover sheet (Form FDA-1571)

The Form 1571 is a required part of the initial IND and every subsequent submission related to the IND application. Each *IND Amendment, IND Safety Report, IND Annual Report,* or general correspondence with the FDA regarding the IND must include Form 1571. It serves as a cover sheet for IND submissions and provides the FDA with basic information about the

submission—name of the sponsor, IND number, name of the drug, type of submission, serial number, and the contents of the application. Each submission to the IND must be consecutively numbered, starting with the initial IND application, which is numbered 0000. The next submission (response to clinical hold, correspondence, amendment, etc.) should be numbered 0001, with subsequent submissions numbered consecutively in the order of submission. There are significant commitments involved with the signing of FDA Form 1571 and the sponsor should be aware that signing the Form 1571 is more than a formality. Making a willfully false statement on the Form 1571 or accompanying documentation is a criminal offence. Detailed information on completing the Form 1571 can be found on the FDA website, in section 312.23(a)(1) and from the FDA review division responsible for reviewing the IND.

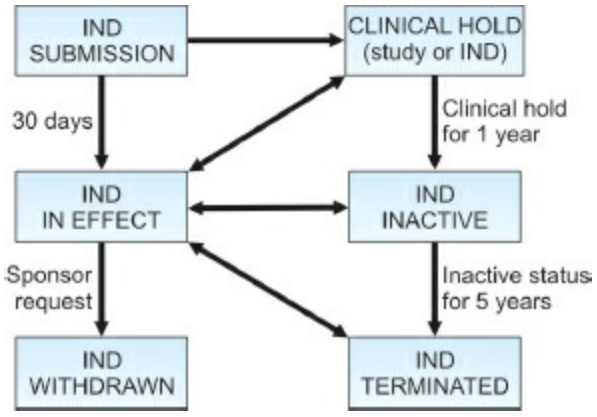



**Fig. 30.2:** Schematic representation of the IND application process

## *IND Amendment*

A submission to the IND file that adds new or revised information. Every submission adds to, revises, or affects the body of information within the IND and is, therefore, considered an IND amendment. *Protocol amendments and information amendments* are two examples of information that are filed to an IND in the course of clinical development.

*Protocol amendment:* A protocol amendment is submitted when a sponsor intends to conduct a new study, wishes to modify the design or conduct of a previously submitted study protocol, or adds a new investigator to a protocol. However, there is no 30 day review period after the initial IND review. When adding new investigators or updating investigator information

(e.g. FDA Form 1572) to a previously submitted protocol, the investigator, and information may be batched and submitted every 30 days.

*Information amendment:* An information amendment is used to submit new toxicology, pharmacology, clinical, or other information that does not fall within the scope of a protocol amendment, annual report, or IND safety report.

## *IND Safety Report*

An expedited report sent (within 15 calendar days of sponsor's receipt) to the FDA and all participating investigators of a serious and unexpected adverse experience associated with use of the drug or findings from nonclinical studies that suggest a risk to human subjects.

## *IND Annual Report*

A brief report to the FDA on the progress of the clinical investigations. It is submitted each year within 60 days of the anniversary date that the IND went into effect.

*A table of contents:* This should be a comprehensive listing of the contents of the IND broken down by section, volume, and page number. The table of contents should include all the required sections, appendices, attachments, reports, and other reference material.

*Introductory statement and general investigational plan:* This part should include information about the indication(s) to be studied, name of the drug and all active ingredients, the drug's pharmacological class and structural formula (if known), the formulation of the dosage form(s) to be used, route of administration, and broad objectives and planned duration of the proposed clinical investigation(s). If the drug has been previously administered to humans, the introductory statement should include a brief summary of human clinical experience to date, focusing mainly on safety of the drug in previous studies and how that supports studies proposed in the IND. If the drug was withdrawn from investigation or marketing in any country for safety reasons, the name of the country and reasons for withdrawal should also be briefly discussed in the introductory statement.

*Investigator's brochure (IB):* The investigator's brochure is a key document that is provided to each clinical investigator and the institutional review board at each of the clinical sites. The IB presents, in summary form,

the key nonclinical (safety), clinical, and CMC (quality) data that support the proposed clinical trial.

*Protocols:* A clinical protocol describes how a particular clinical trial is to be conducted. It describes the objectives of the study, the trial design, the selection of subjects, and the manner in which the trial is to be carried out. The initial IND is required to have a clinical protocol for the initial planned study. Protocols for Phase 1 studies may be less detailed and more flexible than those for Phase 2 and 3 studies. In Phases 2 and 3 detailed protocols describing all the aspects of the study should be submitted. Content and format of the protocols are provided in detail in 21 CFR, Section 312.23 and ICH guidance E6: Good Clinical Practices guidance document.

*Chemistry, manufacturing, and control (CMC) information:* This key section of an IND describes the quality information, comprising the composition, manufacturing process and control of the drug substance and drug product. The CMC section must provide sufficient details and information to demonstrate the identity, quality, purity, and potency of the drug product. The amount of information to be submitted depends upon the scope of the proposed clinical investigation. For a phase 1 IND the CMC information provided for the raw materials, drug substance, and drug product should be sufficiently detailed to allow the FDA to evaluate the safety of the subjects participating in the trial. A safety concern or lack of data, which make it impossible for the FDA to conduct a safety evaluation, are the only reasons for a clinical hold based on the CMC section. FDA requires that any drug product intended for administration to humans be manufactured in conformance with cGMP. So, all the drug products which are subject of IND should be manufactured in compliance with cGMP requirements. Along with the detailed information about the drug substance, inactives and final drug product, this section should also include a copy of all labels and labeling to be provided to each investigator and a claim for categorical exclusion from environmental assessment.

*Pharmacology and toxicology information:* The pharmacology and toxicology section of the IND includes the nonclinical safety data that the sponsor generated to conclude that the new drug is reasonably safe for clinical study. Such information is required to include the identification and qualifications of the individuals who evaluated the results of such studies and concluded that it is reasonably safe to begin the proposed investigations and a

statement of where the investigations were conducted and where the records are available for inspection. As drug development proceeds, the sponsor is required to submit informational amendments, as appropriate, with additional information pertinent to its safety. Requirements on content of this section can be found in 21 CFR, Section 312.23. A declaration that each nonclinical safety study reported in the IND was performed in full compliance with Good Laboratory Practices (GLP) or if a study was not conducted in compliance with GLP, a brief statement of why it was not, and a discussion on how this might affect the interpretations of the findings is required to be submitted.

*Previous human experience with the investigational drug:* A summary of previous human experience known to the applicant, if any, with the investigational drug should be included.

*Additional information:* In certain applications, as described below, information on special topics may be needed. Such information shall be submitted in this section as follows:

i. *Drug dependence and abuse potential:* If the drug is a psychotropic substance or otherwise has an abuse potential, a section describing the relevant clinical studies and experience, and studies in test animals should be included.

ii. *Radioactive drugs:* If the drug is radioactive, sufficient data from animal or human studies to allow a reasonable calculation of radiation-absorbed dose to the whole body and critical organs upon administration to a human subject should be included. Phase 1 studies of radioactive drugs must include studies which will obtain sufficient data for dosimetric calculations.

iii. *Pediatric studies:* Plans for assessing pediatric safety and effectiveness should be incorporated.

iv. *Other information:* A brief statement of any other information that would aid the evaluation of the proposed clinical investigations with respect to their safety or their design and potential as controlled clinical trials to support marketing of the drug should be incorporated.

*Relevant information:* Any information specifically requested by the FDA that is needed to review the IND application. It is common to place the meeting minutes from any pre-IND meeting or discussion in this section. This is especially useful if the information is referenced elsewhere in the

IND.

*Labeling investigational drugs:* Labels of investigational drugs are required to carry the following statement: “Caution: New Drug-Limited by Federal (or United States) law to investigational use”.

## *Promotion and Charging of an IND*

A sponsor or investigator, or any person acting on behalf of a sponsor or investigator, shall not represent in a promotional context that an investigational new drug is safe or effective for the purposes for which it is under investigation, or otherwise promote the drug. Administration of investigational drugs should be free of charge. Charging for an investigational drug in a clinical trial under an IND is not permitted without the prior written approval of FDA. In requesting such an approval, the sponsor shall provide a full written explanation of why charging is necessary in order for the sponsor to undertake or continue the clinical trial, e.g. why distribution of the drug to test subjects should not be considered a part of the normal cost of doing business. The sponsor may not commercialize an investigational drug by charging a price larger than that necessary to recover costs of manufacture, research, development, and handling of the investigational drug.

## New Drug Application (NDA)

When the sponsor of a new drug believes that enough evidence on the drug's safety and effectiveness has been obtained to meet FDA's requirements for marketing approval, the sponsor submits to FDA a New Drug Application (NDA). The application must contain data from specific technical viewpoints for review, including chemistry, pharmacology, therapeutics, biopharmaceutics, and statistics. If the NDA is approved, the product may be marketed in the United States.

The new drug application is the single most important filing necessary to obtain marketing approval for a drug in the United States. The NDA is organized into specific, technical sections, which are evaluated by specialized review teams of highly qualified experts. Together, the review teams will make a decision to approve or disapprove the NDA. The center for drug evaluation and research (CDER) is the group at the FDA charged with reviewing NDAs for drugs and some biologic products. The FDA's authority to require and review an NDA (prior to an applicant marketing the product in the United States) is clearly stated in section 505 of the food, drug and cosmetic Act (21 USC 355). The Act requires that the application contain:

(A) full reports of investigations which have been made to show whether or not such drug is safe for use and whether such drug is effective in use; (B) a full list of the articles used as components of such drug; (C) a full statement of the composition of such drug; (D) a full description of the methods used in, and the facilities and controls used for, the manufacture, processing, and packaging of such drug; (E) such samples of such drug and of the articles used as components thereof as the Secretary may require; (F) specimens of the labeling proposed to be used for such drug and (G) any assessments required under section 505B. The applicant shall file with the application the patent number and the expiration date of any patent which claims the drug for which the applicant submitted the application or which claims a method of using such drug and with respect to which a claim of patent infringement could reasonably be asserted if a person not licensed by the owner engaged in the manufacture, use, or sale of the drug.

Based on the requirements of the Act, FDA has published the detailed requirements on the content of NDA in 21 CFR, Part 314, which are as follows:

**Application form (FDA form 356H):** This form is to be submitted with each original NDA and every supplement thereof. The form can be downloaded from the FDA website. FDA publishes and updates this form periodically. The form contains information about the sponsor, the drug and the proposed indication, as well as a checklist of the items contained in the NDA. If the applicant is not located in the United States, the form must name an agent with a US address.

**Index (for paper copies only):** The archival copy (copy that is retained by FDA as reference) of the application is required to contain a comprehensive index by volume number and page number to the summary under paragraph (c) of this section, the technical sections under paragraph (d) of this section, and the supporting information under paragraph (f) of this section.

**Summary:** The application summary (required for only original application) is an abbreviated version of the entire application. Application summary should be written in enough detail that the reader may gain a good general understanding of the data and information in the application, including an understanding of the quantitative aspects of the data. The summary is required to contain the following information:

**The proposed text of the labeling**, including, if applicable, any Medication Guide. Requirements for labeling are very specific. Regulations require that the labeling contain the specific information under the following headings and subheadings in the following order:

**Highlights of prescribing information:** This section should not be more than one page. It is incorporated to facilitate the information reading by prescribers.

- Product names, other required information
- Boxed warning
- Recent major changes
- Indications and usage
- Dosage and administration
- Dosage forms and strengths
- Contraindications
- Warnings and precautions

- Adverse reactions
- Drug interactions
- Use in specific populations

## *Full Prescribing Information: Contents*

a. Full prescribing information
b. Boxed warning
    1. Indications and usage
    2. Dosage and administration
    3. Dosage forms and strengths
    4. Contraindications
    5. Warnings and precautions
    6. Adverse reactions
    7. Drug interactions
    8. Use in specific populations
        8.1 Pregnancy
        8.2 Labor and delivery
        8.3 Nursing mothers
        8.4 Pediatric use
        8.5 Geriatric use
    9. Drug abuse and dependence
        9.1 Controlled substance
        9.2 Abuse
        9.3 Dependence
    10. Overdosage
    11. Description
    12. Clinical pharmacology
        12.1Mechanism of action
        12.2Pharmacodynamics
        12.3Pharmacokinetics

13. Nonclinical toxicology
    a. Carcinogenesis, mutagenesis, impairment of fertility
    b. Animal toxicology and/or pharmacology
14. Clinical studies
15. References
16. How supplied/storage and handling
17. Patient counseling information.

**Medication guide:** A medication guide is required for prescription drug or biological products when FDA has determined that either the product poses a serious and significant public health concern or for that it is needed to ensure patients' safe and effective use of the product. The medication guide content outline is:

- Brand name (established name).
- Most important information to know about the drug
- Who should not take the drug
- How to take the drug
- What to avoid while taking the drug
- Possible or reasonably likely side effects of the drug
- General information.

Manufacturer/distributor should ensure that a separate medication guide is made available to each patient to whom the drug product is dispensed. Care should be taken to ensure that the content of the medication guide is not promotional in nature.

1. A statement identifying the pharmacologic class, scientific rationale for the drug, its intended use, and the potential clinical benefits of the drug product.
2. Foreign marketing history.
3. A summary of the chemistry, manufacturing, and controls section of the application: This section should include CMC information on drug substance, inactives and drug product.
4. A summary of the nonclinical pharmacology and toxicology section of the application.

5. A summary of the human pharmacokinetics and bioavailability section of the application.
6. A summary of the microbiology section of the application (for anti-infective drugs only).
7. A summary of the clinical data section of the application, including the results of statistical analyses of the clinical trials.
8. A concluding discussion that presents the benefit and risk considerations related to the drug, including a discussion of any proposed additional studies or surveillance the applicant intends to conduct post-marketing.

## *Technical Sections*

The application is required to contain the technical sections described below. Each technical section is required to contain data and information in sufficient detail to permit the agency to make a knowledgeable judgment about whether to approve the application or whether grounds exist under Section 505(d) of the act to refuse to approve the application. The required technical sections are as follows:

- Chemistry, manufacturing and controls (CMC): A section describing the composition, manufacture, and specification of the drug substance and the drug product. The applicant may, at its option, submit a complete chemistry, manufacturing, and controls section 90 to 120 days before the anticipated submission of the remainder of the application. FDA will review such early submissions as resources permit.
- *Nonclinical pharmacology and toxicology section.*
- *Human pharmacokinetics and bioavailability section*.
- *Microbiology section* (If the drug is an antiinfective drug)
- *Clinical data section.*
- *Statistical section:* A section describing the statistical evaluation of the clinical data.
- *Pediatric use section:* A section describing the investigation of the drug for use in pediatric populations.

## *Samples and Labeling*

Samples of the drug substance, drug product, reference standards, blanks and final marketed package. Samples are to be provided in sufficient quantity so as to enable FDA to perform the tests specified in the application up to three times.

Copies of the label and all labeling for the drug product (including, if applicable, any Medication Guide required) for the drug product (4 copies of draft labeling or 12 copies of final printed labeling).

## *Case Report forms and Tabulations*

## *Patent Information*

An applicant is required to disclose all patent information that is related to the drug for which the NDA is being filed and to verify that the sponsor has all rights necessary to legally manufacture, use, and sell the drug, if the NDA is approved. The patent inquiry is a broad one and covers drug substance (active ingredient) patents, drug product (formulation and composition) patents, and method-of-use patents.

## *Patent Certification and Claimed Exclusivity*

The NDA is required to contain information regarding all financial interests or arrangements between clinical investigators, their spouses and immediate family members, and the sponsor of the clinical trials that support the NDA. The applicant should submit an FDA Form 3454 to certify which investigators had no financial interests or arrangements. An FDA Form 3455 is submitted to disclose any financial interests or arrangements with an investigator that could affect the outcome of the study and a description of the steps taken to minimize the potential bias of the study results.

## *User Fee Cover Sheet (FORM FDA-3397)*

FDA charges the sponsors of NDA a fee for reviewing and approving an NDA. The fee thus collected is utilized for funding the staff employed with FDA and other activities of FDA. A User Fee Cover Sheet is to be completed and submitted with each new drug or biologic product NDA. The form provides a cross-reference to the user fee paid by the applicant.

## *Debarment Certification*

FDA uses its powers to debar persons from involving themselves in any activity related to the applications filed with FDA. This action is taken against those individuals, who FDA thinks/finds have made false, misleading commitments to the FDA intentionally or unintentionally, which can render the approved drug unsafe for human consumption. All applications filed with FDA are required to contain a statement certifying that the applicant did not and will not use in any capacity the services of any person debarred by the FDA.

Once an NDA is approved, any significant change in the manufacturing, control, packaging, or other physical properties of the drug, or any change in its labeling that may have an effect on safety and effectiveness relative to either the drug itself or the manner in which it is used must be covered by a supplemental new drug application.

The requirements of an NDA for prescription and over-the-counter drugs are similar. It is extremely rare, however, that any new chemical entity would be approved by the FDA for over-the-counter sale for reasons of lack of sufficient data to support the safety in use of a totally “new” drug. Since the basic test is the ability to provide on the label adequate directions for use, most applications specifically for over-the-counter drugs would tend to be supplemental applications, mainly on the particular showing that the drug might be safely used without direct medical supervision.

Once approved, several conditions must be met in connection with an NDA. First, advertisement must be submitted to FDA on a routine basis. It is especially important that introductory advertising be submitted with some promptness. Second, reports of human clinical experience are required on the following basis. The first reports are due at intervals of three months, beginning with the date of approval of the application, during the first three years, and annually thereafter.

Records reflecting clinical experience with the drug must be retained, and must be available for inspection. The reports required at the aforementioned intervals must include unpublished reports of clinical experience, studies, investigations, and tests conducted by the applicant or reported to him, unpublished reports of animal experience, experience involving the chemical, or physical properties of the drug, and any information that might affect the

safety or efficacy of the product. These may be submitted on forms FD-1639, FD-2253, or FD-2252.

The FDA requires immediate reporting (called NDA Field Alert Report) of information concerning any mix-up in a drug or its labeling with another article; information concerning any bacteriologic or any significant chemical, physical, or other changes, or deterioration in the drug; or any failure of one or more distributed batches of the drug to meet the specifications established for it in the New drug application. This information could, of course, form the basis for a recall. Such information is to be submitted to the FDA within 3 working days from the day, when such information came into the notice of the applicant.

Certain other information must be transmitted as soon as possible-in any event, within 15 working days of receipt. This information would pertain to serious and unsuspected side effects, injury, toxicity, sensitivity reaction, or incidents associated with clinical use, studies, investigations, or tests, whether or not they are determined to be attributable to the drug, or information concerning any unusual failure of the drug to exhibit its expected pharmacologic activity.

## Abbreviated New Drug Application (ANDA)

An Abbreviated New Drug Application (ANDA) contains data that, when submitted to FDA's Center for Drug Evaluation and Research, and Office of Generic Drugs, provides for the review and ultimate approval of a generic drug product. Generic drug applications are called "abbreviated" because they are generally not required to include preclinical (animal) and clinical (human) data to establish safety and effectiveness. Instead, a generic applicant must scientifically demonstrate that its product is bioequivalent (i.e. performs in the same manner as the innovator drug). Once approved, an applicant may manufacture and market the generic drug product to provide a safe, effective and low cost alternative to the American public.

Kefauver-Harris Drug Amendment (1962) and Hatch-Waxman Act (1984) are two regulations that facilitated the growth of the generic industry. Paper NDA (a provision under Kefauver-Harris Drug Amendment, 1962) were the early equivalents to today's ANDA, wherein the submission used to contain copies of published literature (on safety and efficacy) and bioequivalence studies comparing the generic drug to the innovator drug as the evidence of safety and efficacy. Applications under the "abbreviated" tag were accepted only for drugs approved prior to 1962 and which were labeled safe and effective under DESI review. Hatch-Waxman Act made an attempt to balance the commercial and competitive aspects of the generic and innovator drugs. It created provisions for Marketing Exclusivity, patent term-extension for NDA holder and options of Suitability petition under Section 505(b)2 and, 180-day generic drug exclusivity to the post-innovator applicants.

NDA applicants were provided a protection period through Marketing exclusivity (5 years for NCE exclusivity, 3 years for new clinical investigations). During this period no other competitor was allowed to enter into the market. The provision was aimed at letting the innovator recover the investments made while developing the drugs. All the NDA were required to list the relevant patents for the filed application. Besides this, NDA drugs were also protected through patents. NDA applicants were required to identify and submit the patents with the NDA at the time of initial filing. New patents could be added to this list over the life cycle of the drug.

Generic drug applications are required to prove that they are

therapeutically equivalent to innovator drugs. This is accomplished by demonstrating:

- Bio-equivalence
- Pharmaceutical equivalence
- Compliance with cGMP requirements
- Sameness of labeling with the innovator approved labeling: Generic labels could not include the exclusivity or patent protected information. Only differences allowed between the labeling of generic drugs and innovator drugs were those related to the manufacturers like description of the finished products, ingredients of the formulations, NDC number of the formulations, etc.

Generic drugs were also required to include certifications of the status of all patents to the reference innovator drug. Accordingly the generic submissions are categorized as:

**Paragraph I filings:** No patent information has been previously submitted to the FDA.

**Paragraph II filings:** All the submitted patents relevant to drug have expired.

**Paragraph III filing:** If the applicant states that he intends to market the drug only after the expiry of the listed patents.

**Paragraph IV filings:** If the applicant certifies that any product or used patent is invalid or will not be infringed by the generic drug approval. In such cases, the applicant must give a notice to the innovator applicant. If the innovator thinks that generic drug approval will infringe upon his patents, then he can sue the generic applicant within 45 days of receiving the notice. Under such circumstances, an automatic stay period of up to 30 months (pending resolution) is applied on to the generic drug application. Generic drug application is not reviewed by the FDA during this period.

FDA has issued a checklist on the content of ANDA filings. All the ANDAs are required to comply with this checklist if submitted in CTD (common technical document) format.

## Biologic License Application (BLA)

Biological products are approved for marketing under the provisions of the Public Health Service (PHS) Act. The Act requires a firm who manufactures a biologic for sale in interstate commerce to hold a license for the product. A BLA is a submission that contains specific information on the manufacturing processes, chemistry, pharmacology, clinical pharmacology and the medical effects of the biological product. If the information provided meets FDA requirements, the application is approved and a license is issued allowing the firm to market the product.

## Over-the-counter Drugs (OTC)

OTC drugs are those that are available to consumers without a prescription. There are more than 80 therapeutic categories of OTC drugs, ranging from anti-acne drugs to weight control drugs. As with the prescription drugs, CDER oversees OTC drugs to ensure that they are properly labeled and that their benefits outweigh their risks.

*OTC drugs generally have these characteristics*

- Their benefits outweigh their risks
- The potential for misuse and abuse is low
- Consumer can use them for self-diagnosed conditions
- They can be adequately labeled
- Health practitioners are not needed for the safe and effective use of these drugs.

*Two regulatory mechanisms exist for the legal marketing of OTC drugs*

1. NDA (regulations described in 21 CFR Part 314).
2. OTC drug monograph (regulations described in 21 CFR Part 330).

OTC drugs marketed under either mechanism must meet the established standards for safety and effectiveness. Although we assess compliance with these standards differently under the two mechanisms, neither mechanism establishes higher standards for safety or effectiveness than the other. Under both mechanisms, products must be manufactured according to current good manufacturing practices (cGMPs) as defined in 21 CFR Part 210 and must comply with the labeling content and format requirements in 21 CFR Part 201 Subpart C.

1. **NDA:** Legal marketing is under the authority of an approved product-specific New Drug Application (NDA) or an Abbreviated New Drug Application (ANDA). An OTC drug product with active ingredient(s), dosage form, dosage strength, or route of administration new to the OTC marketplace is regulated under the NDA process. For example, a drug product previously available only by prescription (Rx) can be marketed OTC under an approved “Rx-to-OTC switch” NDA.

   FDA must approve the NDA for an OTC drug product before that

product can be marketed under the OTC category. A drug manufacturer submits data in an NDA demonstrating that a drug product is safe and effective for use by consumers without the assistance of a healthcare professional. FDA must review the data within an established timeframe, and the data submitted in an NDA remains confidential.

The drug manufacturer can only market the product with the specific formulation and exact labeling approved by FDA. To make a change, the manufacturer must submit an NDA supplement and the FDA must approve that supplement.

2. **OTC drug monograph:** Legal marketing is in compliance with an OTC drug monograph. Unlike NDAs which are based on drug *products*, monographs specify the active ingredients that can be contained within OTC drugs. An OTC drug product containing ingredients that comply with standards established in an applicable monograph is considered to be "generally recognized as safe and effective" (GRASE) and does not require specific FDA approval before marketing. For example, OTC sunscreen drug products can be legally marketed if they contain ingredients which comply with the standards established in the OTC sunscreen monograph for formulation, labeling, and testing.

Labeling of the OTC drug must comply with the "Drug Facts" format. Labeling of the OTC drugs must contain "Adequate directions for use and warnings". An OTC drug label (seven-point label) must contain the following:

1. Name of the active ingredient
2. Information on uses
3. Warnings
4. Inactive ingredients
5. Purpose
6. Directions for use
7. Other information like directions for storage.

*OTC drug labels are required to contain the following warning statements*

“Keep out of the reach of children”.

“In case of overdose, get medical help or contact a poison control center right away”.

If any changes related to active ingredient, directions for use, warnings or precautions, or indications are made to the originally approved label than the changes must be flagged to identify such changes.

OTC drug products must be provided in Tamper-resistant/evident packaging. Tamper-resistant feature must be identified on the label. Two-piece hard gelatin capsules are required to have double tamper-resistant features.

## REVIEW PROCEDURES

Once a complete application (NDA, BLA, ANDA) is filed with FDA for review, it is sent to the concerned Center for Review. FDA charges an application fee for reviewing NDA and BLAs (ANDAs are not charged any review fee). The application fee is charged as per the projected financial budget and number of applications to be filed with the FDA. Application fee for the financial year 2008 is set at $ 1,78,000. After receiving the application for review, the Center designates a project manager who is responsible for coordinating the review. Project manager is the primary point of contact for the applicant for any type of communication. Following are typical phases of the review cycle:

Application is checked in the document control centre. Application is checked for completeness of content and technical details. This helps the center to determine whether a worthwhile review is possible (Time limit: Within 60 days of receiving the application). If application is found to be complete an “Acceptance to file” letter is sent to the applicant. If application is found incomplete, a “Refuse to File (RTF)” is sent to the applicant. If an “RTF” is received, the applicant is required to provide all the required information. FDA will initiate the review only after receiving the complete documents.

Different sections of the application are distributed to the review team. Applications are taken up for review based on the priority procedure. Applications are labeled standard (S category) or priority (P category) review based upon the type of application. AIDS drug applications are high priority (AA category) review drug applications. During the review cycle, a 90 day meeting can be held between the applicant and FDA to discuss the progress and issues.

*Issuance of action letter:* An action letter is issued by FDA (project manager) to the applicant (time line: Within 180 days of “Acceptance to file” letter). This letter is issued after technical evaluation of different sections of the application is completed by respective offices. If the application is completely acceptable, an “Approval Letter” is issued. If there are some problems, a “Not Approvable Letter (major problems)” is issued which means that the application is not approvable, or an approvable letter (minor problems) is issued if application can be approved after addressing the

deficiencies. A complete response letter is issued in case of BLAs. These problem letters summarieze the deficiencies in the application and also tells what needs to be done (if something can be) in order to approve the application.

Figure 30.3 outlines the review process of an NDA application:

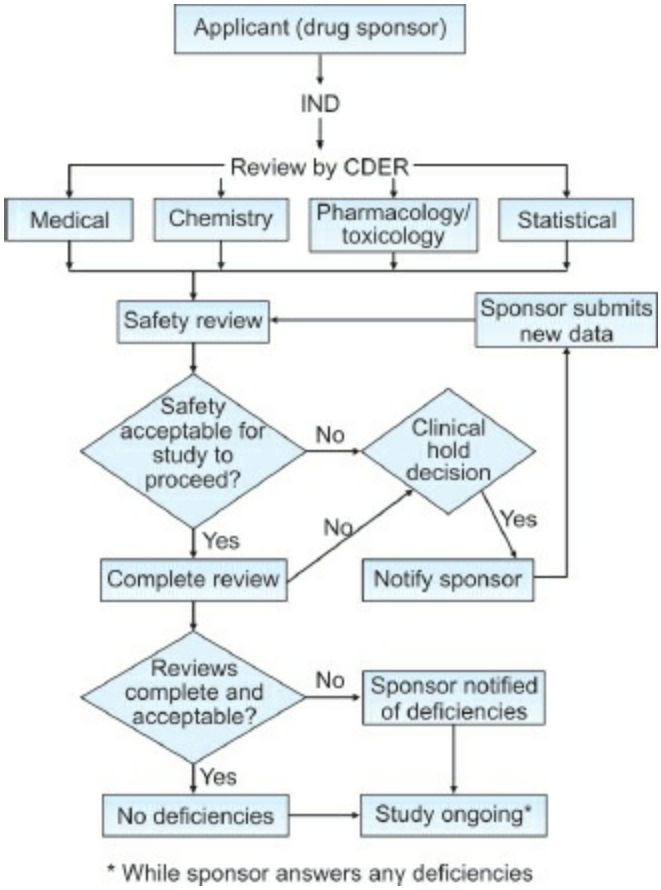



**Fig. 30.3:** Schematic representation of the NDA review process

## ANDA Review Procedure

The review procedure is same for ANDAs, but there are some additional issues which should be kept in mind:

- Review cycle might not be that strictly followed, as it is done in case of NDAs.
- A 45 days gap between “acceptance to file” letter and initiation of review will be there in case of Para IV filings.
- There is a strict time line that the applicant has to follow, in order to respond to different deficiencies raised during the course of review.
- *Telephonic amendments:* FDA may ask for some additional clarifications/data to support efficient review of the application. Response has to be submitted within 10 working days.
- *Minor amendments:* Response has to be submitted within 60 working days of receiving the deficiency letter.
- *Major amendments:* Response has to be submitted within 180 working days of receiving the deficiency letter.

Review of the application is stopped when the applicant is responding to minor or major deficiency. If the applicant does not respond to minor deficiencies within 60 days, category is changed to major amendment. Review is not stopped in case of telephonic amendment.

## Filing Formats and Contents

The format discussed above is the one that is developed and accepted by the FDA for filing applications with FDA. Regulatory authorities from USA, EU and Japan came together due to pressure from the pharmaceutical manufacturers and to supply the latest treatments to their respective populations, and developed a common format for filing drug approval applications in these three regions. The common body is known as International Conference on Harmonization (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use and the format for filing application as CTD. This is to be noticed that the format is called Common Technical Document, because harmonization has been achieved only for the format of the technical part of the filing. Format of administrative sections of the three regions are still different and are to be prepared as per the regional requirements. Few parts of the technical sections as required by the regulatory authorities are still different from the CTD format. The guidances issued by the regulatory authorities clarify when these are needed and where in the CTD format they should be placed.

The ICH is a joint initiative involving both regulators and research-based industry representatives of the European Union, Japan, and the United States in scientific and technical discussions of the testing procedures required to assess and ensure the safety, quality, and efficacy of medicines. It is important for applicants to understand that the CTD format provides guidance on the organization of documents in the submission. The content of the CTD is determined by regulations and discussions with regional regulatory authorities.

CTD divides a drug application into the following modules:

**Module 1 (administrative and prescribing information):** No common format has been finalized.

**Module 2 (summaries):** This section includes:

- *Quality overall summary:* It provides reviews on overview of the data incorporated in Module 3
- *Nonclinical overview:* The nonclinical overview should provide an interpretation of the data, the clinical relevance of the findings cross-linked to the quality aspects of the pharmaceutical, and the

implications of the nonclinical findings for the safe use of the pharmaceutical.

- *Nonclinical written and tabulated summaries:* The nonclinical written and tabulated summaries should provide a comprehensive, factual synopsis of the nonclinical data.
- *Clinical overview:* The clinical overview should provide a succinct discussion and interpretation of the clinical findings that support the application together with any other relevant information such as pertinent animal data or product quality issues that may have clinical implications.
- *Clinical summary:* The clinical summary should provide a detailed factual summarization of the clinical information in the application.

**Module 3 (quality):** Module 3 contains the detailed technical information regarding CMC information. This module contains a TOC (for paper submission only) for module 3 only, detailed information regarding the drug substance and drug product.

**Module 4 (nonclinical study report):** Module 4 contains a TOC (for paper submission only) for module 4 only, nonclinical study reports contained in the application, and literature references.

**Module 5 (clinical study reports):** Module 5 contains a TOC (for paper submission only) for module 5 only and a tabular listing of all clinical studies, clinical study reports, and literature references.

Requirements for preparing module 1 can be found at the official websites of regulatory authorities of respective regions.

## ELECTRONIC SUBMISSION

Paper submission (submission of data in paper) is the traditional way in which applications are submitted to the authorities across the world. But over a period of time and with the evolution of information technology, regulatory authorities have realized that it will be far more convenient (and efficient) to review the applications if the applications are received in electronic format (submission of data in PDF, i.e. portable document format files).

Electronic submission could be in Electronic CTD (eCTD) format or non-eCTD format. In non-eCTD format PDF documents are prepared (authored) as per the requirements for electronic submissions. Requirements include those specified for:

- PDF file size
- Scanning resolution
- Hypertext and book marking
- File security properties
- Pagination, etc.

Non-eCTD submissions contain electronic documents arranged as per the CTD hierarchy. When submissions are made in eCTD format, they are prepared using a particular software format called XML (extensible markup language). This allows for easy navigation and review of the submissions. eCTD submission also provides for information submission throughout the life-cycle of the product. Most of the regulatory agencies in the regulated world are working towards making it the mandatory format for electronic submissions.

## CERTIFICATION OF DRUGS

Insulin and antibiotics are the two classes of drugs requiring certification, although certain antibiotics may be exempt from certification. The FDA, acting under the mandate of the Act, has adopted regulations to ensure that each batch has the characteristics of identity, strength, quality, and purity prescribed in such regulations.

Application for certification must be made to the Commissioner. For insulin, the information that must be contained in a request for certification is set forth in 21 CFR Section 164.2.

In 1982, the FDA published a regulation that exempts virtually all antibiotics from the requirement of obtaining batch certification. For 20 years, companies marketing approved antibiotics had been required to submit assay results and samples of each manufacturing batch for FDA testing and evaluation. The purpose of this exercise was to assure batch-to-batch integrity and uniformity, and this testing was paid for by the companies involved and was both costly and timeconsuming (a batch could not be released for sale until FDA passed it). Apparently finding the exercise no longer justified, the FDA dropped it. Batch confirmation continued to be required for insulin products.

## DRUG EFFICACY STUDY IMPLEMENTATION (DESI) REVIEW

A particularly controversial provision of the 1962 Amendments concerns drugs already on the market that had effective NDAs under the old law, i.e. during the period from 1938 to 1962. A two-year period of grace was provided to the manufacturer to supply substantial evidence of effectiveness to support the claims on the previously cleared labeling, since the 1938 Act required proof of safety only.

To cope with the overwhelming burden of such a review, FDA enlisted the services of the National Academy of Sciences and National Research Council (NAS/NRC) for help. The NAS/NRC set up a series of panels to which drugs were assigned on the basis of the medical disciplines in which they were used, e.g. Ophthalmic panel, Dermatology panel, and Cardiology panel. Manufacturers of drugs affected were invited to submit data to these panels in addition to the data already contained in their NDAs. Upon evaluation of these data, the panels rated the drugs on effectiveness for their labeled indications as follows:

1. Effective
    - Effective, but...
2. Possibly effective
3. Probably effective
4. Ineffective

The panels worked independently, and frequently a drug would be judged effective for one of its uses by one panel and less than effective, i.e. ratings la through 4, for another of its uses by another panel. These apparent differences were not resolved by joint panel meetings, but were treated as independent judgments having equal weight.

The extent of the problem can be seen in the fact that the NAS/NRC Panels returned 2,824 reports covering approximately 3,700 drugs manufactured by 237 companies. The panels began submitting reports on October 11, 1967, and the last report was submitted on April 15, 1969. The FDA attempted to implement these reports by publishing findings in the Federal Register, and in the case of drugs that were probably or possibly

effective, the drug would be allowed to remain on the market for periods of 12 months and 6 months, respectively, to allow for further clinical trials to support their efficacy. After these periods, if no work had been done, a notice of withdrawal was prepared and published in the Federal Register. This gave rise to requests for hearings under the appropriate sections of the Act. In all cases, hearings were denied based on a finding by the Secretary that no issue of fact had been presented on the face of the petitioning papers to require a hearing.

The apparent slowness with which the FDA was implementing the NAS/NRC Reports gave rise to litigation, wherein the American Public Health Association and the National Council of Senior Citizens brought an action against the then acting Secretary of the department of Health, Education and Welfare, and the Commissioner of the FDA to make public all the reports theretofore not public, and in addition, to speed up the process. This resulted in a memorandum by Judge Bryant of the United States District Court for the District of Columbia, wherein the court concluded that there is no compelling reason why the NAS/NRC reports which were not then public should not be immediately released, and that it would be an abuse of agency's discretion to refuse to make such reports public. In addition, the court would set a deadline for the FDA to complete its evaluation of all drugs with regard to efficacy, and in October 1972, Judge Bryant threw open to public inspection all of the remaining reports. He set certain timetables (1) for implementing either a removal of less than effective classifications or a withdrawal of the NDAs, (2) for allowing certain drugs to remain on the market if good reason was shown by the FDA, pending a longer-term implementation and (3) for requiring reports to the court of the progress of the DESI review implementation program. Therefore, with the exception of less than effective drugs, which may be continued past the cutoff date because of public health considerations, all implementation would be finished by final order of FDA at least by October 10, 1976. The FDA did not appeal Judge Bryant ruling.

The FDA once again failed to perform under court order, and the plaintiffs went before Judge Bryant for a further order, which resulted in a settlement stipulation calling for the final regulatory action in all cases to be taken by June 1984. The stipulation set certain priorities within which the classes of drugs to be dealt with were specified. Periodic reports of progress were to be made by the FDA to Judge Bryant.

In its own attempt to give some notice to the practitioners that the drug they were planning to prescribe had received a rating of less than effective by the NAS/NRC panels, the FDA promulgated a regulation that became final on February 12, 1972. It required that the less than effective evaluation be called to the practitioner's attention by including a statement prominently in the labeling, and surrounding it with an appropriate border. Since most package insert labeling is printed black on white, this regulation promptly became known as the "Black Box" regulation. Thus, during the period of implementation ordered by Judge Bryant, all products rated less than effective must be so designated, in the "Indications" section of the package insert or on the label within an appropriate border, as follows:

> Based on a review of this drug by the National Academy of Sciences/National Research Council and/or information, FDA has classified the indication(s) as follows:
> (A statement of the probably or possibly effective status in paragraph form appears here) Final classification of the less than effective indications requires further investigation.

From the beginning, enforcement of the NAS/NRC findings was not without its problems and resolving these problems ultimately required recourse to the Supreme Court of the United States. First, there was the spectre of the so-called "grandfather clause" of the Kefauver-Harris Bill, which it was claimed, protected certain drugs within its definition from ever having to be proven effective. Second, it was claimed that certain drugs that were identical or similar to NDA drugs reviewed by NAS/NRC were unaffected by the findings of such review because they had been marketed in imitation of such drugs, but had not themselves had an NDA ("me-too" drugs). This claim was made in the face of FDA's proposed notice in February and final notice in October of 1972 that DESI notices of findings were equally applicable to identical, related, or similar drug products even though such products had never had an NDA.

Third, there was a controversy regarding which body, the FDA or the courts, had jurisdiction to decide whether a drug was a "new drug"—that is it was not generally recognized as safe and effective. Finally, there was the question whether a manufacturer had an absolute right to a hearing before FDA prior to withdrawal of his NDA or whether FDA could administratively decide that no hearing was warranted because, on the face of the documents offered in support of the drug, there was no "substantial evidence" of safety and/or efficacy.

These cases were fought vigorously through the lower courts, and the importance of their outcomes was deemed to be so great that the Supreme Court granted writs of certiorari, that is, it agreed to hear argument on the issues. They were resolved in June 1973.

The first case is the so-called Bentex case. In this case, a group of drug companies marketing pentylenetetrazol had filed suit against the FDA, alleging that their products were generally recognized as safe and effective, and therefore were not new drugs. They asked to be exempt from the regulatory results of an NAS/NRC panel report, which had found that other companies pentylenetetrazols, which were the subject of new drug applications, lacked substantial evidence of efficacy. This meant that the question before the court was whether a so-called “me-too” drug could be subject to the NAS/NRC results, even if it had never held an NDA and even if it might meet the literal requirements of the grandfather clause. The Supreme Court held that such products are subject to the NAS/NRC results.

Thus, “me-too” products must meet the standards announced by the FDA for NDA drugs, notwithstanding any other legal arguments that the “me-too” manufacturers might make. The court held that the FDA has jurisdiction to decide, with administrative finality, the new drug status of individual drugs or classes of drugs. The FDA, according to the Supreme Court, should not be required to litigate on a case-by-case basis the new drug status of each drug on the market. All of this means that henceforth FDA may say with great finality whether a product is a “new drug” and, in particular, can say that “me-too” products are new drugs and must come off the market if the NDA drug they are imitating is required to come off the market.

The second case before the supreme court involved a bioflavonoid product. In this instance, a lower court had held that the courts had jurisdiction to determine whether a product was protected by the grandfather clause, and that the FDA did not have authority to decide this question conclusively. The lower court also had held that the manufacturer’s own “me-too” versions of its NDA drug were subject to the NDA procedures. The Supreme Court held that the phrase “any drug,” as used in the grandfather clause, is used in a generic sense, which means that the “me-toos,” whether products of the same or of different manufacturers, covered by an effective NDA, are not exempt from the efficacy requirements of the 1962 Amendments. Thus, the grandfather clause has been further narrowed so that

it applies only to those drugs that meet the definition in the grandfather clause and are not “me-toos” of other drugs that did have NDAs.

The third decision involves the important question of a drug company’s rights to administrative hearings before the FDA. This is the Hynson, Westcott, and Dunning decision. In this case, a lower court had ruled that the company had presented enough evidence, in its request for a hearing to revoke a new drug application, to entitle that company to a hearing. In this decision, the supreme court sustained the summary judgment mechanism instituted by FDA to determine whether administrative hearings must be held. Under this mechanism, a drug company must present, in its hearing request, substantial evidence that the drug is effective, and if the FDA concludes that such evidence has not been presented, no hearing is granted, and the drug is removed from the market without hearing. The supreme court has sustained this mechanism. In the event that the FDA denies a hearing, the Supreme Court has ruled that a court of appeals, in reviewing this administrative action, “must determine whether the commissioner’s findings accurately reflect the study in question, and, if they do, whether the deficiencies he finds *conclusively* render the study inadequate or uncontrolled...”

Perhaps the most important statement in the Hynson, Westcott and Dunning decision regards a problem other than hearing requests. It had long been accepted that a drug could be an “old drug” if there was sufficient expert medical opinion declaring it to be generally recognized as safe and effective. It was further thought that this opinion could be just that opinion and need not be based on any particular type of scientific data. The Supreme Court has now severely narrowed this thinking and has ruled that a drug can be an old drug, based on a consensus of expert opinion, “only when that expert consensus is founded upon substantial evidence as defined in Section 505(d).” This means that a drug can be an old drug only if there is substantial evidence of efficacy by well-controlled studies that demonstrate that the product is in fact effective. Lower federal courts have followed this discussion, and there now remains no question as to the validity of this doctrine.

The final case before the Supreme Court involved CIBA. Here, a lower court had ruled that the FDA had the authority, in an administrative hearing, to determine whether a product was a new or old drug. CIBA had appealed this decision. The Supreme Court held, in a manner consistent with the

previous three decisions, that the agency does have such jurisdiction, and that, therefore, when a company has had one opportunity to litigate the new drug issue before FDA, it may not later try to litigate it in another forum, such as a federal court.

A final chapter in this controversy, the status of “me-too” or generic equivalent drugs, was finally decided in 1983. All through the late fifties and sixties, and well into the seventies, companies continued to market their unapproved versions of drugs that had gone through the NDA process, relying on various legal arguments and an FDA policy of allowing such marketing a policy that seemingly conceded that NDAs were unnecessary for generic drugs. As the DESI Review moved on and the ANDA policy was put into effect, however, a more definitive resolution of this problem was needed.

Following a series of lawsuits involving the generic drug industry, one finally reached the Supreme Court. There the Court ruled that generic versions of drugs that had gone through an NDA process (be it NDA or ANDA) also had to go through a similar procedure. The main reason for this decision was the Court’s reliance on the fact that changes in inactive ingredients such as colors, stabilizers, binders, or the life could affect the safety or efficacy of an active ingredient. Thus, each should be separately approved by the FDA. With that ruling, the last issue of the DESI Review was resolved, so that the question of whether there are any drugs that do not require FDA approval is largely academic.

## DRUG LISTING ACT OF 1972

By the terms of the Kefauver-Harris Amendment of 1962, all producers of drugs and drug products were required to register their establishments with FDA. The effect of registration was twofold (1) it gave the FDA an updated list of all legitimate manufacturers and (2) by the operation of the amendment, it required inspection of these producers at least once every two years. It did not, however, provide the FDA with a list of the types of drugs produced at each of these establishments, or with formulas for the many thousands of such drugs. Since the 1962 Amendment required that drugs, whether prescription or over-the-counter, which were deemed to be less than effective come off the market, FDA was faced with the almost insurmountable job of locating all such drugs and taking appropriate action. Regulations have been promulgated, however, under Section 2 of the drug listing Act to ensure that sufficient data be sent to the FDA and be computerized for easy handling. ***Drug listing information is required to be filed only in the electronic format from June 1, 2009***.

Once the data are fully computerized and classes of drugs can be identified, the FDA intends to use the broad new powers, which it gained in June of 1973 by dint of the four landmark Supreme Court decisions already cited, to move against drug entities deemed to be less than effective. With the DESI and OTC Reviews affecting virtually every drug currently on the market, and with the Supreme Court mandate, FDA feels it need not proceed on a case-by-case basis, but merely needs to locate violative drugs and notify the producer that it intends to take action. If the producer persists in marketing the violative drug, not only will seizure and injunction follow, but use of the criminal penalties of the Act will increase.

As mentioned earlier, most of the landmark advances in the regulatory development were triggered by adverse events. A brief look at these events (Table 30.1) would be helpful in developing a better understanding/appreciation of the current regulations.

Above listed events along with others have resulted into the regulations that are imposed on the drug products today in USA.

**Table 30.1:** Landmark advances in the regulatory development along with their key features

| Title/Year | Incident/Key feature |
|---|---|
| Drug Importation Act (1848) | Incident: Quinine used by American troops in Mexico was found to be adulterated |
| | Enactment of the Act now required US customs service inspection to stop entry of adulterated drugs from overseas |
| Biologics Control Act (1902) | Authorized US public health service to license establishment and ensure purity and safety of serums, vaccines, and similar products in interstate sales |
| Novel “The Jungle” By Novelist Upton Sinclair (1906) | Shocking disclosures of in sanitary conditions in meat-packing plants. This novel initiated the process of forming regulations for Consumer Protection |
| Food and Drugs Act 1906 | Passed by congress on June 30. It prohibits interstate commerce in misbranded and adulterated foods, drinks and drugs |
| | Did not authorize to evaluate the safety and efficacy of the drug or to ban the sale of unsafe drug |
| | This law did not cover Cosmetics |
| | In **US v Johnson (1911)**, the Supreme court rules that food and Drug Act 1906, does not prohibit false therapeutic claims but only false and misleading statements about the ingredients or identity of a drug |
| Sherley Amendment (1912) | Incident: Mrs Winslow’s soothing syrup for teething and colicky babies, unlabeled yet laced with morphine, killed many infants |
| | Sherley amendment prohibits labeling medicines with false therapeutic claims intended to defraud the purchaser. But in absence of safety and efficacy data, authenticity of the therapeutic claims was difficult to prove |

| | |
|---|---|
| | The name of the food, drug, and insecticide administration is shortened to **Food and Drug Administration (FDA)** under an agricultural appropriations act<br><br>FDA recommends a complete revision of the obsolete 1906 Food and drugs Act. The first bill is introduced into the senate, launching a five-year legislative battle<br><br>**Elixir of sulfanilamide** manufactured by SE Massengil company, containing the poisonous solvent diethylene glycol, kills 107 persons, many of whom are children, dramatizing the need to establish drug safety before marketing and to enact the pending food and drug law |
| The Federal Food, Drug, and Cosmetic (FDC) Act of 1938 | The most significant feature of the Act was that it required the drugs to be proved safe before they can be approved for marketing. Other significant features of the Act were:<br>• Extended control to cosmetics and therapeutic devices<br>• Required new drugs to be shown safe before marketingstarting a new system of drug regulation<br>• Eliminating the Sherley Amendment requirement to prove intent to defraud in drug misbranding cases<br>• Providing that safe tolerances be set for unavoidable poisonous substances<br>• Authorizing standards of identity, quality, and fill-of-container for foods<br>• Authorizing factory inspections<br>• Adding the remedy of court injunctions to the previous penalties of seizures and prosecutions |

| | |
|---|---|
| | Concept of new drug: Any drug for which there was lack of genera recognition by qualified experts as to drug's safety for intended purpose was labeled "New Drug". Created the requirement of filing a New Drug Application (NDA) for marketing such drugs. |
| Insulin Amendment Act (1941) | This amendment to FD and C Act requires FDA to test and certify purity and potency of this lifesaving drug for diabetes |
| 1943 | In **US v Dotterweich**, the Supreme Court rules that the responsible officials of a corporation, as well as the corporation itself, may be prosecuted for violations. It need not be proven that the officials intended, or even knew of, the violations |
| | This initiated the process of Debarment and Conviction of the officials responsible for violations |
| Public Health Service Act (1944) | This act covered a broad spectrum of health concerns, including regulation of biological products and control of communicable diseases |
| Durham Humphrey | Brought in the concept of "Prescription Drugs" |
| Amendment (1951) | Kinds of drugs that cannot be safely used without medical supervision were identified |
| | Law now required "Prescription Drugs" to carry the Federal Legend "CAUTION: Federal Law prohibits dispensing without medical prescription" |
| Factory Inspection | This amendment updated and clarified FDA's inspection authority |
| Amendment (1953) | Removed the need for FDA to obtain consent prior to initiating an inspection |
| | Requires FDA to give manufacturers written reports of conditions observed during |

| | |
|---|---|
| | inspections and analyses of factory samples |
| Kefauver Harris Amendment/Drug Efficacy Supplement (1962) | **Thalidomide**, a new sleeping pill, was found to have caused birth defects in thousands of babies born in western Europe. News reports on the role of Dr Frances Kelsey, FDA medical officer, in keeping the drug off the US market, arouse public support for stronger drug regulation<br><br>**Kefauver-Harris drug amendments** was then passed to ensure drug efficacy and greater drug safety. For the first time, drug manufacturers are required to prove to FDA the effectiveness of their products before marketing them. Other significant features of the amendment were:<br>• Required Manufacturers to comply with GMP requirements<br>• Made FDA approval of NDA a pre-requisite for marketing<br>• Placed prescription drug advertisement under FDA's supervision<br>• Required drug manufacturing facility registration and periodic inspections (at least once every two years)<br>• Required manufacturers to include full information on product adverse events and contraindications to provide a balanced picture for healthcare professionals, so that practitioner might safely prescribe the drug, which led to package inserts |
| DESI Review (1966) | FDA contracts with the National Academy of Sciences/National Research Council to evaluate the **effectiveness of 4,000 drugs** approved on the basis of safety alone between 1938 and 1962 |
| Fair Packaging and | Requires all consumer products in interstate |

| | |
|---|---|
| Labeling Act (1966) | commerce to be honestly and informatively labeled, with FDA enforcing provisions on foods, drugs, cosmetics, and medical devices |
| Poison Prevention Packaging Act (1970) | Required special packaging for controlled substances and prescription drugs to enhance patient safety, especially that of children |
| Drug Listing Act (1972) | This Act required firms engaged in drug manufacture, preparation, propagation, compounding or processing to register their establishments and list all of their commercially marketed drug products with FDA. The act also required each manufacturer to submit a semi-annual list of drugs introduced or discontinued since the last submission |
| Over-the-counter Drug Review (1972) | Begun to enhance the safety, effectiveness and appropriate labeling of drugs sold without prescription. Based on available data drugs were classified into class I, II and III. Class I drugs were allowed to continue marketing, class II drugs were withdrawn and class III drugs were allowed to remain till further data to establish the safety and efficacy becomes available |
| Tamper Resistant Packaging Act (1980) | Issued by FDA to prevent poisonings such as deaths from cyanide placed in Tylenol capsules. The Federal Anti-Tampering Act passed in 1983 makes it a crime to tamper with packaged consumer products<br>It was further amended in 1988 to provide two tamper resistant features to Hard Gelatin Capsules dosage forms |
| Drug Price Competition and Patent Term Restoration Act/Hatch Waxman Act (1984) | Expedited the availability of less costly generic drugs by permitting FDA to approve applications to market generic versions of brand name drugs without repeating the |

| | |
|---|---|
| | research done to prove them safe and effective. At the same time, the brand-name companies can apply for up to five years additional patent protection for the new medicines they developed to make up for time lost while their products were going through FDA's approval process Established the process for approving Abbreviated New Drug Applications (ANDA). The law required that generic manufacturers should demonstrate that the Generic product is Bioequivalent to the pioneer product (Now referred to as Reference Listed Drug, i.e. RLD) |
| Prescription Drug User Fee Act (PDUFA) | Under Pressure of complaints from various groups, which were concerned that due slow review process at FDA, new therapies are not being made available to the public in a timely manner This legislation required that Applicant's of drugs and biologics application pay fee for their applications. This fees was used to hire more reviewers to effect timely review of applications Generic Prescription drugs and OTC drugs were exempted from this requirement |

## Other Federal Laws Affecting the Industry

There are other laws with which the industrial pharmacist should be familiar. Perhaps the most important is the Fair Packaging and Labeling Act, which is administered by both the Food and Drug Administration and the Federal Trade Commission (FTC). The Fair Packaging and Labeling Act is designed primarily to protect the consumer by requiring a prominent display of statements of net contents and ingredients on the principal display panel of consumer packaging. The law has certain specific requirements concerning the placement and size of this information. Violation of the provisions of this Act might lead to seizure by FDA or to a cease and desist order from the FTC.

Often the product development chemist is involved in helping to create advertising claims and copy. To do this, he should understand the regulatory agency concerned with these matters. The FTC (Federal Trade Commission), operating under the Federal Trade commission Act, has jurisdiction over the advertising and promotion of all consumer products including drugs and cosmetics. This authority extends to all advertising media and is concerned with deceptive advertising practices and with promotion that is deemed to be false and misleading. Since there is some overlap between the jurisdictions of the FTC and the FDA, there has been an agreement between these agencies to the effect that, in general, the FTC would monitor advertising of OTC drugs and cosmetic products in so far as false and misleading claims might be concerned, whereas the FDA would have responsibility for drug labeling and all advertising relating to prescription drugs. The reason for this agreement, of course, is to avoid unnecessary duplication of enforcement procedures. The agencies have a close working relationship, with the FTC relying in large measure on FDA for its scientific expertise.

Since the industrial pharmacist is generally involved with FDA inspections, it might be interesting to note in passing that the federal statutes relating to crimes and criminal procedure make it a crime to forcibly assault, resist, oppose, impede, intimidate, or interfere with an FDA inspector during the performance of his official duties. Any such offense carries with it a fine of not more than $5,000 or imprisonment for not more than 3 years, or both. If any of the aforementioned acts entails the use of a dangerous weapon, the fine may be as high as $10,000 and the imprisonment up to 10 years or both.

The color additives amendments of 1960, which are a part of the federal food, drug and cosmetic Act, specify that no product is to contain any unsafe color additive. An unsafe color additive is defined in Section 706(a) 21 USC 376 of the Act as a drug ingredient that has been added for purposes of color only, but not in conformity with a regulation issued under the Act, or not from a batch certified in accordance with proper regulation, or not otherwise exempt from regulation by FDA.

The post office department prohibits shipments of poisons, explosives, or injurious or fraudulent materials through the mails and has jurisdiction over drugs falling within this definition. Violations are subject to seizure and fine.

The alcohol tax division of the internal revenue service has jurisdiction over the use of alcohol in the formulation of drugs. Specially denatured alcohol may be purchased for industrial use at a fully taxed price. After manufacture of the drug, using such denatured alcohol, the manufacturer may file a claim showing that it has been used in the manufacture of a product unfit for beverage purposes. This qualifies the manufacturer for a recovery of most of the tax, known as a tax “drawback”.

Should the industrial pharmacist be involved in the manufacture of narcotics or other controlled substances, such manufacture is within the ambit of the Controlled Substances Act and regulations thereunder. The drug enforcement agency (DEA) maintains a quota system for the manufacture of controlled substances and generally monitors this production by the use of established yield ratios. Periodic inspection by agents of the DEA are made for the purposes of verifying accountability and security against loss. Regulations are not only concerned with manufacture of controlled substances, but also with provisions against the diversion of these drugs.

If a hearing becomes necessary under any provision of the federal food, drug and cosmetic Act or the controlled substances Act, it is conducted under the provisions of the administrative procedure Act, unless a specific hearing procedure is set forth in the principal Act. Records, letters, memoranda, and other documents pertaining to particular drugs and/or to matters pertinent to a hearing may be requested under the provisions of the freedom of information Act. The FDA generally allows inspection of all such documents except (1) internal memoranda relating to the regulatory aspects of the administration, (2) INDs or NDAs and the material contained therein withdrawn or otherwise expired NDAs or INDs lie in a gray area, with the FDA leaning toward

disclosure, but with the overwhelming sentiment of the industry being that these should remain confidential and (3) materials marked confidential upon submission on request of the FDA by a regulated company. Such material will remain confidential unless called for by a third party, at which time the party claiming confidentiality must prove to the satisfaction of the FDA that the material is indeed confidential; the test is generally whether or not the release of such material will place the person, firm, or corporation submitting it at a competitive disadvantage.

Finally, there are several other acts that are too specialized for discussion here, but that may also be of interest to the industrial pharmacist. Their titles are fairly indicative of the subject matter of the contents:

- Animal welfare Act
- Caustic poison Act
- Controlled substances Act
- Controlled substances import and export Act
- Drug abuse office and treatment Act
- Hazardous substances Act
- Insecticide, fungicide and rodenticide Act (economic poisons Act)
- Poison prevention packaging Act of 1970 virus serum and toxin Act of 1944.

## STATE AND LOCAL REGULATIONS

The right of any governmental authority to pass laws for the protection of its citizens is an inherent one. This right is called the “police power,” and it is the basis upon which laws regulating drugs and drug products, and their manufacture, distribution, and sale are predicated. It is not unusual for such laws to exist in large municipalities as well as at the state and federal levels. These laws generally deal with misbranding and adulteration, false advertising, and the maintenance of sanitary conditions. In order to sort out the proper jurisdiction under which a particular manufacturing operation is conducted, a body of law known as “preemption” has evolved.

The doctrine of preemption means that when a greater and a lesser authority (i.e. a state and municipality or the federal government and the state) both regulate a particular activity such as the production of pharmaceutical products, and find they are in conflict, the highest authority is most often deemed to have preempted the regulatory field, and the law of the higher authority prevails.

As with all general rules, the doctrine of preemption has subrules. For example, if an operation is strictly an intrastate one, and there is no interstate commerce involved, the federal law cannot attach. Similarly, if there is an overriding “local” concern, as for example a more stringent control of the holding or transportation of narcotics for legitimate local reasons, then the more stringent local rule would apply.

Lest it be thought that a strictly intrastate business, if it embraces the practice of industrial pharmacy, might be conducted only under state law, the courts have increasingly taken the view that the federal law is applicable on a very broad scale. For example, it has been held that if any component used in the manufacture of a drug has at any time been shipped in interstate commerce, federal jurisdiction attaches. Since most components in a drug have been purchased and shipped from an out-of-state supplier, one can readily understand that an almost universal federal jurisdiction applies to the regulation of drug products.

There is a uniform state food, drug and cosmetic bill, which is patterned largely after the federal statute, and which has been endorsed by the executive committee of the association of food and drug officials of the United States. It has been adopted in many states, and with certain

modifications, those states dealing with the manufacture of drugs, even if they themselves do not use the uniform statute, lean heavily on federal law and definitions. The district of columbia is, of course, regulated by the federal Act.

Most states provide purity, labeling, and packaging requirements that are applicable to a “drug,” and these are defined most frequently in language identical or similar to that of the federal Act. All but 17 states prohibit the marketing of a new drug (which is defined in federal terms) until a new drug application has been filed with FDA and has been approved. The drug labeling requirements of each of the states, as governed by their local laws, appear in Table 30.2.

**Table 30.2:** Drug labeling requirements chart

| State | Name and and place business | Quantity of contents | Name of drug | Names of ingredients | Quantity or proportion of certain ingredients | Directions for use | Habit forming warning | Precaution against deterioration | Warnings where use may be dangerous | Special requirements for official drugs |
|---|---|---|---|---|---|---|---|---|---|---|
| AL | | X | | | X | | X | | | |
| AK | X | X | X | X | X | X | X | X | X | X |
| AZ | X | X | X | X | X | X | X | X | X | X |
| AR | X | X | X | X | X | X | X | X | X | X |
| CA | X | X | X | X | X | X | X | X | X | X |
| CO | X | X | X | X | X | X | X | X | X | X |
| CT | X | X | X | X | X | X | X | X | X | X |
| DE | | | | | X | | | | | X |
| DC | | | | | | | | | | |
| FL | X | X | X | X | X | X | X | X | X | X |
| GA | X | X | X | X | X | X | X | X | X | X |
| HI | X | X | X | X | X | X | X | X | X | X |
| ID | X | X | X | X | X | X | X | X | X | X |
| IL | X | X | X | X | X | X | X | X | X | X |
| IN | X | X | X | X | X | X | X | X | X | X |
| IA | X | X | X | X | X | X | X | X | X | X |
| KS | X | X | X | X | X | X | X | X | X | X |
| KY | X | X | X | X | X | X | X | X | X | X |
| LA | X | X | | | X | | X | X | | X |
| ME | X | X | X | X | X | | | | | |
| MD | X | X | X | X | X | X | X | X | X | X |
| MA | X | | X | X | X | X | X | X | X | X |
| MI | X | | | | | | | | | |
| MN | X | X | | | | | | | | X |
| MS | | | | | | | | | | |
| MO | X | X | X | X | X | X | X | X | X | X |
| MT | X | X | X | X | X | X | X | X | X | X |
| NE | | | | | X | | | | | X |
| NV | X | X | X | X | X | X | X | | X | X |
| NH | X | X | X | X | X | X | X | X | X | X |
| NJ | X | | X | X | X | X | X | X | X | X |
| NM | X | X | X | X | X | X | X | X | X | X |
| NY | X | X | X | X | X | X | X | | X | X |
| NC | X | X | X | X | X | X | X | X | X | X |
| ND | X | X | X | X | X | X | X | X | X | X |
| OH | X | X | X | X | X | X | X | X | X | X |
| OK | X | X | X | X | X | X | X | X | X | X |
| OR | X | X | X | X | | X | X | | X | |
| PA | V | X | X | X | X | X | X | X | X | X |
| RI | X | X | X | X | X | X | X | X | X | X |
| SC | | | | | X | | | | | |
| SD | | | | | X | | | | | |
| TN | X | X | X | X | X | X | X | X | X | X |
| TX | X | X | X | X | X | X | X | X | X | X |
| UT | X | X | X | X | X | X | X | X | X | X |
| VT | X | X | X | X | X | X | X | X | X | X |
| VA | X | X | X | X | X | X | X | | X | X |
| WA | X | X | X | X | X | X | X | X | X | X |
| WV | | | | | | | | | | |
| WI | | | | | | | | | | |
| WY | X | | X | X | X | | | | | X |

* Any specific provision of any state law as applied to a particular manufacturing plant or operation should, of course, be checked by a person qualified to make a judgment as to the applicability of the statute and its provisions.

The provisions of the Federal Act are applicable to commerce within the District of Columbia.

## FDA ENFORCEMENT POWERS

Inasmuch as the Federal Food, Drug and Cosmetic Act is a criminal statute, FDA activities are discussed in terms of enforcement. If necessary, legal action can be taken against any party violating the Act. Such action can take one of the following forms.

### *1. Seizure*

Whenever FDA believes that a drug product that has traveled in interstate commerce, or is being held for shipment into interstate commerce, is in some way violative of the Act, a so-called monition or complaint against the articles themselves is drawn up and presented to the United States attorney in the district where the goods are located. He then proceeds to make this case *ex parte* (without an adversary present) before a Federal Judge sitting in that district. The monition is an *in rem* proceeding, that is, it is a proceeding against the articles to be seized for condemnation. Since the articles themselves are the defendant in the case, the owner, usually the manufacturer, if he wishes to resist the seizure, becomes an "intervener," that is, alleging ownership, he files a claim for the goods and thus becomes the litigant. Under the law, seizure can be made in only one district of one quantity of the product unless (1) the Commissioner of Food and Drugs finds that the goods represent a significant hazard to health or (2) he finds that the drug is claimed to be a new drug, in which case multiple seizures may take place. Drugs seized can be destroyed if there is no intervenor or if intervention is not successful. After intervention, the drugs can, under proper safeguards, be released back to the intervenor for relabeling to bring them in compliance with the Act, or they can be returned to the manufacturer under bond for reworking if that will cure the misfeasance.

### *2. Injunction*

An injunction is a legal term describing a court-directed prohibition against a person, firm, or corporation doing a particular act. Generally, in the case of industrial pharmacy, it can be against the production of a drug not in compliance with GMP or the selling or shipping of such drug in interstate commerce, or against certain advertising or promotion found to be violative of the Act. The issuance of an injunction is not a function of the FDA. FDA must go to court and prove to a federal district judge that there is a violation

of the law sufficient to cause him to invoke the extraordinary remedy of injunction. The burden of proof is on FDA, and because of the extraordinary nature of the remedy, it is a heavy burden. Injunction is a regulatory tool rarely used.

### *3. Criminal Prosecution*

Criminal action may be taken against persons, corporations, or responsible officers. A penalty of $1,000 or one year in jail for an unintentional first offense may be exacted, or $10,000 and three years if the offense was committed with intent to defraud or mislead, or by a previous offender. The latter form of action is relatively rare, and would normally involve a wanton disregard of the law.

Prior to the institution of criminal proceedings, however, Section 305 of the Act provides for an informal hearing procedure whereby the manufacturer may present to an FDA hearing officer explanations for and circumstances behind the activities cited. The hearing officer may receive documentary as well as oral evidence, after which he prepares a report, uniformly dictated in the presence of the respondent. He asks the respondent to agree that the facts so stated are those presented by him, and in addition, gives the respondent a chance to add explanatory matter to the report. The hearing report, with the hearing officer’s recommendation, is then forwarded to the compliance branch in Washington for final disposition. This informal hearing procedure is important, both for the manufacturer and for the FDA. It affords the manufacturer a hearing at a level higher than that of the inspector, and it gives the FDA a chance to reevaluate its regulatory thinking in the light of such explanation. Thus, the procedure avoids unnecessary litigation and acts as a buffer against precipitous action, either by the regulating authority or by the regulated industry.

## INTELLECTUAL AND INDUSTRIAL PROPERTY

The term “intellectual and industrial property” generally is understood to refer to patents, trademarks, and copyrights. There is often uncertainty, however, regarding the rights each of these embraces and the manner in which each is obtained. There is also that nebulous something called “know-how” and that further something called “trade secrets”. Perhaps most important for the industrial pharmacist is the ability to recognize the rights, duties, and obligations inherent in each of these concepts as applied to others as well as to himself.

Briefly, *a patent* is a statutory grant of monopoly for a period of years during which an inventor can exclude others from making, using, or selling his invention. The right is absolute as long as the validity of the patent is upheld. A *trademark*, on the other hand, does not derive its value or indeed its validity from a statutory grant. The right to the exclusive use of a trademark is derived primarily from the common law in that it must be properly used by the person claiming ownership, and such use may be challenged by another even though there is a federal or state registration. *Registration* is merely a procedural tool in that it serves as notice of use and claim of ownership. A *copyright* (often confused with a trademark) is a statutory grant given to creators of artistic or literary works, so that they might enjoy the fruits of their intellectual labors for a period of years without fearing plagiarism or other copying.

*Know-how* is a term widely used to describe the accumulated knowledge of a person, a group of persons, or, in some instances, a corporate entity. It describes the general intelligence concerning either the manufacturing, or particular steps in the manufacturing, of a product or products and embraces the experience of years, largely empiric, that is generally claimed to impart particular qualities of elegance or superiority to the product manufactured in that manner. “Know-how,” if not patentable, is carefully husbanded by the owner and usually passed on only to those who need to know in order to continue the production of the product. In general, it is knowledge that is common, but that involves a particular application within the production procedure to accomplish its beneficial results.

*Trade secrets*, on the other hand, are generally those bits of knowledge or ingenuity that may well be patentable, but for reasons best known to the

keeper of the secrets, are not imparted to the general public. One explanation for the keeping of a trade secret is that if it is exposed in a patent, it can only be monopolized for a short period of years, generally 17, after which it can be used by others with impunity. As long as the secret is kept a secret, its use can be exclusive for an indefinite period. The law recognizes and protects genuine trade secrets. Each of these subjects are discussed subsequently, with emphasis being placed on patents because they represent the most likely area of industrial property and intellectual protection for the industrial pharmacist.

# Patents

## *Nature of the Grant*

In the United States, the protection of industrial property by patent has its genesis in the Constitution, Article I, Section 8, which states:

> The congress shall have power... to promote the progress of science in useful arts, by securing, for limited times to authors and inventors the exclusive right to their respective writings and discoveries.

This recognition, by those farsighted founders of this country, of the essential means of promoting industrial advance has been the backbone of technologic and economic growth in the United States since its inception. By rewarding the inventive mind with a monopoly, others are prompted to absorb the teaching and improve on it. A patent is a right granted by the United States government to an inventor to exclude others from making, using, and selling his invention for a period of years (usually 17) in exchange for the benefits to be gained by disclosing the invention to others and to have the knowledge in the public domain following the period of exclusive use.

Most countries of the world have patent laws. They vary so widely in their requirements that it would be impossible to enumerate the differences in this chapter. It is useful to know, however, that since the general rule in foreign countries is to grant a patent to the first party filing for it. Such filing should be done as rapidly as possible after the issuance of the United States Patent. Some measure of protection is afforded to the United States patent owner by an international Convention to which the United States belongs. The convention provides that a patent application filed in a foreign country within one year of its first filing in a convention country is given the earlier filing date. The following discussion relates only to United States Patent law.

## *Requirements*

To be patentable, an invention must be new, useful, and not obvious to one skilled in the art to which the invention pertains. To be new, the article must be novel, that is, of a novelty requiring more than the use of journeymen skills in the field to which the invention applies. And newness alone does not support patentability.

The requirement that the development not be obvious is the main criterion for patentability. An invention or development may not be obvious for several reasons. The results obtained from the use of the development are unexpected, that is, they are outside the general theory pertaining to the combination of the various ingredients. This concept is also related to lack of obviousness that might come about by way of synthesization, with unexpected results arising from the utilization of two or more known features in a combination resulting in something greater than a mere sum of known effects. The invention may relate to new compositions of matter as in the synthesis of a new chemical entity, which must not be an obvious variation of the prior art compounds or compositions. The invention might also be a new approach to solving technical problems or even for the production of a new plant strain.

To be useful, the invention must be able to be reduced to practice, and must perform in a predictable manner in accordance with the claims made for such performance in the patent application. For other than pharmaceutical inventions, utility is rarely a problem. The United States patent office has taken an increasingly stringent position with respect to utility data for pharmaceuticals, particularly those intended for the treatment of serious illness. The claims for utility should be written clearly and concisely with as much detail as possible to prove that the product will perform in a uniform manner with a predictable result. Patent coverage can be obtained on compositions of matter, apparatuses, new processes of manufacture, new article of manufacture, and asexually reproduced plants.

For the sake of completeness, it should be mentioned that design patents may be issued. These are generally for nonfunctional elements of an article of manufacture. The period of design patent exclusivity is much shorter, running usually for a period of 7 years.

Finally, in addition to being new, useful, and not obvious, the invention must not be barred by having been previously known, used, or sold publicly in the United States, or patented or described in a publication anywhere in the world before the invention by the applicant, or more than one year prior to the filing of the application obtaining a patent.

Since strong proof of the date of conception is the inventor’s best insurance for obtaining his patent, he should keep careful notes, preferably in a bound notebook, giving a brief description of the invention and the manner

in which it is expected to work. This record should be dated by the inventor, and should be witnessed by at least one other person who is able to understand the invention. Often, in an interference or a litigation involving two or more claimants, this early record can be controlling.

An *interference* is a procedural device used by the patent office and is declared when different inventors have applied for the same invention. Since uncorroborated testimony of the inventor as to the date of invention is not admissible, a failure to make early disclosure records as noted above can lead to the loss of valuable rights. The attesting witness might have to give testimony at a later date, and therefore the fullest explanation of the invention and the manner in which it operates should be given to the witness, since he will later have to describe in detail that which was explained to him.

Although it is not necessary to reduce an invention to practice, there must be in the application enough detail to enable one skilled in the art to practice it. Patent applications for pharmaceuticals generally are not filed until the utility of the pharmaceutical has been determined, and this, in a sense, is “reduction to practice”.

When an invention is made and sufficient data gathered to prove it useful, the next step should be a disclosure to a patent attorney, who will then follow several prescribed steps to obtain the patent. He will usually want to do a search of the prior art in an effort to determine novelty. Such a search is not required by the rules of the patent office, but a careful patent attorney will want the information so as not to waste the time and money of his client or the time of the patent examiner.

Once satisfied that the invention has all the elements of patentability, the patent attorney draws up the application, the principal portions of which are the specification and the claims. The specification is a detailed description of the invention, including general and preferred versions, with necessary supporting data when applicable. A specification also may include drawings, graphs, or any other aids to the understanding of the invention. If the prior art search has uncovered references, these may also be included in the specification, and distinguishing features of the invention for which patent is sought should be set forth. All of this is designed to enable one skilled in the art to understand the invention and to employ it.

Perhaps the most critical part of the application is the section containing claims, which-appear as a series of numbered paragraphs setting forth the

scope of the patent protection sought. These claims are important because, when a patent is granted, the claims form the basis of exclusivity and determine whether another has infringed on the issued patent. Beginning with the first of the numbered paragraphs, claims are presented in the order of the broadest protection sought to the narrowest embodiment of the invention. Since the claims determine infringement, they should include all of the elements of the invention, because no patent protection is accorded subject matter disclosed in the specification, only subject matter specifically claimed in the claims section.

Once all specifications and claims are written, the necessary formal documents, including an oath by the inventor, are completed, an attorney is appointed, and the application is filed with the Patent Office and receives an application number and a filing date. It is thereafter examined by a patent examiner, who is generally a person trained or skilled in the art relating to the invention. Should questions arise in the mind of the examiner, he asks in writing for clarification or limitation or other remedial steps. These are called office actions, to which responses by the inventor are invited-indeed, are mandatory if the patent is to be issued. Following these actions, the application is allowed in all of its parts, in some of its parts, or in a restricted way, or it is finally rejected. There is an appeal procedure from final rejection, but if it is not successful, the application either becomes abandoned or may be further appealed to the Patent Office Board of Appeals, which is the final agency review. The board is a tribunal of three highly qualified patent office experts who pass on the merits of the examiner’s and the applicant’s positions. The decision of the board may be appealed to the court of customs and patent Appeals or to any United States District Court having jurisdiction. Elapsed time from application to allowance without an appeal is generally considered to be more than two years but less than four. Appeals, of course, may make the time span longer.

## *General Comments*

Since the benefits derived from a patent are manifold, certain restrictions must be kept in mind both by the inventor and by those who would use the invention. A patent grants the right to exclude others from making, using, and selling the invention, but this right may be restricted if the patent is subordinate to another. For instance, a new chemical entity that has been granted a patent may be taken by a second inventor and combined with

another compound to form a new and useful compound, the properties of which were not obvious to those skilled in the art. Thus, to practice the teaching of the second invention, one would of necessity have to reward the original inventor.

Licensing is a system of rewarding the original inventor and to otherwise authorize the practice of the invention by persons other than the inventor. The license grants the right to make, to use, or to sell the invention, or any appropriate combination of these rights. A licensor is usually given periodic payments, called royalties, for the term of the patent. Of course, it is also possible, after negotiation, to purchase the rights to a patent from the inventory by an outright lump sum payment.

Patent licensing is a delicate endeavor today, since there have been decisions in the antitrust area which tend to restrict the types of license that might be granted. For instance, it has been held in certain cases to be a violation of the antitrust laws to license an invention in only one area of the marketplace and not in another-for example, to license the use of the invention in oral dosage forms to one licensee and in parenteral dosage forms to another licensee. Inattention to these important interdictions (which at this writing tend to become more and more restrictive) could lead to the imposition of fines, both on the licensor and the licensee, or in more aggravated cases, the loss of patent protection entirely.

Applications for patents in the United States are submitted in the name of the inventor, and he personally receives the grant of patent. In the larger industrial centers, including pharmaceutical manufacturers, it is common practice for the inventor to assign the patent application to his employer in consideration of his continued employment, that is, inasmuch as he was hired to do the research and was given the facilities to do it by the employer, he is expected to make the proper assignment. Usually, a contract covers this right of the employer to receive and benefit from the inventions of his employee. When no such contract exists and the inventor balks at an assignment, a legal doctrine of “shop right” becomes operative. This is a right that belongs to the employer, who has implicitly been given the right to use the invention because he supplied the necessary support for the inventor, and allows him to use the patent even though title remains in the name of the inventor.

If there is more than one inventor, that is, if two or more individuals have contributed to the conception and/or the reduction to practice of the

invention, each contributor is separately named and becomes a “co-inventor”. The important thing to remember about co-inventors is that each may separately exploit the patent, unless there is a contractual agreement among them to act in concert. Therefore, one having an interest as a licensee in a patent for which there are co-inventors, especially if that interest is exclusive, should either obtain a license from all or have the noncontracting inventors agree on the licensing document to the exclusivity of the license granted.

An inventor can stop an infringement by applying to a Federal court for relief, which is generally available in two forms. One is *injunction*, that is, a judicially imposed restriction against the infringer from continuing to make, use, or sell the invention; the other involves an award of damages to the inventor for past infringement.

Since there can be no infringement prior to the issuing of the patent, the inventor who wishes to exclude others even during the pendancy of his application can mark his product with the phrases “Patent Pending” or “Patent Applied For”. These phrases serve as warnings to putative infringers that they make, use, or sell the invention at their peril. Notices on goods that bear patent numbers, of course, refer to already issued patents, and infringement after issue can result in damages generally based on the measure of sales lost by the patent holder.

Finally, it should not be overlooked that issued patents in any particular field are a rich source of research data. Further, because countries foreign to the United States may issue patents earlier than the US Patent Office, foreign inventions should be monitored, as they are often the first indication of new technologic developments in the United States.

## Trademarks

A trademark, unlike a patent, is not a statutory grant of monopoly. A trademark is any word, name, symbol, or device that a manufacturer or trader places on his goods to distinguish them from goods manufactured or sold by others. The validity of a trademark is tested by its proper use, not by the fact of registration. The registration of a trademark may be accomplished either in a state having a trademark registration statute or at the Federal level. To be registered in the United States Patent Office, the trademark must be used on goods that have actually travelled in interstate commerce.

Thus, the application for a trademark must await the first use of that mark in interstate commerce, the date of which, along with the date of its first use anywhere, must appear on the application. The trademark must not be descriptive of the goods or any of its properties (although it may be “suggestive”); it must not be so similar to prior registrations as to cause confusion as to the source of origin of the goods; it must not be immoral or deceptive; and it must not utilize the flag or insignia of the United States or the portrait or likeness of any living individual without that individual’s consent. Trademark protection extends over a period of 20 years, and is renewable indefinitely as long as the trademark continues to be used by its owner. Ownership of a trademark is transferable by assignment, and the use thereof is able to be licensed.

A trademark represents good will. The considerable value of such good will often cannot be determined in units of absolute dollars. For instance, if by some quirk of massive misfortune, all the Coca-Cola bottling plants in the United States should burn down today, there would be no question that because of the extensive good will in that trademark alone, the necessary financing could be had to begin to rebuild the plants tomorrow.

A registered trademark may be infringed by others who use the same mark or a similar mark on the same or similar goods or on goods that travel in the same channels of trade. The legal determination for an infringement is whether the second user of the mark confuses the buying public regarding the originator of the goods on which the mark appears.

Since the axiom is that trademark rights are gained only by proper use, it is possible to have an infringement of an unregistered mark. An action for such infringement is generally couched in terms of the law of unfair

competition, of which the trademark laws have uniformly been held to be a subdivision.

A word trademark is a name; it is a special part of the language having unusual characteristics. It should always be used in conjunction with a generic or descriptive term, for instance, Librium, brand of chlordiazepoxide. The use of a descriptive or generic term preserves the function of the trademark as a symbol pointing to the source of origin and not as a descriptive term for the article itself. The trademark might be rendered invalid if it is not used at all or if it is used improperly, that is, if it falls into the common language to describe the article. Prime examples of trademarks that have been lost by such improper use are aspirin, cellophane, and escalator. Trademark owners jealously guard their rights, and more popular trademarks become the subject of letters to publishers or printers to remind them that the trademark should always be printed either in all caps or at least in initial caps, generally with quotation marks and never in the possessive.

A claim to trademark rights is generally printed by the claimant on the goods and in advertisements relating to the goods. Prior to Federal registration, the symbol “TM” is generally used in conjunction with the mark. After registration, the symbol ® or the phrases “Trademark Registered,” “Trademark Registered in the United States Patent Office,” or “TM Reg. US Pat. Off.” are used. These notices serve, as with patent notices, to tell would-be infringers that they adopt this mark at their peril. Infringement and unfair competition actions may be brought in United States District Courts in the case of Federally registered trademarks or in state courts on counts of unfair competition. Damages are awarded against infringers based on an assessment of the damage that the infringement has caused to the goodwill of the trademark owner.

A trade name, for which registration also is permitted, is distinguishable from a trademark in that the trade name identifies the manufacturer or trader rather than the goods themselves. For instance, Schering-Plough Inc. is a trade name identifying a manufacturer; Coricidin is that manufacturer’s well-known brand name for a product for the relief of the symptoms of colds.

## Unfair Competition

The law of unfair competition is almost as old as the common law itself. Essentially, it is the effort of the law to referee in the marketplace and to maintain fair and open competition among traders. It covers a vast array of commercial wrongs, some of which, e.g. the Patent and Trademark laws, have been codified. Complaints in the area of unfair competition generally seek relief under the principles of both equity and law. That is, they seek specific performance or specific nonperformance (injunction), which is an outgrowth of the ancient laws of equity, or they seek money damages (law), or both. In all of the jurisdictions in the United States today, both law and equity relief can be sought in one court rather than in separate courts, which our common law adopted from the ancient English practice, but which has now been uniformly replaced by the single jurisdiction system.

One possible grievance in an unfair competition action is interference with a contractual relationship. Such interference could take the form of commercial disparagement, that is, disparagement of the name of a company or of its product; the hiring of an employee to gain insight into the prior employer's business practices or customers; or some other inducement of a breach of contract by one party to the detriment of another.

The law of unfair competition also provides for protection of know-how and trade secrets, which are as much a property right as a patent or a trademark. Although the theft or disclosure of these is actionable, anyone innocently receiving knowledge of a trade secret or know-how in the usual course of his business may practice such knowledge with impunity. This points up the peril of attempting to keep commercial secrets, since their only protection lies in their freedom from theft. For this reason, when secret information is received or given, it is usually done on the basis of a "confidential disclosure." It should be given in writing, with the receiver acknowledging and promising to maintain its confidentiality.

Finally, the law of unfair competition prevents the palming-off of the goods of one trader as being those of another, as in false advertising and in misleading statements made by sales personnel. Recent decisions have narrowed the scope of this type of protection in that elements of an article not protected by United States patents or trademark registrations may be copied with impunity. Thus, nonfunctional features, even though they are unique

with the original user and might serve to indicate origin, cannot be held exclusively unless the originator can show that these nonfunctional features have gained “secondary meaning,” and are attributed in the minds of purchasers with the originator only and with no one else. In other words, it serves to identify the goods as coming from a particular manufacturer or trader.

## Copyrights

A copyright is a statutory grant of monopoly for a period of years. Generally, it is for 27 years renewable for another 27 years by the copyright owner or his heirs or assigns. After that, the copyrighted material falls into the public domain and can be used by anyone.

Copyright protection may be granted for works of art and literature, and a copyright carries with it, as does a patent, the right to exclude others from reproducing the work of art or literature without compensation to the copyright owner. A copyright is obtained by the filing of an application on which is stated the nature of the work; in the case of published literature or music, a copy is filed with the Library of Congress. For works of art, such as painting and sculpture, generally a picture will suffice. The danger inherent in the publication of a copyrightable work lies in the fact that rights gained through copyright might be lost if a proper copyright notice is not placed on the article *before* it is published or otherwise made public. Thus, a musical composition performed before an audience or a book disseminated widely without a printed claim of copyright almost invariably causes that work of art to fall immediately into the public domain. Since it is so easy to lose copyright rights, competent advice should be sought whenever a new work of art or literature is to be made available for public sale or use. The Copyright Office in the Library of Congress supplies the forms necessary for copyright application, but since application is an after-the-fact event, advice concerning proper notice should be sought at an early stage in the development of the work of art or literature.

Copyright rights are commonly claimed by the use of the phrase: Copyright, year, name of the copyright claimant. The symbol © might also be used as: © year, name of claimant.

## PRODUCT LIABILITY

The manufacture of a product that is to be purchased for use or consumption carries with it the responsibility to make that product as free from hazard as possible. The law of product liability is a common law concept. Broadly stated, it gives redress to a purchaser based on the concept of breach of contract or of negligence, or some combination of the two. The breach of contract, a legal fiction, holds that there is an implied warranty of merchantability of the goods for their intended use, that is, if used in the manner prescribed by the manufacturer or seller, there will be no hazard inuring to the purchaser or user. Increasingly, state laws are being so broadened as to remove the need for a direct relationship (as at the old common law) between the purchaser and the seller. The anomalous situation, for instance, wherein a child hurt by a product bought by a parent could not recover because the purchase was not made by the person injured (the privity doctrine), has largely been abandoned. The implication of liability has been so broadened that virtually anyone using the product and being hurt can, upon a showing that the product was properly used (or at least not abused), be compensated for the injury.

The other concept is that of negligence, that is, the product or food was manufactured in such a way that it contained a defect that should have been eliminated or discovered by the manufacturer. Thus, the article, when placed in the hands of the consumer, carries with it an inherent danger even if properly used. The burden on pharmaceutical manufacturers is especially heavy, their product by its very nature being potentially dangerous; however, strict adherence to current good manufacturing practices and stringent quality control are the pharmaceutical manufacturer’s best protection.

## RECALLS

While the FDA has no statutory authority to require the recall of a pharmaceutical, the agency nevertheless has published, in the interest of public safety, extensive regulations concerning recalls and has set up a detailed mechanism for the recall of drugs from the marketplace. Recalling a drug is one of the most difficult decisions facing a manufacturer or distributor, for such an Act has important ramifications, ranging from a company's relationships with the FDA to potential product liability litigation. Therefore, recalls should be entered into only after the most careful evaluation of the problem, and once begun, they must be completed as rapidly and efficiently as possible. FDA regulations provide extensive guidelines to determine (1) what products should be recalled (and to what level) and (2) the precise mechanism for recall that would most likely achieve the desired goal. In addition to these regulations, the FDA has also made available its own regulatory procedures manual, which defines in even greater detail the recall mechanism and FDA's involvement in it. All companies manufacturing or distributing pharmaceutical products should have officials and employees that are fully informed as to these regulations and FDA procedures, so that accurate and appropriate decisions can be made.

The FDA classifies all recalls into one of three categories. Category I recall represents an emergency situation in which the drug poses a hazard that is immediate, long-range, and life-threatening. Such recalls require special mechanisms and a close relationship with the FDA, since a public warning must be issued, and the product must be recalled to the consumer level.

The second type of recall is Category II, a priority situation in which the consequences of the offending drug remaining on the market may be immediate, long-range, or potentially life-threatening. In general, such recalls must be made to the retail or dispensing level, and occasionally, the FDA issues press releases to inform the public. The final recall classification is Category III. This is a routine situation in which the threat of life is remote or nonexistent. Such recalls include products that are technically illegal because they are adulterated or misbranded in some respect or another, for example, labeling violations not involving a health hazard. Such recalls are required only to the wholesale level, and press releases are usually not issued.

In addition to such formal recall mechanisms, there are occasions that

require companies to remove products from the market when there are no violations of the law involved. This is often done for reasons totally unrelated to the FDA or to the integrity of the product. It is important for companies to appreciate that such market withdrawals are not recalls and are not subject to FDA regulations.

Removal of a product that is still entirely under the direct control of the manufacturer, even though it may have been shipped interstate to one or more branch warehouses, also is not classified as a recall, provided no stock has been distributed to the trade. These removals are termed “stock recoveries” and do not appear on the public recall list. Usually, however, checks are made on the adequacy of these retrievals to cover ultimate disposition of the merchandise.

# Appendix

# WEB ADDRESS OF WORLDWIDE REGULATORY BODIES

| Country | Web address | Regulatory body |
|---|---|---|
| **Worldwide** | http://www.ich.org | The International Conference on Harmonisation |
| **Worldwide** | http://www.who.org | World Health Organization |
| **Australia** | http://www.tga.gov.au | Department of Health and Ageing Therapeutic Goods Administration |
| **Argentina** | http://www.anmat.gov.ar | Administracion Nacional de Medicamentos, Alimentosy Tecnologia Medica |
| **Austria** | http://www.bmsk.gv.at/cms/siteEN | Federal Ministry of Labour, Social affairs and Consumer Protection |
| **Belgium** | https://portal.health.fgov.be/portal/ | Ministry of Social affairs, Public Health and Environment |
| **Brazil** | http://w3.datasus.gov.br/datasus/index.php | Departamento De Informatica Do SUS |
| **Brazil** | http://www.anvisa.gov.br/eng/index.htm | National Health Surveillance Agency (Anvisa; Agencia Nacional de Vigilancia Sanitaria) |
| **Canada** | http://www.hc-sc.gc.ca/index-eng.php | Health Canada |

| | | |
|---|---|---|
| **Chile** | http://www.minsal.cl/ | Ministerio de Salud, Gobierno De Chile |
| **Columbia** | http://www.minsalud.gov.co/ | Ministry of Health |
| **Cost Rica** | http://www.netsalud.sa.cr/ms/ | Ministry of Health |
| **Czech Republic** | http://www.sukl.cz/index.php?lchan=1andlred=1 | State Institute for Drug control |
| **Denmark** | http://www.dkma.dk/1024/visUKLAnsi.asp?artikelID=728 | Danish Medicines Agency |
| **Estonia** | http://www.sam.ee/ | State Agency of Medicines is a governmental body under the Ministry of Social Affairs (Ravimiamet;State agency of Medicines) |
| **Europe** | http://www.eudra.org/en_home.htm | EMEA: The European Agency for the Evaluation of Medicinal Products. |
| | http://www.hma.eu/ | Heads of Medicines Agencies is a network of the Heads of the National Competent Authorities whose organisations are responsible for the regulation of Medicinal Products for human and veterinary use in the European Economic Area |
| **Finland** | http://www.nam.fi/ | Laakelaitos Lakemedelsverket (National Agency for Medicines) |

| | | |
|---|---|---|
| **France** | http://agmed.sante.gouv.fr/ | Agency for Medicine |
| **Germany** | http://www.bfarm.de/cln_028/DE/Home/startseite__node.xhtml__nnn=true | Bundesinstitu Fur Arzneimittel und Medizinprodukte (Federal Institute for Drugs and Medical Devices) |
| | http://www.bmg.bund.de/EN/Ministerium/ministry.xhtml__nnn=true | The Federal Ministry of Health |
| **Guam** | http://www.admin.gov.gu/pubhealth | Department of Public Health and Social Services |
| **Hong Kong** | http://www.psdh.gov.hk/eps/webpage.jsp | Pharmaceutical Service Department of Health, The Government of Hong Kong Special Administrative Region |
| **Hungary** | http://www.ogyi.hu/main_page/ | National Institute of Pharmacy |
| **Iceland** | http://eng.heilbrigdisraduneyti.is/ | Ministry of Health |
| **Indonesia** | http://www.depkes.go.id/en/index_en.htm | Departemen Kesehatan Republic Indonesia |
| **India** | http://cdsco.nic.in | Central Drugs Standard Control Organization under Ministry of Health and Family Welfare |
| **Ireland** | http://www.dohc.ie/ | Department of Health and Children |
| | http://www.imb.ie/ | Irish Medicines Board |
| **Israel** | http://www.health.gov.il/ | Ministry of Health |
| **Italy** | http://www.sanita.it/farmaci/ | Ministry of Health |
| **Japan** | http://www.mhw.go.jp/english/index.html | Ministry of Health and Family Welfare |

| | | |
|---|---|---|
| **Jordan** | http://www.nis.gov.jo/infres/owa/Section_21_2?v_Main_code=12 | Ministry of Health |
| **Kenya** | http://www.kenyastatehouse.go.ke/Ministry/ministry_health.htm#top | Ministry of Health |
| **Latvia** | http://www.vza.gov.lv/ | State agency of medicines |
| **Lithuania** | http://www.vvkt.lt/defaeng.htm | State medicines control agency |
| **Luxembourg** | http://www.etat.lu/MS/ | Ministry of Health |
| **Malaysia** | http://dph.gov.my/ | Department of Public Health |
| **Mexico** | http://www.ssa.gob.mx/ | Ministry of Health |
| **Morocco** | http://www.sante.gov.ma/ | Ministry of Health |
| **Netherlands** | http://www.cbg-meb.nl/ | Medicines Evaluation Board |
| **New Zealand** | http://www.medsafe.govt.nz/ | Medicines and Medical Devices Safety Authority |
| **Nordic** | http://www.nln.se/ | Nordic Council of Medicines |
| **Norway** | http://odin.dep.no/shd/engelsk/index-b-n-a.xhtml | Ministry of Health and Social Welfare |
| **Philippines** | http://www.pchrd.dost.gov.ph/ | Philipine council for Health Research and Development (PCHRD) |
| **Poland** | http://www.mzios.gov.pl/ | Ministry of Health and Social Welfare |
| **Portugal** | http://www.infarmed.pt/ | The national Institute of Pharmacy and Medicines (Infarmed) |
| **Russia** | http://views.vcu.edu/views/fap/medbio/medbio.htm | Public Health Institute |
| **Slovak Republic** | http://www.health.gov.sk/<br>http://www.sukl.sk/ | Ministry of Health<br>State Institute for Drug |

| | | |
|---|---|---|
| | | Control (SIDC) |
| **Slovania** | http://www.gov.si/mz/ur-zdrav/uzintang.htm | Ministry of Health Agency for Medicinal Products |
| **Spain** | http://www.msc.es/ | Ministry of Health |
| **Switzerland** | http://www.swissmedic.ch/ | Swiss Agency for Therapeutic Products |
| **Taiwan** | http://www.mpa.se/ | Department of Health |
| **Thailand** | http://www.moph.go.th/ | Ministry of Public Health |
| **Turkey** | http://www.health.gov.tr/ | Ministry of Health |
| **UAE** | http://www.moh.gov.ae/ | Ministry of Health |
| **UK** | http://www.mhra.gov.uk/ | Medicines and Healthcare Products Regulatory Agency |
| **USA** | http://www.fda.gov/ | Food and Drug Administration |

# OTHER IMPORTANT WEB PORTALS

| Web address | Remark |
|---|---|
| **www.aaps.org** | Information on officers, activities and membership from the American Association of Pharmaceutical Scientists. |
| **www.abpi.org.uk/_private/welcome/default.htm** | The Association of the British Pharmaceutical Industry for companies producing prescription medicines. |
| **www.wizard.pharm.wayne.edu** | American Chemical Society Division of Medicinal Chemistry is an organization for scientists involved in drug research and development. |
| **www.who.int/dap/ DAP_Homepage.xhtml** | Information on this World Health Organization program supporting and coordinating comprehensive national drug policies. |
| **www.aacp.org** | American body representing the interests of pharmaceutical educators and education. With activity details and a software library. |
| **www.bpsweb.org** | BPS recognizes specialties in pharmacy and provides advanced practice specialty certification for qualifying pharmacists. |
| **www.ualberta.ca/~csps/** | Canadian Society for Pharmaceutical Sciences is a organization established to foster research. Includes journal articles and links. |

| | |
|---|---|
| **www.iacprx.org** | International Academy of Compounding Pharmacists, a professional body with members in the US, Canada, Australia and Chile. |
| **www.crsadmhdq.org** | International society focusing on the advancement of science and technology of drug delivery systems. |
| **www.fda.gov** | News, publications and research reports, with links to other sites. Includes the Center for Drug Evaluation and Research. |
| **www.ndmainfo.org** | Hosted by the Nonprescription Drug Manufacturers Association.With news, information on the industry, facts and figures, and info on OTC issues. |
| **www.npa.co.uk** | Organization of Britain's community and retail pharmacies.Information on services, publications and membership benefits. |
| **www.pharmweb.net** | Guide to pharmacy and related resources on the Internet. Includes jobs, forums and directories of individuals and organizations. |
| **http://www.geocities.com/pharmalinks/** | Directory of pharmaceutical companies and institutions, with information on products and services, research, industry and government.pharmacology.miningco.com COB=looksmart Get facts on legal drugs, trials, and treating specific ailments. Search databases geared |

| | |
|---|---|
| | toward consumers and professionals. |
| **www.sci.lib.uci.edu/~martindale/Pharmacy.xhtml** | Huge website with pharmacy, pharmacology, clinical pharmacology and toxicology information. Features journals, student resources and databases. |
| **www.fda.gov** | Food and drug administration of USA deals with all prescription/non-prescription human medicines, animal products and dietary supplements sold across USA. |
| **http://www.medsch.wisc.edu/clinsci/antigram.xhtml** | Guidelines for the selection of antimicrobials and their cost-effective use in hospitals. With advice to maximize patient care. |
| **www.bio.com/bio.xhtml** | Directory of pharmaceutical and biotechnology companies and institutes, with information on products and services, research,industry and government. |
| **info.cas.org/welcome.xhtml** | Guides, tutorials, articles, software and research tools for molecular modeling, from this center which works with NIH on research. |
| **chrom.tutms.tut.ac.jp/JINNO/DRUGDATA/00database.xhtml** | Drug database with function, with details of their chemical composition, physical properties and UV spectrum. |
| **pharminfo.com/drugdb/db_mnu.xhtml** | Information from this database includes recent announcements and warnings about side effects of drugs, written mainly for professionals. |

| | |
|---|---|
| **www.dd-database.org** | Searchable bibliography of abstracts, journal articles, and books on this method for studying the effects of drugs. |
| **http://www.druginfonet.com** | Information resource for consumers and professionals on health care and pharmaceutical-related businesses. |
| **www.merck.com/pubs/mmanual/** | Manual of diagnosis and therapy covering most disorders,describing symptoms, common clinical procedures, and drug treatment. |
| **http://dtp.nci.nih.gov/epnrd/3dis.xhtml** | National Cancer Institute provides this collection of 3D structures for over 400,000 drugs. |
| **www.phrma.org** | Organization representing America's pharmaceutical research companies provides details of drug development, industry news,and health guides. |
| **www.bio.net/hypermail/PHARMA-CEUTICAL-BIOTECHNOLOGY/** | Search this newsgroup archive by keyword or phrase for information on pharmaceuticals and biotechnolgy. |
| **www.boomer.org/pkin/** | Information and resources about these two disciplines studying the dynamics of drug and metabolite levels in the body. |
| **www.exit109.com/~zaweb/pjp/pharm.htm** | This guide to pharmacy resources on the Internet gives a description of links and includes mailing lists. |
| **www.medmarket.com/tenants/rainfo.htm** | Find any regulatory body, organization, meeting, document or guideline in this extensive directory. Covers regulated product industries. |
| **www.rxlist.com** | First stop for finding informative |

| | |
|---|---|
| | reports about almost any drug.Lists over 4,000 drugs and common side effects. |
| **www.li.net/~edhayes/rxschool.xhtml** | US pharmacy schools listed by state. With their address, telephone numbers and Internet links. |
| **members.tripod.com/~ChristopherMarres/** | Internet resource for drug development and outsourcing information for pharmaceutical professionals. |
| **www.pharmacy.org/** | Join in the live discussions of pharmacy and related topics on Internet Relay Chat. |
| **www.ifpw.com** | International Federation of Pharmaceutical Wholesalers provides industry information to pharmaceutical and related companies. |
| **www.pharmpro.com** | Database of new technology for the pharmaceutical industry featuring new products and product directory. |
| **www.arcwebserv.com/pharm/** | Weekly pharmacy articles by registered pharmacists, advice about careers, tips, and a useful database of links for students doing research. |
| **www.usc.edu/hsc/lab_apk/** | Information, software and event details pertaining to pharmacokinetic systems and individualized drug therapy. |
| **http://157.142.72.143/ gaps/pkbio/pkbio.xhtml** | Online course pharmacokinetics and biopharmaceutics, covering drug administration, data analysis, and modeling systems. |
| **www.drugs.indiana.edu/pubs/newsline/searching.xhtml** | Covers searching and offering strategies for locating accurate and |

| | |
|---|---|
| | scientifically accepted information on pharmaceuticals. |
| **www.rxlist.com** | First stop for finding informative reports about almost any drug.Lists over 4,000 drugs and common side effects. |
| **www.usp.org/standard/9901.htm** | Official newsletter from the United States Pharmacopeia provides details of new drug regulations and standards plus member news. |
| **www.worldpharmaweb.com** | Pharmaceutical industry news service has stories from around the world updated daily, plus contact information. |
| **www.pharmaceuticalonline.com** | Source for product, technology, regulatory, and management information for the pharmaceutical industries. |
| **http://pkpd.icon.paloalto.med.va.gov** | Repository of software, computer controlled drug administration programs and simulation programs, for these specialty areas. |
| **www.fda.gov/cder/** | Find out information about new drugs, clinical trials and the process of regulation from the Food and Drug Administration. |
| **www.pharmweb.net/pwmirror/pwk/pharmwebx.html** | Listing of international pharmaceutical regulatory bodies including the US Food and Drug Administration. |
| **http://pharmacology.tqn.com/library/newdrug/blindx.htm** | Search for information about newly-approved or newly-released medicines in the United States. |
| **www.medmarket.com/tenants/rainfo/rainfo.htm** | Finland's regulatory body, |

| | |
|---|---|
| | organization, meeting, document or guideline in this extensive directory. Covers regulated product industries. |
| **www.rsi-nc.com** | Organization that offers clinical monitoring, auditing, and training relevant to clinical drug research. With links to other drug trial sites. |
| **www.cvmg.com.** | Clinical research facility provides access to case reviews. With information on specific research and trials. |
| **www.ama-assn.org/scipub.htm** | Keep critically informed by registering for the AMA's free online news service and journals. |
| **amanda.uams.edu/index2.0.xhtml** | Molecular modeling and visualization information from this center using computational techniques for research and education. |
| **www.ualberta.ca/~csps/** | Canadian Society for Pharmaceutical Sciences is a organization established to foster research. Includes journal articles and links. |
| **http://ourworld.compuserve.com/homepages/datagen/** | Clinical/data management organization which focuses on medical informatics, clinical data management and statistical consulting. |
| **www.phrma.org** | Organization representing pharmaceutical research companies. With details of drug development and health guides. |

# SOLVENT MISCIBILITY TABLE

| Dielectric constant | UV (nm) cut-off | Refractive index | Viscosity (°C) | Boiling point (°C) | Solvent | iso-Octane | n-Hexane | n-Heptane | di-Ethyl ether | Cyclohexane | Ethyl acetate | Toluene | Chloroform | Tetrahydrofuran | Benzene | Acetone | Dichloromethane | Dioxane | n-Propanol | Ethanol | Dimethylformamide | Acetonitrile | Acetic acid | Dimethyl sulfoxide | Methanol | Water |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1.94 | 210 | 1.3914 | 0.50 | 98 | iso-Octane | | | | | | | | | | | | | | | | | | | | | |
| 1.88 | 195 | 1.3749 | 0.31 | 69 | n-Hexane | | | | | | | | | | | | | | | | | | | | | |
| 1.92 | 200 | 1.3876 | 0.41 | 98 | n-Heptane | | | | | | | | | | | | | | | | | | | | | |
| 4.33 | 218 | 1.3524 | 0.24 | 34 | di-Ethyl ether | | | | | | | | | | | | | | | | | | | | | |
| 2.02 | 200 | 1.4262 | 1.00 | 80 | Cyclohexane | | | | | | | | | | | | | | | | | | | | | |
| 6.02 | 256 | 1.3724 | 0.45 | 76 | Ethyl acetate | | | | | | | | | | | | | | | | | | | | | |
| 2.38 | 284 | 1.4969 | 0.59 | 110 | Toluene | | | | | | | | | | | | | | | | | | | | | |
| 4.81 | 245 | 1.4458 | 0.57 | 60 | Chloroform | | | | | | | | | | | | | | | | | | | | | |
| 7.58 | 212 | 1.4072 | 0.55 | 65 | Tetrahydrofuran | | | | | | | | | | | | | | | | | | | | | |
| 2.27 | 278 | 1.5011 | 0.65 | 80 | Benzene | | | | | | | | | | | | | | | | | | | | | |
| 20.70 | 330 | 1.3587 | 0.36 | 56 | Acetone | | | | | | | | | | | | | | | | | | | | | |
| 8.93 | 233 | 1.4241 | 0.44 | 40 | Dichloromethane | | | | | | | | | | | | | | | | | | | | | |
| 2.25 | 215 | 1.4224 | 1.37 | 101 | Dioxane | | | | | | | | | | | | | | | | | | | | | |
| 20.33 | 210 | 1.3856 | 2.30 | 98 | n-Propanol | | | | | | | | | | | | | | | | | | | | | |
| 25.80 | 210 | 1.3610 | 1.20 | 78 | Ethanol | | | | | | | | | | | | | | | | | | | | | |
| 36.70 | 268 | 1.4305 | 0.92 | 153 | Dimethylformamide | X | X | X | | | | | | | | | | | | | | | | | | |
| 37.50 | 190 | 1.3411 | 0.38 | 82 | Acetonitrile | X | X | X | | X | | | | | | | | | | | | | | | | |
| 6.30 | 230 | 1.3720 | 1.26 | 118 | Acetic acid | | | | | | | | | | | | | | | | | | | | | |
| 4.70 | 260 | 1.4830 | 2.24 | 189 | Dimethyl sulfoxide | X | X | X | X | | | | | | | | | | | | | | | | | |
| 32.70 | 205 | 1.3284 | 0.55 | 65 | Methanol | X | X | X | | X | | | | | | | | | | | | | | | | |
| 81.10 | 210 | 1.3330 | 1.00 | 100 | Water | X | X | X | X | X | X | X | X | | X | | X | | | | | | | | | |

X - Immiscible (in some proportions two phases will be produced)

# LIMITS OF RESIDUAL SOLVENT

Residual solvents in pharmaceuticals (according to ICH Q3C guidelines) are defined as organic volatile chemicals that are used or produced in the manufacture of drug substances or excipients, or in the preparation of drug products. The solvents are not completely removed by practical manufacturing techniques. Residual solvents were evaluated for their possible risk to human health and placed into one of three classes as follows:

## Class 1 Solvents: Solvents to be Avoided

Known human carcinogens, strongly suspected human carcinogens, and environmental hazards.

## Class 2 Solvents: Solvents to be Limited

Non-genotoxic animal carcinogens or possible causative agents of other irreversible toxicity such as neurotoxicity or teratogenicity.

## Class 3 Solvents: Solvents with Low Toxic Potential

No health-based exposure limit is needed. Class 3 solvents have permitted daily exposure (PDE) of 50 mg or more per day.

**Table 1:** Class 1 solvents in pharmaceutical products (solvents that should be avoided)

| Solvent | Concentration limit (ppm) | Concern |
|---|---|---|
| Benzene | 2 | Carcinogen |
| Carbon tetrachloride | 4 | Toxic and environmental hazard |
| 1,2-Dichloroethane | 5 | Toxic |
| 1,1-Dichloroethene | 8 | Toxic |
| 1,1,1-Trichloroethane | 1500 | Environmental hazard |

ppm = parts per million

**Table 2:** Class 2 solvents in pharmaceutical products

| Solvent | PDE (mg/day) | Concentration limit (ppm) |
|---|---|---|
| Acetonitrile | 4.1 | 410 |
| Chlorobenzene | 3.6 | 360 |
| Chloroform | 0.6 | 60 |
| Cyclohexane | 38.8 | 3880 |
| 1,2-Dichloroethene | 18.7 | 1870 |
| Dichloromethane | 6.0 | 600 |
| 1,2-Dimethoxyethane | 1.0 | 100 |
| N,N-Dimethylacetamide | 10.9 | 1090 |
| N,N-Dimethylformamide | 8.8 | 880 |

| | | |
|---|---|---|
| 1,4-Dioxane | 3.8 | 380 |
| 2-Ethoxyethanol | 1.6 | 160 |
| Ethyleneglycol | 6.2 | 620 |
| Formamide | 2.2 | 220 |
| Hexane | 2.9 | 290 |
| Methanol | 30.0 | 3000 |
| 2-Methoxyethanol | 0.5 | 50 |
| Methylbutyl ketone | 0.5 | 50 |
| Methylcyclohexane | 11.8 | 1180 |
| N-Methylpyrrolidone | 5.3 | 530 |
| Nitromethane | 0.5 | 50 |
| Pyridine | 2.0 | 200 |
| Sulfolane | 1.6 | 160 |
| Tetrahydrofuran | 7.2 | 720 |
| Tetralin | 1.0 | 100 |
| Toluene | 8.9 | 890 |
| 1,1,2-Trichloroethene | 0.8 | 80 |
| Xylene* | 21.7 | 2170 |

* usually 60% m-xylene, 14% p-xylene, 9% x-xylene with 17% ethyl benzene

**Table 3:** Class 3 solvents which should be limited by GMP or other quality-based requirements

| | |
|---|---|
| Acetic acid | Heptane |
| Acetone | Isobutyl acetate |
| Anisole | Isopropyl acetate |
| 1-Butanol | Methyl acetate |
| 2-Butanol | 3-Methyl-1-butanol |
| Butyl acetate | Methylethyl ketone |
| tert-Butylmethyl ether | Methylisobutyl ketone |

| | |
|---|---|
| Cumene | 2-Methyl-1-propanol |
| Dimethyl sulfoxide | Pentanol |
| Ethanol | 1-Pentanol |
| Ethyl ether | 1-Propanol |
| Ethyl ether | 2-Propanol |
| Ethyl formate | Propyl acetate |
| Formic acid | |

**Table 4:** Solvents for which no adequate toxicological data was found

| | |
|---|---|
| 1,1-Diethoxypropane | Methylisopropyl ketone |
| 1,1-Dimethoxymethane | Methyltetrahydrofuran |
| 2,2-Dimethoxypropane | Petroleum ether |
| Isooctane | Trichloroacetic acid |
| Isopropyl ether | Trifluoroacetic acid |

# PREPARATION OF STANDARD BUFFER SOLUTION

## Hydrochloric Acid Buffer

Place 50.0 ml of the 0.2 M potassium chloride in a 200-ml volumetic flask, add the specified volume of 0.2 M hydrochloric acid (Table 1) and then add water to volume.

**Table 1**

| pH | 0.2M HCl ml | pH | 0.2M HCl ml |
|---|---|---|---|
| 1.2 | 85.0 | 1.8 | 20.4 |
| 1.3 | 67.2 | 1.9 | 16.2 |
| 1.4 | 53.2 | 2.0 | 13.0 |
| 1.5 | 41.4 | 2.1 | 10.2 |
| 1.6 | 32.4 | 2.2 | 7.8 |
| 1.7 | 26.0 | | |

## Acid Phthalate Buffer

Place 50.0 ml of the 0.2 M potassium hydrogen phthalate in a 200-ml volumetic flask, add the specified volume of 0.2 M hydrochloric acid (Table 2) and then add water to volume.

**Table 2**

| pH | 0.2 M HCl ml | pH | 0.2 M HCl ml |
|---|---|---|---|
| 2.2 | 49.5 | 3.2 | 15.7 |
| 2.4 | 42.2 | 3.4 | 10.4 |
| 2.6 | 35.4 | 3.6 | 6.3 |
| 2.8 | 28.9 | 3.8 | 2.9 |
| 3.0 | 22.3 | 4.0 | 0.1 |

## Neutralised Phthalate Buffer

Place 50.0 ml of 0.2 M potassium hydrogen phthalate in a 200-ml volumetic flask, add the specified volume of 0.2 M sodium hydroxide (Table 3) and then add water to volume.

**Table 3**

| pH | 0.2 M NaOH, ml | pH | 0.2 M NaOH, ml |
|---|---|---|---|
| 4.2 | 3.0 | 5.2 | 28.8 |
| 4.4 | 6.6 | 5.4 | 34.1 |
| 4.6 | 11.1 | 5.6 | 38.8 |
| 4.8 | 16.5 | 5.8 | 42.3 |
| 5.0 | 22.6 | | |

## Phosphate Buffer

Place 50.0 ml of 0.2 M potassium dihydrogen phosphate in a 200-ml volumetic flask, add the specified volume of 0.2 M sodium hydroxide (Table 4) and then add water to volume.

**Table 4**

| pH | 0.2 M NaOH, ml | pH | 0.2 M NaOH, ml |
|---|---|---|---|
| 5.8 | 3.6 | 7.0 | 29.1 |
| 6.0 | 5.6 | 7.2 | 34.7 |
| 6.2 | 8.1 | 7.4 | 39.1 |
| 6.4 | 11.6 | 7.6 | 42.4 |
| 6.6 | 16.4 | 7.8 | 44.5 |
| 6.8 | 22.4 | 8.0 | 46.1 |

## Alkaline Borate Buffer

Place 50.0 ml of 0.2 M boric acid and potassium chloride in a 200-ml volumetic flask, add the specified volume of 0.2 M sodium hydroxide (Table 5) and then add water to volume.

**Table 5**

| pH | 0.2 M NaOH, ml | pH | 0.2 M NaOH, ml |
|---|---|---|---|
| 8.0 | 3.9 | 9.2 | 26.4 |
| 8.2 | 6.0 | 9.4 | 32.1 |
| 8.4 | 8.6 | 9.6 | 36.9 |
| 8.6 | 11.8 | 9.8 | 40.6 |
| 8.8 | 15.8 | 10.0 | 43.7 |
| 9.0 | 20.8 | | |

## STANDARD REAGENTS

1. Boric acid and Potassium chloride, 0.2 M: Dissolve 12.366 g of boric acid and 14.911 g of potassium chloride in water and dilute with water to 1000 ml.
2. Disodium hydrogen phosphate, 0.2 M: Dissolve 71.630 g of disodium hydrogen phosphate in water and dilute with water to 1000 ml.
3. Hydrochloric acid, 0.2 M: Hydrochloric acid diluted with water to contain 7.292 g of HCl in 1000 ml.
4. Potassium chloride, 0.2 M: Dissolve 14.911 g of potassium chloride in water and dilute with water to 1000 ml.
5. Potassium dihydrogen phosphate, 0.2 M: Dissolve 27.218 g of potassium dihydrogen phosphate in water and dilute with water to 1000 ml.
6. Potassium hydrogen phthalate, 0.2 M: Dissolve 40.846 g of potassium hydrogen phthalate in water and dilute with water to 1000 ml.
7. Sodium hydroxide, 0.2 M: Dissolve sodium hydroxide in water to produce 40–60% w/v solution and allow to stand. Siphon off the clear supernatant liquid and dilute with carbon dioxide-free water such that it contains 8.0 g of NaOH in 1000 ml.

## Diluting Concentrated Acids to 1 M Solution

**Hydrochloric acid (36%):** Take 83.5 ml of 36% hydrochloric acid and add distilled water up to 1 litre.

**Nitric acid (70%):** Take 62 ml of 70% nitric acid and add distilled water upto 1 litre.

**Sulfuring acid (98%):** Take 54 ml of 98% concentrated sulfuric acid and add distilled water up to 1 litre.

**Acetic acid (99.5%):** Take 57 ml of the Glacial acetic acid and add distilled water up to 1 litre.

# References


Adamson, AW (1976) Physical chemistry of surfaces, 3rd Ed., Interscience, New York.

Aguiar, AJ, and Zelrher, JE (1969) J Pharm. Sci., 58, 983.

Albert, A and Serjeant, EP (1962) Ionization Constants of Acids and Bases, Methuen, London.

American Society for Testing Materials (ASTM) (1975) Standard Recommended Practice for Dealing with Outlying Observations, E 178–75, 183.

Amidon, GL et al. (1980) J Pharm. Sci., 69, 1363.

Annual Book of ASTM Standards, (1978) American Society for Testing and Materials, Philadelphia, Part 20, (D775) 180, (D880) 229, (D999) 282.

Armstrong, NA et al. (1983) J Pharm. Pharmacol., 35, 320.

Armstrong, NA et al. (1970) J Pharm. Pharmacol. (Suppl), 22, 8S.

Aulton, ME (1981) The Design and Manufacture of Medicine, 3rd Ed., Churchill Livingstone, Elsevier.

Banker, GS, and Rhodes, C (2002) Modern Pharmaceutics, CRC Press.

Becher, P (1965) Emulsions Theory and Practice, 2nd Ed., Reinhold, New York.

Beringer, P et al. (2005) Remington: The Science and Practice of Pharmacy, 21st Ed., Lippincot Williams and Wilkins, Baltimore.

Blythe, RH (1957) Product Formulation and Stability Prediction, Presented at the Production Section of the Canadian Pharm. Mfrg. Assoc.

Bodin, J, and Taub, A (1955) J Am Pharm. Assoc., Sci. Ed., 44, 296.

Bogardus, JB, and Backwood, RK (1979) J Pharm. Sci., 68, 188.

Bolton, S (1983) J Pharm. Sci., 72, 362.

Bond, FC (1952) Min. Eng., 4, 484.

Carstensen, JT (1974) J Pharm. Sci., 63, 1.

Chemburkar, PG, and Joslin, RS (1975) J Pharm. Sci., 64, 414.

Cheymol, J, Buffet, J, and Lechat, P (1947) Ann. Pharm. Franc, 5, 59.

Chong, CW (1963) J Soc. Cosm. Chem., 74, 123.

Collins, AP, Hohmann, JR, and Zopf, LE (1957) Am. Profess. Pharmacist,

23, 231.
Connors, KA, Amidon, GL, and Kennon, L (1979) Chemical Stability of Pharmaceuticals. John Wiley and Sons, New York.
Copsidas, E, and Ward-McQuaid, J (1979) J Int. Med. Res., 7, 592.
Cox, DR (1965) Planning of Experiments. John Wiley and Sons, New York, p. 221.
Cutie, AJ, Ertefaie, S, and Sciarra, JJ (1984) Aerosol Age, 29, 24.
Davies, O (1963) Design and Analysis of Industrial Experiments. Hafner, New York.
De Leede, et.al. (1983) Rate-controlled rectal drug delivery with a hydrogel preparation: I Drug release in vitro. In Rate-Controlled and Site-Specific Rectal Drug Delivery. Edited by De Leede, et.al. 1983, p. 89.
De Martine, ML, and Cussler, EL (1975) J Pharm. Sci., 64, 976.
Dixon, WJ, and Massey, FJ, Jr. (1969) Introduction to Statistical Analysis, 3rd Ed., McGraw-Hill, New York.
Dodge, HL (1962) Ann. Math. Statistics, 14, 264.
Dressman, JB, Fleisher, D, and Amidon, GL (1984) J Pharm. Sci., 73, 1274.
Edwards, JJ (1950) Trans Farad. Soc., 46, 723.
Enesco, J, Branisteanu, D, and Dangeaunu, J (1939) Bull. Acad. Med. Roumanie, 4, 1.
Federal Register (1971) 36, 133.
Federal Register (1978) 43, 45077.
Federal Register (1982) 47, 50441.
Felmeister, A, and Dischler, CA (1964) J Pharm. Sci., 53, 756.
Flavor Research and Food Acceptance, Arthur D Little Inc., Reinhold, New York, 1958.
Florence, AT, and Siepmann, J (2009) Modern Pharmaceutics: applications and advances, 5th Ed., Informa Health Care, New York.
Fonner, DE, Buck, JK, and Banker, GS (1976) J Pharm. Sci., 59, 1587.
Free, SM (1955) Considerations in Sampling for Stability-Presented before the American Drug Manufacturers Association.
Friberg, S (1979) J Soc. Cosmetic Chemists, 30, 309.
Garnett, CJ, et al. (1983) J Chem. Soc, 79, 953.
Garrett, ER, and Carper, RF (1955) J Pharm. Sci., 44, 515.
Garrett, ER (1957) J Am. Pharm. Assoc, Sci. Ed., 46, 584.
Garrett, ER (1962) J Pharm. Sci., 51, 811.
Garrett, ER (1966) J Pharm. Pharmacol., 18, 589.

Griffin, WC (1979) Emulsions in Kirk Othmer Encyclopedia of Chemical Technology, Vol. 8, 3rd Ed., Wiley Interscience, New York, 1979, p. 922 Griffith, AA (1921) Phil. Trans. Roy. Soc. London, Ser. A, 221, 163.
Grizzle, JE (1965) Biometrics, 21, 467.
Groot, RC, and Void, RD (1967) Proc. 4th Int. Congress Surface Activity, 11, 1233.
Gross, HM, and Becker, CH (1953) J Am. Pharm. Assoc, Sci. Ed., 42, 498.
Gstimer, F (1969) Mitt. Deutsch. Pharm. Ges., 39, 21.
Guav, F, and Bisaillon, S (1979) Drug Develop. Ind. Pharm., 5, 107.
Guenther, WC (1981) The American Statistician, 35, 243.
Guidelines for the Analytical Validation of HPLC Method. Pharmacopeial Forum, 9:2789, 1983. Guillory, JK (1972) J Pharm. Sci., 61, 26.
Guttman, DE (1962) J Pharm. Sci., 51, 1162.
Guttman, DE, Hamlin, WE, Shell, JW, and Wag-ner, JG (1961) J Pharm. Sci., 50, 305.
Haleblian, J, and McCrone, W (1969) J Pharm. Sci., 58, 911.
Hamlin, WE, Chulski, X, Johnson, RH, and Wagner, JG (1960) J Am. Pharm. Assoc, Sci. Ed., 49, 253. Harris, CC (1967) Trans. AIME, 238, 17.
Hatch, TJ (1933) Franklin Inst., 215, 27.
Hatch, TJ, and Choate, SP (1929) Franklin Inst., 207, 369.
Heckel, RW (1961) Trans. Metallurgy Soc., Am. Inst. Mech. Eng., 221, 671.
Hendersen, NL, Meer, PM, and Kostenbauder, HB (1961) J. Pharm. Sci., 50, 788.
Hiestand, EN, Well, JE, Pool, CB, and Ocks, JF (1977) J Pharm. Sci., 66, 510.
Hiestand, EN (1978) Proc. Int. Conf. Powder Tech. and Pharmacy, Basel.
Hiestand, EN, and Smith, DP (1984) Powd. Technol., 38, 145.
Higuchi, T, and Lachman, L (1955) J Am. Pharm. Assoc, Sci. Ed., 44, 532.
Higuchi, T, et al. (1954) J Am. Pharm. Assoc., Sci. Ed., 43, 349.
Higuchi, T, Rao, AN, Busse, LW, and Swintosky, JV (1953) J Am. Pharm. Ass., Sci. Ed., 42, 194.
Higuchi, T, and Stella, V (1975) Prodrugs as Novel Drug Delivery Systems. American Chemical Society, Washington, DC.
Hildebrand, JH, and Scott, RL (1950) The Solubility of Non-electrolytes, Reinhold, New York.
Hixson, A, and Crowell, J (1931) Ind. Eng. Chem., 23, 923.

Hollander, M, and Wolfe, DA (1973) Nonparametric Statistical Methods, John Wiley and Sons, New York.
Industry Specifications for Fabricated Aerosol Cans, In CSMA Aerosol Guide, 6th Ed., Chemical Specialties Manufacturers Association, Washington, DC, March 1971, p. 37.
Jaffe, J, and Foss, NE (1959) J Am. Pharm. Ass., Sci. Ed., 48, 26.
Jain GK, et al., (2010) J Pharm. Biomed. Anal., 52, 110.
Jain GK, et al. (2011) The Theory and Practice of Physical Pharmacy, Elsevier Health Sciences, India.
Janovsky, HL (1960) Drug and Cosmetic Industry, 86, 335.
Jones, D (2008) Pharmaceutics – Dosage Form and Design, Pharmaceutical Press.
Jones, TM, and Pilpel, N (1966) J Pharm. Pharmacol., 18, 429.
Kennon, L (1964) J Pharm. Sci., 53, 815.
Kick, F (1885) Das Gesetz der proportionalen Widerstande und seine Anwendung. Arthur Felix, Leipsig, Germany
Kim, C (2004) Advanced Pharmaceutics Physicochemical Principles, CRC Press, Taylor & Francis.
Kim, C (1999) Controlled Release Dosage Form Design, CRC Press, Taylor & Francis.
Kramer, SF, and Flynn, GL (1972) J Pharm. Sci., 61, 1897.
Lachman, L, and Higuchi, T (1957) J Am. Pharm. Assoc, Sci. Ed., 46, 32.
Lachman, L, Guttman, D, and Higuchi, T (1957) J Am. Pharm. Assoc, Sci. Ed., 46, 36.
Lachman, L, Ravin, LJ, and Higuchi, T (1956) J Am. Pharm. Assoc, Sci. Ed., 45, 290.
Lachman, L, Swartz, CJ, and Cooper, J (1960) J Am. Pharm. Assoc, Sci. Ed., 49, 213.
Langenbucher, F, and Lange, B (1969) Pharm. Acta. Helv., 45, 572.
Levy, G, Tsuchiya, T, and Amsel, LP (1972) Clin. Pharmacol. Ther., 13, 258.
Levy, G (1963) Biopharmaceutical considerations in dosage form design and evaluation, In Prescription Pharmacy, Edited by J. B. Sprowls. J. B. Lippincott, Philadelphia, 1963, p. 31.
Lieberman, A (1982) Med. Device and Diagn. Ind., 4(1), 43.
Loo, JCK, and Riegelman, S (1968) J Pharm. Sci., 57, 918.
Marcus, AD, and Taraszka, AJ (1959) J Am. Pharm. Assoc, Sci. Ed., 48, 77.

Marszall, L (1978) Cosmetics Toiletries, 93, 53.
Martin, AN, Swarbrick, J, and Cammarata, A (1969) Physical Pharmacy, 2nd Ed., Lea and Febiger, Philadelphia.
Martin, A, Swarbrick, J, and Cammarata, A (1983) Physical Pharmacy: Physical Chemical Principles in the Pharmaceutical Sciences, 3rd Ed., Lea and Febiger, Philadelphia.
Mc Cormick K (2002) Quality: Pharmaceutical Engineering Series, 1st Ed., Butterworth – Heinemann, Woburn MA.
McBain, JW (1942) Advances in Colloid Science, Vol. I, Interscience, New York.
McCabe, WL, and Smith, JC (1985) Unit Operations and Chemical Engineering, 4th Ed., McGraw-Hill, New York.
Meyer, MC, Slyka, GWA, Dann, RE, and Whyatt, PL (1974) J Pharm. Sci., 63, 1693.
Meyer, RJ, and Cohen, L (1959) J Soc. Cosm. Chem., 10, 143.
Micro Filtration Systems Catalog 1982. Micro Filtration Systems, Dublin, CA, 1981.
MIL STD 105D, Military Sampling Procedures and Tables for Inspection by Attributes. US Government Printing Office, Washington, DC, 1963.
Mitrevej, A, and Augsburger, LL (1980) Drug Dev. Ind. Pharm., 6, 331.
Modern Packaging Encyclopedia and Planning Guide, Vol. 46, McGraw-Hill, New York, 1973. Modern Plastics Encyclopedia, Vol. 49, McGraw-Hill, New York, 1972–1973.
Mollica, JA, Rehm, CR, and Smith, JB (1969) J Pharm. Sci., 58, 635.
Mooney, KG, et al. (1981) J Pharm. Sci., 70, 22.
Mooney, KG, et al. (1981) J. Pharm. Sci., 70, 13.
Nelson, E, Busse, LW, and Higuchi, T (1955) J Am. Pharm. Ass., Sci. Ed., 44, 223.
Nelson, E, Naqvi, SM, Busse, LW, and Higuchi, T (1954) J Am. Pharm. Ass., Sci. Ed., 43, 596. Ober, SS, Viscent, HC, Simon, DE, and Frederick, KJ (1958) Pharm. Assoc. Sci. Ed., 47, 667. Package Engineering Encylcopedia. Cahners, Boston, 1982.
Perrie, T and Rades, T (2009) Pharmaceutics: Drug Delivery and Targeting, Pharmaceutical Press. Perry, RH et al. (1963) Chemical Engineers' Handbook, 4th Ed., McGraw-Hill, New York. Physician's Desk Reference. 36th Ed. Medical Economics, Oradell, New Jersy, 1982, p.

403.
Pisano DJ and Mantus DS (2008) FDA Regulatory Affairs, 2nd Ed., Informa Health Care, New York.
Plossl, GW, and Wight, OW (1967) Production and Inventory Control, Principles and Techniques.
Prentice Hall, Englewood Cliffs, New Jersy.
Porush, I, and Maison, G (1959) US Pat. No. 2,868,691.
Proposed Guidelines for Stability Studies for Human Drugs and Biologicals, Food and Drug Administration, Washington, DC, March 1984.
Quality Assurance in Drug Manufacturing, Eastern Regional Proceedings of PMA-FDA Seminar, Rutgers University, March 1969, Workshop No. E, p 55.
Rankell, AS, and Higuchi, T (1968) J Pharm. Sci., 57, 574.
Riegelman, S, and Crowell, WJ (1958) J Am. Pharm. Assoc, Sci. Ed., 47,115.
Riegelman, S (1960) J Pharm. Sci., 49, 339.
Riegelmen, S, and Fisher, EZ (1962) J. Pharm. Sci., 53, 206.
Rippie, EG, and Danielson, DW (1981) J. Pharm. Sci., 70, 476.
Rudolfo, AS, et al. (1959) Am. J Med. Sci., 237, 585.
Schanker, LS, Shore, PA, and Brodie, BB (1957) J Pharm. Expd. Therap., 120, 528.
Schuhmann, R, Jr. (1940) Mining Technol, 4, Tech. Publ. No. 1189 Schulman, JJ, and Cockbain, EG (1940) Trans. Faraday Soc, 36, 651.
Sciarra, JJ (1970) Aerosols, In Prescription Pharmacy, 2nd Ed., Edited by JB Sprowls. JB Lippincott, Philadelphia, p 280.
Sciarra, JJ (1974) Types of aerosol systems, In The Science and Technology of Aerosol Packaging, Edited by JJ Sciarra and L Stoller. John Wiley and Sons, New York, p 37.
Shapiro, I (1944) PhD Thesis, Univ. Minnesota.
Sherman, P (1968) Rheology of Emulsions, Academic Press, London.
Shichiri, M, Yamasaki, Y, Kawamori, R, et al. (1978) J Pharmacol., 30, 806.
Shin-Etsu Chemical Industry., Ltd., Jpn. Kokai Tokyo Koho Japanese Patent 58/201724 A2 [83/201724] (1983).
Shinoda, K, and Kunieda, H (1983) Phase properties of emulsions: PIT and HLB, In Encyclopedia of Emulsion Technology, Vol. I, Edited by P Becher, Marcel Dekker, New York, p 337.
Siegel, S (1956) Non-Parametric Statistics for the Behavioral Sciences,

McGraw-Hill, New York, p 166.
Sinko, PJ (2011) Martin's Physical Pharmacy and Pharmaceutical Sciences, 5th Ed., Lippincot Williams and Wilkins, Baltimore.
Snedecor, GW, and Cochran, WG (1980) Statistical Methods, 7th Ed., Iowa State Univ. Press, Ames, IA. Spiegel, AJ, and Noseworthy, MM (1963) J Pharm. Sci., 52, 917.
Swarbrick J (2007) Encyclopedia of pharmaceutical technology, 3rd Ed., Informa Health Care, New York.
Swartz, CJ, and Cooper, J (1962) J Pharm. Sci., 51, 89.
Symposium on Pyrogens. J Pharm. Pharmacol., 6:302 1954
Tables of the Binomial Distribution. National Bureau of Standards, US Government Printing Office, Washingon, DC, 1952.
Tabletting Specification Manual, Industrial Pharmaceutical Technology Standard Specifications for Tabletting Tools. 1981 revision. American Pharmaceutical Association, Washington, DC Taraszka, MJ (1971) J Pharm. Sci., 60, 1414.
The United States Pharmacopeia XX/National Formulary XV. Supplement 3. US Pharmacopeial Convention, Rockville, MD, 1982. p 310.
The United States Pharmacopeia XX/National Formulary XV. US Pharmacopeial Convention, Rockville, MD, 1980, pp 958, 990.
Theil, CG, Porush, I, and Law, RD (1961) US Pat. No. 3,014,844.
Thomas, R (1981) Pharm. Eng. 1, 16.
Turrell, R, Marino, AWM, and Nerb, L (1940) Ann. Surg., 112, 417.
US Food and Drug Administration, Federal Register 43:45077, Sept. 29, 1978.
United States Pharmacopoeia XIX. Mack Publishing, Easton, PA. 1975.
Van Campen, L, Amidon, GL, and Zografi, G (1983) J Pharm. Sci., 72, 1381.
Void, RD, and Groot, RC (1963) J Soc. Cosmetic Chemists, 14, 233.
Von Rittinger, RP (1867) Lehrbuch der Aufber-eitungskunde. Ernst and Korn, Berlin, Germany, 19 Wagner, JG, and Nelson, E (1963) J Pharm. Sci., 52, 610.
Walker, EE (1923) Trans. Farad. Soc., 19, 60.
Walters, WA (1959) Physical Aspects of Organic Chemistry, 4th Ed. Van Nostrand, New York, p 332. Wang, LH, and Paruta, AN (1984) Drug Dev. and Ind. Pharm., 10, 667.
Wesley, F (1957) Pharmaceutical Flavor Guide, Fritzsche Brothers Inc., New

York, 1957. Whitworth, CW, and Larocca, JP (1959) J Am. Pharm. Assoc, Sci. Ed., 48, 353.
Wilcoxon, F, and Wilcox, R (1964) Some Rapid Approximate Statistical Procedures, Lederle Labs., Pearl River, New York.
Windheuser, J (1963) Bull. Parenteral Drug Assoc, 17,1.
Wolfson, BB, and Banker, GS (1965) J Pharm. Sci., 54, 195.
Wood, JH, et al. (1965) J Pharm. Sci., 54, 1068.
Wood, JA, et al. (1970) Gail. J Pharm. Sci., 5, 18.
Wray, PE (1969) Drug Cosmet. Ind., 105, 58.
Wurster, DE (Wisconsin Alumni Research Foundation) (1953) US Patent 2,648,609.
Young, JG, Thiel, CG, and Laursen, GA (1962) Nonaqueous Aerosol Dispersions, Midwestern Regional Meeting, Industrial Pharmacy Section, American Pharmaceutical Association, Chicago, IL.