

Sterile Dosage Forms and Novel Drug Delivery Systems (Theory)

BP805T

A Textbook for Final-Year B.Pharm Students

*As per the new Pharmacy Council of India (PCI)
Syllabus 2026 (NEP 2020)*

Dr. Prevesh Kumar

Associate Professor
IFTM University, Moradabad, Uttar
Pradesh, India

Akriti Singh

Assistant Professor
Lingaya's Vidyapeeth, Faridabad, India

Supriya Singh

Assistant Professor
Jaipur National University, Jaipur,
Rajasthan, India



Copyright © 2026 Wiseleaf Scientific Ventures Pvt. Ltd.

All rights reserved.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher, except for brief quotations in reviews or scholarly works permitted by copyright law.

ISBN: 978-81-687391-4-7



June, 2026

Published by: Wiseleaf Scientific Ventures Pvt. Ltd.

Address: Wework - Berger Delhi One Floor 19, C-001/A2, Sector 16b, Noida, Uttar Pradesh, 201301, India

PREFACE

This book covers the subject *Sterile Dosage Forms and Novel Drug Delivery Systems (Theory)*, course code BP805T, for final-year Bachelor of Pharmacy students. It is written for a reader who already understands physical pharmacy, basic pharmaceuticals, pharmacokinetics and biopharmaceuticals, and who is now ready to see how those foundations are used to engineer medicines that do what conventional dosage forms cannot.

The book is built around a single organising idea, stated in its first chapter and returned to in its last: a novel delivery system is the answer to a specific, named limitation of the conventional dosage form, and a sterile product is shaped at every step by the single requirement that it be safe to place directly into the body. Wherever possible the text leads with the reason a system exists before describing how it is built, because a delivery system understood as the solution to a problem is remembered, while one memorised as a list of features is forgotten by the examination hall.

The content follows the five units of the prescribed syllabus. Unit I establishes the rationale for novel delivery and the polymers, lipids, inorganic carriers and excipients from which such systems are made. Unit II covers oral and mucosal delivery, including gastro-retentive and colon-targeted systems and the in-vitro–in-vivo correlations that link the laboratory to the patient. Unit III covers transdermal delivery and the physical technologies that extend it. Unit IV covers the evaluation, quality-by-design, regulatory and case-study material that turns a laboratory system into an approved medicine. Unit V, the largest, covers sterile parenteral and ophthalmic products in depth, from Water for Injection through sterile manufacturing, containers, and the quality-control tests that release a product for use.

Each chapter follows a consistent structure: learning outcomes, an introduction that sets the problem, a main body that develops the mechanism and the formulation reasoning, quick-recall boxes for the points most often examined, a synthesis summary, a set of short-answer, long-answer and multiple-choice questions in the style of university examinations with answer pointers, and a reference list. The aim throughout is that a student who works through a chapter can both understand the topic and answer examination questions on it.

It is hoped that this book serves not only as preparation for an examination but as a durable way of thinking about delivery: name the limitation, match the strategy, respect the evidence, and never forget that a medicine must work as a product and not only as a mechanism.

Authors

Dr. Prevesh Kumar

Akriti Singh

Supriya Singh

COURSE OBJECTIVES

The objectives of this course are to:

1. Understand the scientific and technological foundations of advanced and novel drug delivery systems, including their classification, design rationale, materials, and formulation development challenges.
2. Examine the role of polymers, lipids, and excipients in formulating biodegradable, targeted, and controlled-release drug delivery systems.
3. Study specialized drug delivery routes such as oral, mucosal, transdermal, ocular, parenteral and drug carriers including and their impact on therapeutic efficacy.
4. Apply principles of Quality by Design (QbD), preclinical evaluation, and regulatory science to assess NDDS development and clinical translation.
5. Explore the emerging field of precision medicine, including its genetic, molecular, and technological underpinnings relevant to pharmacy.

COURSE OUTCOMES

Upon successful completion of this course, the students will be able to:

1. Classify and explain the different types of NDDS and describe their need in overcoming limitations of conventional drug delivery.
2. Identify appropriate pharmaceutical ingredients for designing site-specific or controlled-release delivery systems.
3. Design oral, mucosal, parenteral, and transdermal NDDS, and evaluate their mechanisms and formulation parameters.
4. Apply QbD principles and preclinical evaluation tools to assess NDDS quality and safety.
5. Demonstrate an understanding of the scope, principles, and challenges of precision medicine in modern pharmacotherapy.

SYLLABUS

Unit	Topics	No. of lectures
I	Fundamentals, Polymers & Materials for NDDS 1. Limitations of conventional dosage forms (e.g., solubility, permeability, toxicity, first-pass metabolism) 2. Classification of NDDS: Controlled Release, Targeted, Stimuli-Responsive (Smart), Chronotherapeutic, Transdermal and Mucosal, Implantable and Injectable Depot Systems, Vesicular, Polymeric, Ocular, Pulmonary, and Nasal Drug Delivery Systems 3. Biodegradable and biocompatible polymers: PLA, PLGA, PCL, chitosan, gelatin 4. Lipids and surfactants: lecithin, phospholipids, Span/Tween, SLNs 5. Inorganic systems: silica, gold nanoparticles, iron oxide, MOFs 6. Excipients in NDDS: GRAS substances, IIG database, regulatory constraints	12 hours
II	Oral and Mucosal Delivery Systems 1. Gastro-retentive systems: floating, bioadhesive, expandable systems 2. Colon-targeted systems: pH-triggered, microbially-triggered, time-dependent systems 3. Buccal, nasal, and pulmonary carriers: films, sprays, DPIs, liposomes 4. IVIVC and biorelevant dissolution testing.	6 hours
III	Transdermal NDDS 1. Introduction to Transdermal drug delivery: Need, advantages disadvantages, basic structure and components. Formulation aspects including permeation enhancers. 2. Long-acting systems: in-situ gels, microspheres, implants, ocular inserts 3. Transdermal and microneedle technologies: iontophoresis, sonophoresis, micro needles Aseptic manufacturing, scale-up challenges, and process analytical tools (PAT)	6 hours
IV	Evaluation, Quality, and Translation 1. Preformulation and QbD: QTPP, CQAs, CPPs 2. Characterization methods: particle size, zeta potential, SEM/TEM, in-vitro release, permeability assays (PAMPA, Caco-2) 3. Non-clinical evaluation: PK/PD modeling, biodistribution, toxicology	

	4. Regulatory pathways: 505(b)(2), complex generics, ICH Q8–Q10, EMA pathways 5. Case studies: Liposomal Amphotericin B for Fungal Infections, mRNA-Lipid Nanoparticle Vaccines (e.g., COVID-19 Vaccines), Depot Antipsychotic Injections (e.g., Risperidone Microspheres), Transdermal Patch for Hormone Replacement Therapy, Ocular Inserts for Glaucoma Management, Oral Colon-Targeted Delivery for Inflammatory Bowel Disease, Inhalable Insulin for Diabetes Mellitus, Dendrimers for Targeted Cancer Therapy, Chronotherapeutic Drug Delivery in Hypertension. 6. Nanosystems and precision medicine a) Liposomes, niosomes, transferosomes, polymeric nanoparticles, solid-lipid nanoparticles, dendrimers. b) Precision medicine: definition and scope, evolution from "one-size-fits-all" to personalized approaches	
V	Sterile Dosage Forms (Parenteral & Ophthalmic) Introduction & classification: SVP vs LVP; routes; Advantages/limitations. Water for Injection (WFI): Types, preparation, storage/distribution; pharmacopoeial requirements. Formulation components: Vehicles, buffers, co-solvents, surfactants, antioxidants, preservatives, chelators; isotonicity & osmolality. Manufacturing: Aseptic vs terminal sterilisation; filtration, lyophilisation principles; environmental controls. Containers & closures: Glass/Plastic systems; elastomers; selection factors; evaluation & container–closure integrity. Form fill Seal and Blow Fill Seal Technology Ophthalmics: Solutions/suspensions/ointments/gels/inserts; physiological factors (pH, tonicity, viscosity); residence-time enhancers; preservative-free systems (e.g., BFS). Quality control & compliance: Sterility tests, LAL/endotoxins, particulate matter, clarity, extractable volume; labelling & documentation; overview of cGMP/cleanrooms and environmental monitoring.	14 hours

How to Use This Book

The book is designed to be read in order, because the later units build on the earlier ones. The rationale and materials of Unit I are assumed throughout; the manufacturing and quality concepts introduced in Units III and IV are completed in Unit V. A reader preparing a single topic can, however, use any chapter on its own, since cross-references point to where supporting ideas are developed.

A few features are worth using deliberately. The **learning outcomes** at the head of each chapter define what mastery of that chapter means; they are a useful self-test before an examination. The **quick-recall boxes** distil the points most frequently examined and are the fastest revision aid in the book, but they are summaries, not substitutes for the reasoning in the body. The **figures and tables** are numbered per chapter (Figure 3.2 is the second figure of Chapter 3) and are referenced from the text at the point where they help; they are meant to be studied alongside the prose, not skipped. The **self-assessment** questions at the end of each chapter are graded as short-answer (2-mark), long-answer (5- and 10-mark) and multiple-choice items, mirroring university examination patterns, and the answer pointers indicate what a complete answer should contain rather than giving model essays.

Units are best studied as wholes, since each opens with a short overview that frames its chapters and closes a thread that runs through them.

Table of Contents

Preface	iii
How to Use This Book	vii
Unit I — Fundamentals, Polymers and Materials for NDDS	1
Chapter 1 — Limitations of Conventional Dosage Forms and the Rationale for NDDS	2
Learning outcomes	2
1.1 Introduction: the problem a dosage form is meant to solve	2
1.2 The peak–trough problem	2
1.3 The four barriers of the conventional oral route	4
1.3.1 Dissolution and solubility.....	4
1.3.2 Permeability	5
1.3.3 Pre-systemic (first-pass) metabolism	5
1.3.4 Short biological half-life and chemical instability	5
1.4 First-pass metabolism in quantitative terms.....	6
1.5 Limitations of conventional parenterals and topicals.....	7
1.6 The rationale for NDDS, assembled	8
Chapter summary	8
Self-assessment	9
References	10
Chapter 2 — Classification of Novel Drug Delivery Systems.....	11
Learning outcomes	11
2.1 Why a classification is needed	11
2.2 Classification by rate of release	12
2.3 Classification by spatial targeting	14
2.4 Classification by trigger and timing	15
2.5 Classification by route and carrier	15
2.6 A note on overlap and on naming discipline.....	16
Chapter summary	16
Self-assessment	17
References	17
Chapter 3 — Biodegradable and Biocompatible Polymers	18
Learning outcomes	18
3.1 Two words that are not the same: biodegradable and biocompatible	18
3.2 How a biodegradable matrix releases its drug: erosion mechanisms.....	18
3.3 The polyesters: PLA, PGA and PLGA.....	19
3.3.1 Chemistry and degradation.....	19
3.3.2 Tuning PLGA: ratio, molecular weight, and crystallinity.....	20
3.3.3 Applications	21
3.4 Poly(ϵ -caprolactone) (PCL)	22

3.5 Chitosan: a natural polycation.....	22
3.6 Gelatin	23
Chapter summary	24
Self-assessment	24
References	25
Chapter 4 — Lipids and Surfactants	26
Learning outcomes	26
4.1 Why lipids are central to delivery	26
4.2 Amphiphiles and the hydrophobic effect	26
4.3 Lecithin and the phospholipids	27
4.4 Surfactants and the HLB system	28
4.4.1 The Span and Tween series	29
4.5 Solid lipid nanoparticles and nanostructured lipid carriers	29
4.6 Lipid nanoparticles for nucleic acids	31
Chapter summary	31
Self-assessment	32
References	32
Chapter 5 — Inorganic Systems for Drug Delivery.....	33
Learning outcomes	33
5.1 Why use an inorganic carrier at all.....	33
5.2 Mesoporous silica nanoparticles	33
5.3 Gold nanoparticles.....	34
5.4 Superparamagnetic iron oxide nanoparticles	35
5.5 Metal–organic frameworks	36
Chapter summary	38
Self-assessment	38
References	39
Chapter 6 — Excipients in NDDS: GRAS, the Inactive Ingredient Database, and Regulatory Constraints.....	40
Learning outcomes	40
6.1 Excipients are not passengers.....	40
6.2 GRAS: generally recognised as safe	41
6.3 The Inactive Ingredient Database (the syllabus “IIG”).....	42
6.4 The novel excipient problem.....	42
6.5 Putting the constraints to work.....	44
Chapter summary	44
Self-assessment	45
References	45
Unit II — Oral and Mucosal Delivery Systems	46
Chapter 7 — Gastro-retentive Drug Delivery Systems.....	47

Learning outcomes	47
7.1 The problem: the gut does not wait.....	47
7.2 Floating systems.....	48
7.3 Bioadhesive (mucoadhesive) systems.....	49
7.4 Expandable and high-density systems.....	50
7.5 Limitations and the fed–fasted problem.....	50
Chapter summary	51
Self-assessment	51
References	52
Chapter 8 — Colon-Targeted Drug Delivery Systems	53
Learning outcomes	53
8.1 Why target the colon	53
8.2 pH-triggered systems.....	54
8.3 Microbially triggered systems	55
8.4 Time-dependent systems	56
8.5 Combining triggers and marketed approaches.....	56
Chapter summary	56
Self-assessment	57
References	58
Chapter 9 — Buccal, Nasal and Pulmonary Delivery.....	59
Learning outcomes	59
9.1 Why deliver across a mucosa	59
9.2 Buccal and sublingual delivery	60
9.3 Nasal delivery.....	60
9.4 Pulmonary delivery	61
9.5 Vesicular and particulate carriers for mucosal and pulmonary delivery.....	62
Chapter summary	62
Self-assessment	63
References	63
Chapter 10 — IVIVC and Biorelevant Dissolution Testing	65
Learning outcomes	65
10.1 The problem an IVIVC solves	65
10.2 The levels of IVIVC	65
10.3 Biowaivers.....	67
10.4 Why dissolution media matter: from compendial to biorelevant.....	67
10.5 The biorelevant fluids.....	68
Chapter summary	69
Self-assessment	69
References	70

Unit III — Transdermal NDDS.....	71
Chapter 11 — Transdermal Drug Delivery: Need, Structure, Components and Permeation Enhancers	72
Learning outcomes	72
11.1 Why deliver through the skin.....	72
11.2 The structure of the skin and the barrier	72
11.3 The three routes of permeation.....	73
11.4 What makes a drug a transdermal candidate.....	74
11.5 Components and architecture of a transdermal patch	74
11.6 Chemical permeation enhancers.....	75
Chapter summary	76
Self-assessment	76
References	77
Chapter 12 — Long-Acting Systems: In-Situ Gels, Microspheres, Implants and Ocular Inserts.....	78
Learning outcomes	78
12.1 The depot philosophy	78
12.2 In-situ gelling systems.....	79
12.3 Injectable microspheres.....	81
12.4 Implants.....	81
12.5 Ocular inserts and intravitreal implants	81
Chapter summary	82
Self-assessment	83
References	83
Chapter 13 — Microneedles, Physical Enhancement, Aseptic Manufacturing and PAT.....	84
Learning outcomes	84
13.1 Beyond the passive limit	84
13.2 Microneedles	84
13.3 Iontophoresis and sonophoresis	85
13.4 Aseptic manufacturing for skin-piercing and injectable systems.....	87
13.5 Scale-up challenges	87
13.6 Process analytical technology (PAT)	87
Chapter summary	88
Self-assessment	88
References	89
Unit IV — Evaluation, Quality and Translation	90
Chapter 14 — Preformulation and Quality by Design.....	91
Learning outcomes	91
14.1 Preformulation for novel systems	91
14.2 From quality-by-testing to Quality by Design	91
14.3 The QbD elements in sequence	92

14.4 Design space, control strategy and the edge of failure.....	93
14.5 QbD and the ICH framework	94
Chapter summary	94
Self-assessment	95
References	95
Chapter 15 — Characterization Methods for Novel Delivery Systems.....	97
Learning outcomes	97
15.1 Why characterization matters.....	97
15.2 Particle size and size distribution by dynamic light scattering	97
15.3 Zeta potential and colloidal stability	98
15.4 Morphology by electron microscopy	99
15.5 Drug loading, encapsulation efficiency and in-vitro release.....	99
15.6 Permeability assays: PAMPA and Caco-2	99
Chapter summary	101
Self-assessment	101
References	102
Chapter 16 — Non-Clinical Evaluation: PK/PD, Biodistribution and Toxicology	103
Learning outcomes	103
16.1 What non-clinical evaluation is for	103
16.2 How a delivery system reshapes pharmacokinetics	104
16.3 Biodistribution and the fate of the carrier	105
16.4 Toxicology	106
16.5 Why the whole system, and not the drug alone, is evaluated.....	106
Chapter summary	107
Self-assessment	107
References	108
Chapter 17 — Regulatory Pathways for Novel Delivery Systems	109
Learning outcomes	109
17.1 Three doors to the US market	109
17.2 Why 505(b)(2) fits novel delivery.....	110
17.3 Complex generics	110
17.4 The ICH quality framework	112
17.5 The EMA and the Indian context	112
Chapter summary	113
Self-assessment	113
References	114
Chapter 18 — Case Studies in Novel Drug Delivery.....	115
Learning outcomes	115
18.1 How to read a case study.....	115

18.2 Liposomal amphotericin B (AmBisome)	116
18.3 mRNA–lipid nanoparticle vaccines (Comirnaty, Spikevax)	116
18.4 Depot antipsychotic injection (Risperdal Consta)	117
18.5 Transdermal hormone replacement (estradiol patch)	117
18.6 Ocular insert for glaucoma (Ocusert)	117
18.7 Colon-targeted delivery for inflammatory bowel disease	117
18.8 Inhalable insulin: a success and a failure (Afrezza and Exubera)	117
18.9 Dendrimers for targeted cancer therapy	118
18.10 Chronotherapeutic delivery in hypertension	118
18.11 The common thread	118
Chapter summary	119
Self-assessment	119
References	120
Chapter 19 — Nanosystems and Precision Medicine	121
Learning outcomes	121
19.1 The nanocarrier families gathered	121
19.2 Liposomes recapped, and the niosome	122
19.3 The transferosome	122
19.4 Dendrimers	122
19.5 Precision medicine and where delivery fits	123
Chapter summary	125
Self-assessment	125
References	126
Unit V — Sterile Dosage Forms: Parenteral and Ophthalmic	127
Chapter 20 — Parenterals: Introduction and Classification	128
Learning outcomes	128
20.1 What makes a product parenteral	128
20.2 Small-volume and large-volume parenterals	128
20.3 Routes of parenteral administration	129
20.4 Advantages and limitations of the parenteral route	131
Chapter summary	131
Self-assessment	132
References	132
Chapter 21 — Water for Injection	133
Learning outcomes	133
21.1 Water as the principal ingredient	133
21.2 The hierarchy of pharmaceutical waters	133
21.3 Pharmacopoeial requirements for WFI	134
21.4 Production of WFI	135

21.5 Storage and distribution	135
Chapter summary	137
Self-assessment	137
References	138
Chapter 22 — Formulation Components of Parenterals	139
Learning outcomes	139
22.1 Building a safe injection.....	139
22.2 Vehicles	140
22.3 Buffers	140
22.4 Co-solvents and surfactants.....	141
22.5 Antioxidants and chelating agents.....	141
22.6 Antimicrobial preservatives	141
22.7 Isotonicity and osmolality	142
Chapter summary	143
Self-assessment	143
References	144
Chapter 23 — Sterile Manufacturing: Sterilisation, Filtration, Lyophilisation and Environmental Control.....	145
Learning outcomes	145
23.1 Two philosophies of achieving sterility	145
23.2 Methods of sterilisation and the measures of success	146
23.3 Sterile filtration	146
23.4 Lyophilisation (freeze-drying)	147
23.5 Environmental control and cleanrooms.....	148
Chapter summary	148
Self-assessment	149
References	149
Chapter 24 — Containers, Closures, and Form-Fill-Seal/Blow-Fill-Seal Technology	151
Learning outcomes	151
24.1 The container is part of the product	151
24.2 Glass containers and the pharmacopoeial glass types.....	152
24.3 Elastomeric closures.....	152
24.4 Extractables, leachables, sorption and integrity	153
24.5 Form-Fill-Seal and Blow-Fill-Seal technology.....	153
Chapter summary	155
Self-assessment	155
References	156
Chapter 25 — Ophthalmic Products	157
Learning outcomes	157
25.1 The eye as a delivery target.....	157

25.2 Essential requirements of an ophthalmic product	158
25.3 Dosage forms and residence-time enhancers	159
25.4 Preservatives and the case for preservative-free systems.....	159
Chapter summary	160
Self-assessment	160
References	161
Chapter 26 — Quality Control and Compliance	162
Learning outcomes	162
26.1 Proving a sterile product is fit to use.....	162
26.2 Sterility versus pyrogens: two distinct properties	163
26.3 The sterility test.....	163
26.4 The bacterial endotoxin test (LAL).....	163
26.5 Particulate matter, clarity and extractable volume	164
26.6 Labelling and documentation	164
26.7 cGMP, cleanrooms and environmental monitoring	164
Chapter summary	166
Self-assessment	166
References	167
Glossary of Key Terms	168
Subject Index.....	171

Unit I - Fundamentals, Polymers and Materials for NDDS

Unit overview. A novel drug delivery system is not a clever container bolted onto an old drug. It is a designed answer to a specific failure of the conventional dosage form: a drug that will not dissolve, will not permeate, will not stay in the body long enough, or harms tissues it was never meant to reach. This unit builds the foundation the rest of the book stands on. Chapter 1 sets out exactly where conventional tablets, capsules and simple injections fall short, and why those shortfalls justify the engineering effort that follows. Chapter 2 organises the field into a usable classification. Chapters 3 to 6 then survey the raw materials of the discipline: the biodegradable polymers, the lipids and surfactants, the inorganic carriers, and the excipients and regulatory categories that decide what a formulator is actually allowed to put into a product. Master this unit and the later units on transdermal systems, evaluation, and sterile products become applications of a logic already understood, rather than a list of facts to memorise.

Chapter 1 - Limitations of Conventional Dosage Forms and the Rationale for NDDS

Learning outcomes

After working through this chapter the student will be able to:

1. Explain, using plasma concentration–time profiles, why conventional immediate-release dosing produces peak–trough fluctuation and how that fluctuation links to loss of efficacy and onset of toxicity.
2. Identify the four principal barriers (solubility, permeability, pre-systemic loss, and short biological half-life) that limit the performance of conventional oral and parenteral products, and relate each to the Biopharmaceutics Classification System.
3. Describe first-pass metabolism quantitatively in terms of hepatic extraction and bioavailability, and list the administration routes that avoid it.
4. Justify the rationale for a novel drug delivery system for a given drug, by matching a specific molecular or therapeutic limitation to the delivery strategy that resolves it.
5. State, with a real example, the therapeutic and commercial drivers that make reformulation through NDDS worthwhile.

1.1 Introduction: the problem a dosage form is meant to solve

A drug molecule by itself does nothing useful. To produce an effect it has to arrive at its site of action, at a concentration high enough to act but low enough to be safe, and stay there for as long as the therapy requires. The dosage form is the vehicle that makes this happen. For most of pharmaceutical history that vehicle was a tablet, a capsule, or a solution for injection, designed to release its entire dose as quickly as the body would accept it. These immediate-release products are cheap, stable, and easy to manufacture, and for a great many drugs they remain entirely adequate.

The trouble begins when the drug itself is awkward. Some drugs barely dissolve in the gut fluid. Some dissolve well but cannot cross the intestinal wall. Some are absorbed efficiently only to be destroyed by the liver before they reach the bloodstream. Others are cleared so fast that a patient would need to swallow a tablet every two hours to stay within the therapeutic range. And some are toxic to healthy tissue, so that the dose needed to treat a tumour also damages the heart, the kidney, or the marrow. In every one of these situations the conventional dosage form is the bottleneck, not the drug. The molecule may be pharmacologically excellent; the delivery is what fails.

Novel drug delivery systems (NDDS) exist to remove that bottleneck. The term covers any system engineered to control the rate, the time, or the place at which a drug is released, rather than simply dumping the dose and letting physiology take its course. Before studying how these systems are built, it is worth being precise about what exactly they are built to fix. The rest of this chapter works through the limitations of conventional dosage forms one barrier at a time, and shows how each limitation defines a corresponding design goal.

1.2 The peak–trough problem

Consider a drug with a short half-life given as an immediate-release tablet three or four times a day. Each dose is absorbed quickly, the plasma concentration climbs to a peak, and then

falls as the drug is distributed and eliminated. By the time the next dose is due the concentration has often dropped below the level needed for effect. The resulting plasma profile is a series of spikes and valleys (Figure 1.1).

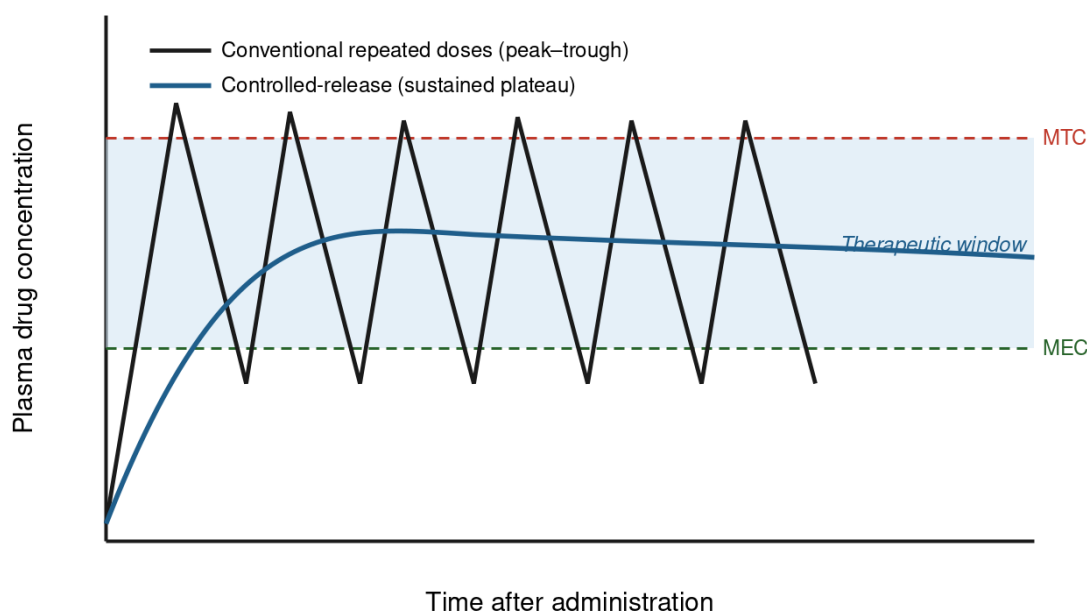


Figure 1.1. Plasma drug concentration against time for repeated immediate-release doses (black sawtooth) compared with a single controlled-release system (blue). The shaded band is the therapeutic window, bounded above by the minimum toxic concentration (MTC) and below by the minimum effective concentration (MEC). The conventional regimen repeatedly overshoots toward toxicity and then falls into sub-therapeutic troughs; the controlled-release system holds the concentration inside the window.

Two boundaries define safe and useful therapy. The **minimum effective concentration (MEC)** is the plasma level below which the drug produces no measurable benefit. The **minimum toxic concentration (MTC)** is the level above which adverse effects appear. The span between them is the **therapeutic window**. A drug with a wide window is forgiving; one with a narrow window, such as digoxin, theophylline, lithium, or many cytotoxics, leaves little room between doing nothing and doing harm.

The peaks of a conventional regimen push the patient toward the MTC, which is where dose-related side effects appear, while the troughs drop below the MEC, which is where the disease is left untreated. A patient on a thrice-daily analgesic may swing between drowsiness an hour after each dose and breakthrough pain before the next. The clinical goal is obvious: flatten the profile so that the concentration sits steadily inside the window. That single objective, smoothing the plasma curve, is the founding rationale for the largest class of NDDS, the controlled- and sustained-release systems. They trade the rapid, complete release of the conventional tablet for a slow, governed release that keeps the concentration on the plateau shown in blue.

The benefit is not only pharmacological. A system that releases over twelve or twenty-four hours can replace three or four daily doses with one, and dosing frequency is one of the strongest predictors of whether a patient actually takes a medicine as prescribed. Reducing a

regimen from four tablets a day to one is, for many chronic diseases, the difference between a therapy that works in the clinic and one that works in real life.

1.3 The four barriers of the conventional oral route

The oral route carries the overwhelming majority of doses because it is convenient, painless, and cheap. It is also the route on which a drug faces the most obstacles. A swallowed dose must survive and clear four barriers in sequence before it can act (Figure 1.2). A weakness at any one of them limits the whole.

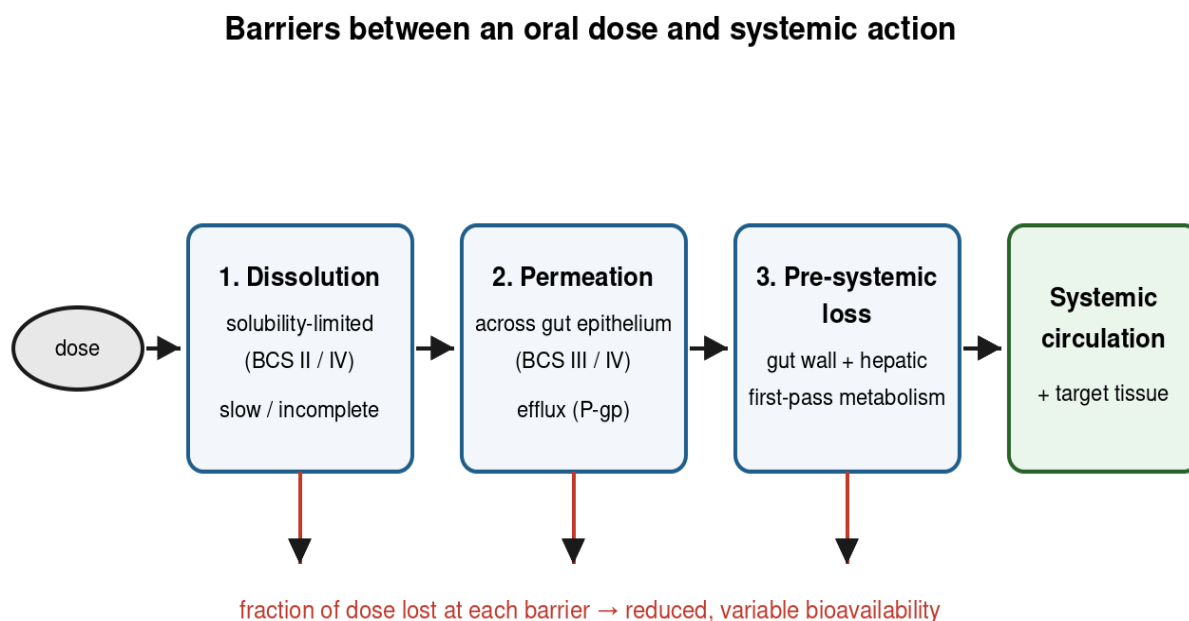


Figure 1.2. The sequence of barriers between an oral dose and systemic action. The drug must dissolve, permeate the gut epithelium, and survive pre-systemic (first-pass) metabolism before reaching the circulation. A fraction of the dose is lost at each step, so the bioavailable fraction is the product of the efficiencies of all three.

1.3.1 Dissolution and solubility

A drug can only be absorbed in solution. A tablet must first disintegrate and then dissolve in the limited volume of gastrointestinal fluid before any absorption can begin. For a poorly water-soluble drug, dissolution is so slow that the dose may pass through the absorption window of the small intestine before it has fully dissolved, and the undissolved fraction is simply excreted. Solubility limitation is now the single most common formulation challenge in development, because a large share of modern drug candidates emerging from high-throughput screening are lipophilic and poorly soluble.

The **Biopharmaceutics Classification System (BCS)** organises this thinking. It sorts drugs by aqueous solubility and intestinal permeability into four classes: Class I (high solubility, high permeability), Class II (low solubility, high permeability), Class III (high solubility, low permeability), and Class IV (low solubility, low permeability).¹ Class I drugs rarely need clever delivery; absorption is limited only by how fast the formulation releases the dose. The problems live in the other three classes. For a Class II drug, absorption is dissolution-rate-limited, and the formulator's task is to raise the effective solubility or the dissolution rate,

through salt formation, particle-size reduction, solid dispersions, lipid-based systems, or nanocrystals.

1.3.2 Permeability

Dissolving is not the same as getting in. Once in solution, a drug must cross the intestinal epithelium, a continuous lipophilic membrane sealed at the cell junctions. Highly water-soluble but poorly lipophilic molecules, including most peptides and many polar small molecules, cross this barrier slowly. These are the BCS Class III drugs, where absorption is permeability-limited. Raising solubility achieves nothing here; the design goal is instead to improve transport, by using permeation enhancers, by exploiting carrier-mediated uptake, or by changing the route entirely. Active efflux compounds the problem: transporters such as P-glycoprotein in the apical membrane pump certain drugs back into the lumen as fast as they enter, capping the absorbed fraction regardless of concentration.

1.3.3 Pre-systemic (first-pass) metabolism

A drug that has dissolved and permeated still has not reached the systemic circulation. Blood draining the gut does not go straight to the heart; it is collected by the hepatic portal vein and delivered first to the liver. The gut wall and the liver together are rich in metabolising enzymes, principally the cytochrome P450 family. A drug that is a good substrate for these enzymes can be largely metabolised on this first pass, before it ever reaches the rest of the body. This is **first-pass (pre-systemic) metabolism**, treated in detail in Section 1.4. It explains why some drugs with excellent intrinsic activity have oral bioavailabilities of only a few percent.

1.3.4 Short biological half-life and chemical instability

The final limitation is not about getting in but about staying useful. A drug with a short elimination half-life disappears quickly, forcing frequent dosing and producing the peak–trough swings of Section 1.2. Separately, some drugs are chemically fragile: proteins and peptides are denatured by stomach acid and digested by gut proteases, which is why insulin cannot simply be swallowed. For these molecules the conventional oral tablet is not merely inefficient; it is impossible, and the entire delivery problem must be solved by another route or by a system that physically protects the drug until it is past the threat.

Table 1.1 collects these limitations and pairs each with the delivery strategy NDDS uses to address it. The rest of this book is, in large part, an expansion of this table.

Table 1.1. *Principal limitations of conventional dosage forms and the NDDS strategies that address them.*

Limitation of conventional form	Consequence	NDDS strategy (later chapters)
Poor aqueous solubility (BCS II/IV)	Slow, incomplete absorption; high variability	Nanocrystals, solid dispersions, lipid carriers, SLN/NLC, micelles (Ch 4)
Poor permeability (BCS III/IV)	Low absorbed fraction; efflux losses	Permeation enhancers, vesicular carriers, mucosal and transdermal routes (Ch 9, 11)
Extensive first-pass metabolism	Very low, variable oral bioavailability	Transdermal, buccal, nasal, pulmonary, parenteral routes (Ch 9,

		11, Unit V)
Short half-life	Frequent dosing, peak–trough swings, poor adherence	Controlled/sustained release; depot injections, implants (Ch 2, 12)
Chemical/enzymatic instability	Drug destroyed before absorption	Enteric and colon targeting; protective carriers; non-oral routes (Ch 8, 9)
Non-specific distribution / toxicity	Healthy tissue harmed; narrow therapeutic index	Targeted and stimuli-responsive systems; nanoparticles (Ch 2, 19)
Fixed release timing	Mismatch with circadian disease rhythm	Chronotherapeutic / pulsatile systems (Ch 2, Ch 18)

1.4 First-pass metabolism in quantitative terms

First-pass metabolism deserves a closer look because it is both a major limitation and one of the clearest justifications for choosing a non-oral route. When a drug is absorbed from the gut, the fraction that escapes metabolism in the gut wall and liver before reaching the systemic circulation determines its oral bioavailability.

The liver's efficiency at removing a drug is captured by the **hepatic extraction ratio (E)**, the fraction of drug removed in a single pass through the organ. A drug with E close to 1 is almost completely extracted on first pass; one with E close to 0 passes through largely untouched. The fraction surviving the liver is therefore $(1 - E)$. For an orally absorbed drug that is completely absorbed and not metabolised in the gut wall, the maximum achievable oral bioavailability is approximately $(1 - E)$. A high-extraction drug such as propranolol, lidocaine, or glyceryl trinitrate has so high an E that oral bioavailability is low and erratic; glyceryl trinitrate is so completely cleared on first pass that it is essentially useless as a conventional swallowed tablet and is given sublingually or transdermally instead.

The escape routes are shown in Figure 1.3. Any route that delivers drug into the systemic veins without first passing through the portal circulation avoids hepatic first-pass loss. The sublingual and buccal mucosae drain into veins that join the systemic circulation directly. The skin does the same, which is the central pharmacokinetic argument for transdermal delivery. Nasal and pulmonary absorption likewise reach the blood without a portal detour, and parenteral injection bypasses the gut entirely. Rectal administration partly avoids the liver, because the lower rectal veins drain to the systemic circulation while the upper rectal vein drains to the portal vein, giving an intermediate and somewhat unpredictable result.

First-pass metabolism and the routes that bypass it

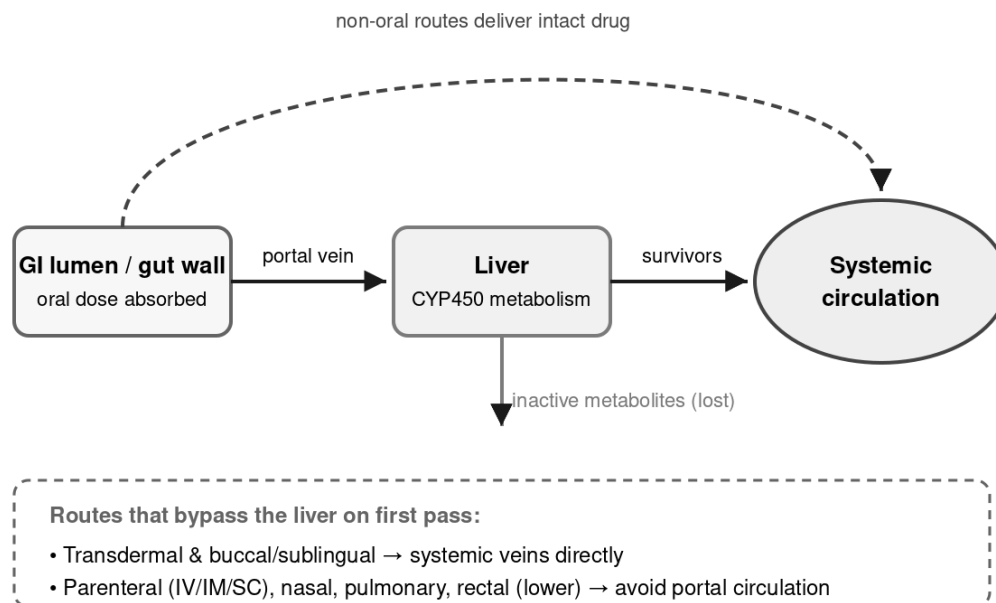


Figure 1.3. First-pass metabolism and the routes that bypass it. An oral dose is carried by the portal vein to the liver, where cytochrome P450 enzymes may destroy a large fraction before it reaches the systemic circulation. Transdermal, buccal/sublingual, nasal, pulmonary and parenteral routes deliver drug to the systemic circulation without this first hepatic pass.

The practical lesson is direct. If a drug fails orally because of first-pass metabolism rather than because of poor solubility or permeability, the rational response is not to reformulate the tablet but to change the route. This single insight underlies transdermal patches for hormones and for glyceryl trinitrate, buccal films, and a large part of the parenteral and pulmonary fields.

1.5 Limitations of conventional parenterals and topicals

It would be a mistake to think the limitations belong only to tablets. Conventional injections solve the absorption problem but introduce their own. A simple intravenous bolus produces the highest possible peak concentration immediately, which for a narrow-window drug means a spike of toxicity, followed by rapid decline that demands repeated injection. Intramuscular and subcutaneous solutions are absorbed quickly and cleared quickly. None of these forms controls where the drug goes once it is in the blood; the dose distributes throughout the body, so that treating a localised tumour exposes the whole organism. The toxicity of conventional cytotoxic chemotherapy is largely a delivery failure of this kind: the drug works, but it cannot tell tumour from healthy tissue. Depot injections, implants, and targeted nanoparticles are the NDDS responses, providing sustained release from a single injection or directing the drug preferentially to the diseased site.

Conventional topical and ophthalmic products face the opposite difficulty: keeping the drug in place. An eye drop is washed away by tears and blinking within minutes, so that only a small percentage of the instilled dose is absorbed and the patient must dose many times a day. Ointments and simple creams on the skin deliver erratically. Residence-time enhancers, in-situ gels, and ocular inserts, covered in later units, all exist to overcome this rapid loss.

1.6 The rationale for NDDS, assembled

Pulling the threads together, a novel delivery system is justified whenever a specific, identifiable limitation of the conventional form prevents a good drug from being used well. The justification is always concrete, never general. One does not reformulate for novelty; one reformulates because this drug has this problem that this system solves.

The drivers fall into three groups. The **therapeutic** driver is improved or safer treatment: a flatter plasma profile, a longer dosing interval, drug delivered to the target and kept away from healthy tissue, or a route that makes an otherwise impossible therapy feasible. The **patient** driver is acceptability and adherence: fewer doses, a less painful or invasive route, and fewer side effects, all of which make a patient more likely to complete therapy. The **commercial and regulatory** driver is real and should not be ignored in a professional course: reformulating an established drug into a superior delivery system can extend a product's useful life, differentiate it from competitors, and, in some jurisdictions, reach the market by a streamlined regulatory route built precisely for modified versions of known drugs. In the United States this is the 505(b)(2) pathway, examined in Unit IV. Each of these drivers is legitimate; the strongest products satisfy all three at once.

A single example ties the logic together. Conventional amphotericin B is a highly effective antifungal but is severely toxic to the kidney, because the free drug interacts with cholesterol in human cell membranes as well as with ergosterol in fungal membranes. The drug is excellent; the delivery is dangerous. Packaging amphotericin B inside a liposome, a lipid vesicle, changes how the drug distributes in the body and reduces its exposure to the kidney, which markedly lowers nephrotoxicity while preserving antifungal activity. The liposomal product AmBisome is the result, and it is widely used in India and internationally for serious systemic fungal infections.² Nothing about the molecule changed. The delivery system did the work, and it did so by answering a precisely identified limitation, namely non-specific distribution and dose-limiting toxicity, with a precisely matched strategy. That is the entire philosophy of this subject in one product.

Quick-recall box — the NDDS justification test Before proposing a novel system, name the limitation it answers. If the answer is “solubility,” think nanocrystals or lipids. If “permeability,” think enhancers or a mucosal route. If “first-pass loss,” change the route. If “short half-life,” think controlled release or a depot. If “non-specific toxicity,” think targeting. A novel system without a named limitation is decoration, not design.

Chapter summary

Conventional immediate-release dosage forms release their whole dose as fast as the body allows. For many drugs that is fine, but for a large and growing number it is the limiting step in therapy. Repeated immediate-release dosing produces a sawtooth plasma profile that swings between toxic peaks and sub-therapeutic troughs, a particular danger for narrow-window drugs and a constant threat to patient adherence. The oral route, though convenient, imposes four sequential barriers: a drug must dissolve, permeate the gut wall, and survive first-pass metabolism in the gut and liver before it reaches the circulation, and even then a short half-life or chemical instability can defeat it. The Biopharmaceutics Classification System frames the first two barriers in terms of solubility and permeability and points to the appropriate remedy for each. First-pass metabolism, quantified through the hepatic extraction ratio, is best overcome not by reformulating the tablet but by choosing a route that bypasses the portal circulation, which is the central argument for transdermal, buccal, nasal, pulmonary

and parenteral delivery. Conventional parenterals and topicals have their own failings, chiefly uncontrolled distribution and rapid loss from the site. Every novel drug delivery system is the answer to one of these specific limitations, and the rationale for building one is always therapeutic, patient-centred, and commercial together, as the liposomal amphotericin B example shows.

Self-assessment

Short-answer (2 marks each)

1. Define the minimum effective concentration and the minimum toxic concentration, and state what is meant by the therapeutic window.
2. Name the four BCS classes and state the rate-limiting step to absorption for Class II and Class III drugs.
3. Define the hepatic extraction ratio and write the expression for the maximum oral bioavailability of a completely absorbed drug in terms of it.
4. List four routes of administration that avoid hepatic first-pass metabolism.
5. Why must a drug be in solution before it can be absorbed from the gut?

Long-answer (5–10 marks each)

6. With a labelled plasma concentration–time diagram, explain why a controlled-release system is preferred over repeated immediate-release dosing for a drug with a short half-life and a narrow therapeutic window.
7. Describe the four sequential barriers faced by an orally administered drug and, for each, identify a category of novel delivery strategy that addresses it.
8. “A novel delivery system is justified only when it answers a specific limitation of the conventional form.” Discuss this statement using liposomal amphotericin B and one other example.

Multiple-choice

9. A drug with a hepatic extraction ratio of 0.9 is best administered by which route to maximise systemic availability? (a) Oral tablet (b) Transdermal patch (c) Enteric-coated tablet (d) Oral solution
10. A BCS Class II drug is characterised by: (a) high solubility, high permeability (b) low solubility, high permeability (c) high solubility, low permeability (d) low solubility, low permeability
11. Peak–trough fluctuation in plasma concentration is most dangerous for a drug with: (a) a wide therapeutic window (b) a long half-life (c) a narrow therapeutic window (d) high solubility

Answer pointers. 1. MEC = lowest plasma level giving effect; MTC = level above which toxicity appears; window = range between them. 2. I, II, III, IV; Class II is dissolution-rate-limited, Class III is permeability-limited. 3. E = fraction removed per hepatic pass; max oral $F \approx (1 - E)$. 4. Transdermal, buccal/sublingual, nasal, pulmonary, parenteral (any four). 5. Only dissolved molecules can diffuse across the epithelial membrane. 6. Expect Figure 1.1 reproduced with MEC/MTC, explanation of plateau keeping concentration in-window and reducing dosing frequency/adherence benefit. 7. Dissolution→solubility strategies; permeation→enhancers/mucosal; first-pass→route change; half-life/stability→controlled release/protection. 8. Should name limitation→strategy mapping for two products. 9. (b). 10. (b). 11. (c).

References

1. Amidon GL, Lennernäs H, Shah VP, Crison JR. A theoretical basis for a biopharmaceutic drug classification: the correlation of in vitro drug product dissolution and in vivo bioavailability. *Pharm Res.* 1995;12(3):413–20.
2. Stone NRH, Bicanic T, Salim R, Hope W. Liposomal amphotericin B (AmBisome): a review of the pharmacokinetics, pharmacodynamics, clinical experience and future directions. *Drugs.* 2016;76(4):485–500.
3. Aulton ME, Taylor KMG, editors. *Aulton's Pharmaceutics: The Design and Manufacture of Medicines.* 6th ed. Edinburgh: Elsevier; 2022.
4. Allen LV, Ansel HC. *Ansel's Pharmaceutical Dosage Forms and Drug Delivery Systems.* 11th ed. Philadelphia: Wolters Kluwer; 2018.
5. Shargel L, Yu ABC. *Applied Biopharmaceutics and Pharmacokinetics.* 7th ed. New York: McGraw-Hill; 2016.

Chapter 2 - Classification of Novel Drug Delivery Systems

Learning outcomes

After working through this chapter the student will be able to:

1. Classify novel delivery systems along four independent axes: rate control, spatial targeting, trigger or timing, and route or carrier, and explain why a single product belongs to several categories at once.
2. Distinguish sustained, controlled, extended, and delayed release using their plasma and release profiles, and identify the kinetic order each aims for.
3. Explain passive and active targeting, including the enhanced permeability and retention (EPR) effect and ligand-mediated uptake.
4. Describe the principle of stimuli-responsive and chronotherapeutic systems and give a verified example of each.
5. Match a delivery requirement to the appropriate class of system and name a marketed product that exemplifies it.

2.1 Why a classification is needed

The field of novel delivery is broad enough that without a map it dissolves into an unmemorable list of product names. A good classification does two things: it groups systems by the design problem they solve, and it makes clear that the categories are not mutually exclusive. A pegylated liposomal doxorubicin product, for instance, is at once a vesicular carrier, a rate-controlling system, and a passively targeted system. Trying to file it under one heading misses the point. The sensible approach is to recognise several independent axes of classification and to place a product on each axis in turn (Figure 2.1).

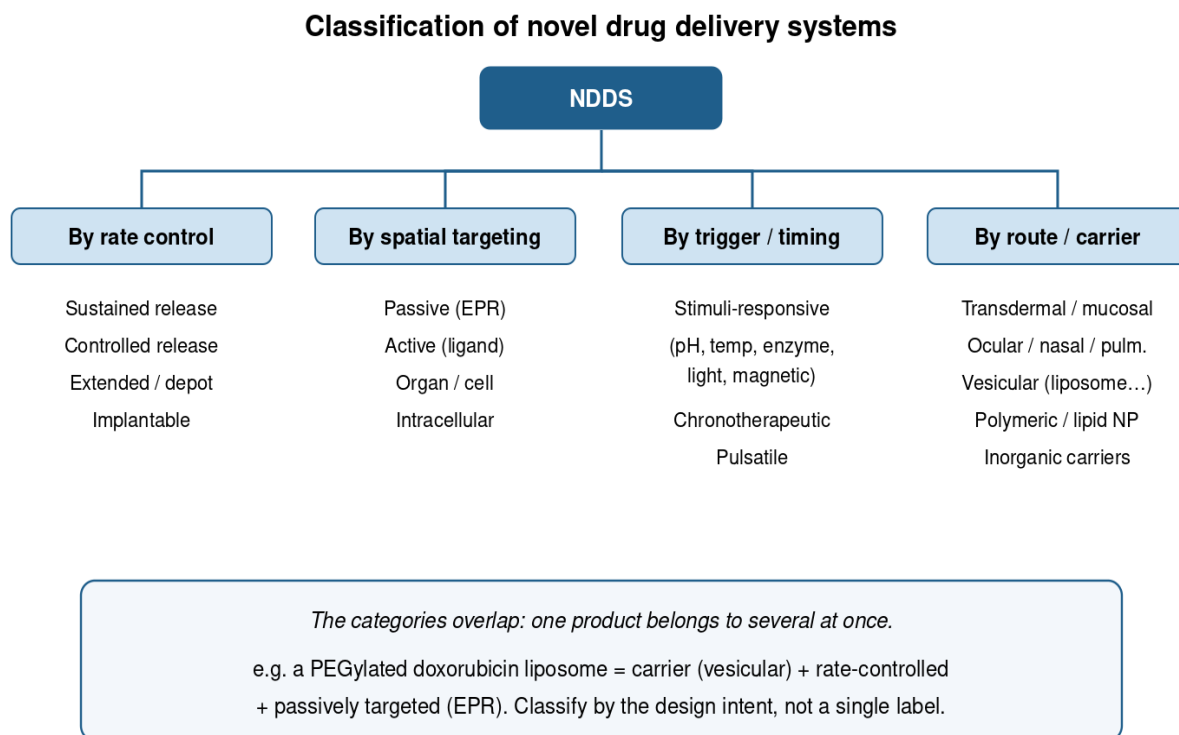


Figure 2.1. The four axes along which novel delivery systems are classified: by rate control, by spatial targeting, by trigger or timing, and by route or carrier. A real product is described by its position on every axis, not by a single label.

The four axes are: control of the **rate** of release; control of the **place** the drug reaches (targeting); control of the **time** or **trigger** of release; and the **route** and **carrier** used to achieve all of these. The sections below take each axis in turn.

2.2 Classification by rate of release

This is the oldest and largest family, descended directly from the peak–trough argument of Chapter 1. The shared aim is to replace the rapid, complete release of an immediate-release form with a slower, governed release. The terminology is often used loosely in marketing, but in a pharmaceuticals course the distinctions matter.

Immediate release is the conventional baseline: the dose is released as fast as the formulation allows, with no attempt to control rate.

Sustained release prolongs release so that the drug appears over an extended period, reducing dosing frequency. It does not guarantee a constant rate; it simply slows and lengthens the release.

Controlled release is the more demanding goal: release at a predetermined, ideally constant rate, so that the input rate matches the elimination rate and the plasma concentration holds steady. The kinetic ideal here is **zero-order release**, where the same amount of drug is released per unit time regardless of how much remains (Figure 2.2). A reservoir system with a rate-controlling membrane, or an osmotic pump, can approach this. Most matrix systems, by contrast, release by diffusion and follow roughly **first-order** or **square-root-of-time (Higuchi)** kinetics, releasing fastest at the start and slowing as the drug depletes.

Cumulative drug released vs time — common kinetic patterns

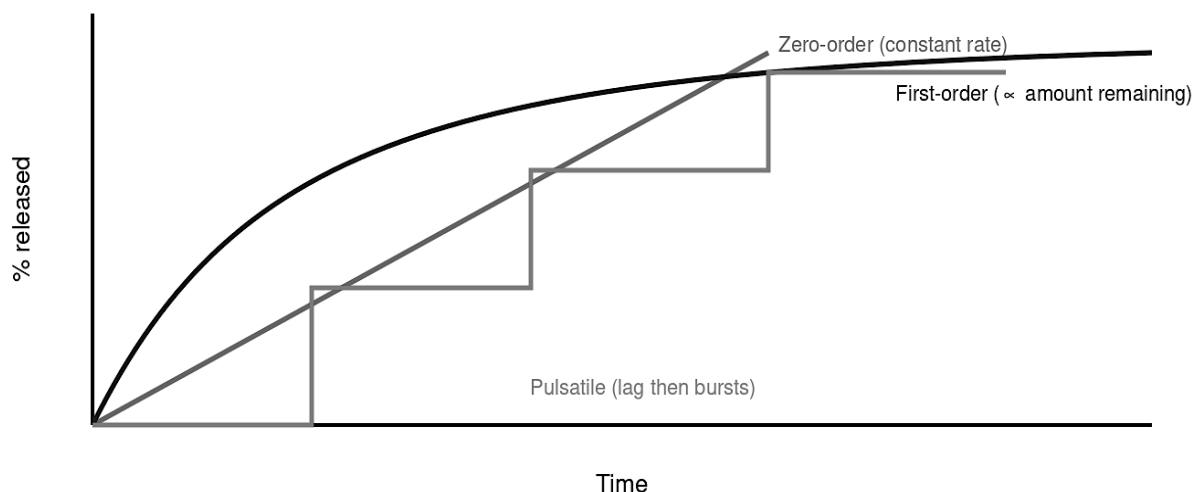


Figure 2.2. Idealised cumulative-release curves. Zero-order release rises linearly (constant rate, the controlled-release ideal). First-order release is fast initially then tapers as the remaining amount falls. Pulsatile release holds the drug back through a lag time and then delivers it in discrete bursts.

Extended release is an umbrella term covering both sustained and controlled release in much regulatory usage. **Delayed release** is different in kind: it releases the dose all at once, but only after a lag, the classic example being an enteric coating that withholds release until the dosage form has left the acidic stomach. **Pulsatile release** delivers the dose in separated bursts rather than continuously, which is the natural fit for chronotherapy (Section 2.4).

Rate control is achieved physically through a small number of recurring mechanisms: diffusion through a membrane or matrix, dissolution or erosion of a carrier, osmotic pumping driven by water influx, and ion exchange from a resin. These mechanisms reappear throughout the book, so it is worth fixing the kinetic vocabulary now. Table 2.2 lists the standard release models and what each tells the formulator.

Table 2.1. Common release-kinetic models and their interpretation.

Model	Equation form	Physical meaning
Zero-order	$Q = k \cdot t$	Constant release rate; reservoir/osmotic systems; the controlled-release ideal
First-order	$\ln(Q \text{ remaining}) \propto -k \cdot t$	Rate proportional to amount remaining; many simple matrices
Higuchi	$Q = k \cdot \sqrt{t}$	Diffusion from a planar matrix; release slows with time
Korsmeyer–Peppas	$Q = k \cdot t^n$	Empirical; exponent n diagnoses diffusion vs erosion control

The student should be able to read Figure 2.2 and say, from the shape of a release curve alone, which mechanism is likely operating.

2.3 Classification by spatial targeting

The second axis concerns where the drug goes, not how fast it is released. Targeting matters most for drugs that are toxic to healthy tissue, anticancer agents above all. The aim is to raise the drug concentration at the diseased site while lowering it everywhere else, widening the effective therapeutic window without changing the molecule.

Passive targeting exploits a property the diseased tissue already has. The most important is the **enhanced permeability and retention (EPR) effect** in solid tumours. Tumour blood vessels are formed hastily and are leaky, with gaps in the endothelium far larger than those of normal vessels, while tumour lymphatic drainage is poor. Nanoparticles in the right size range, broadly tens to around 200 nanometres, are too large to escape normal vasculature but small enough to slip through the leaky tumour vessels, where the poor drainage then keeps them trapped (Figure 2.3, left). The carrier accumulates in the tumour simply by circulating long enough, with no targeting ligand required. Long circulation is itself engineered, usually by coating the particle with poly(ethylene glycol) (PEG), a hydrophilic shield that delays recognition and clearance by the mononuclear phagocyte system. It should be noted that the magnitude and reliability of the EPR effect in human tumours is more variable than early animal work suggested, which is an active and honest debate in the field rather than a settled certainty.¹

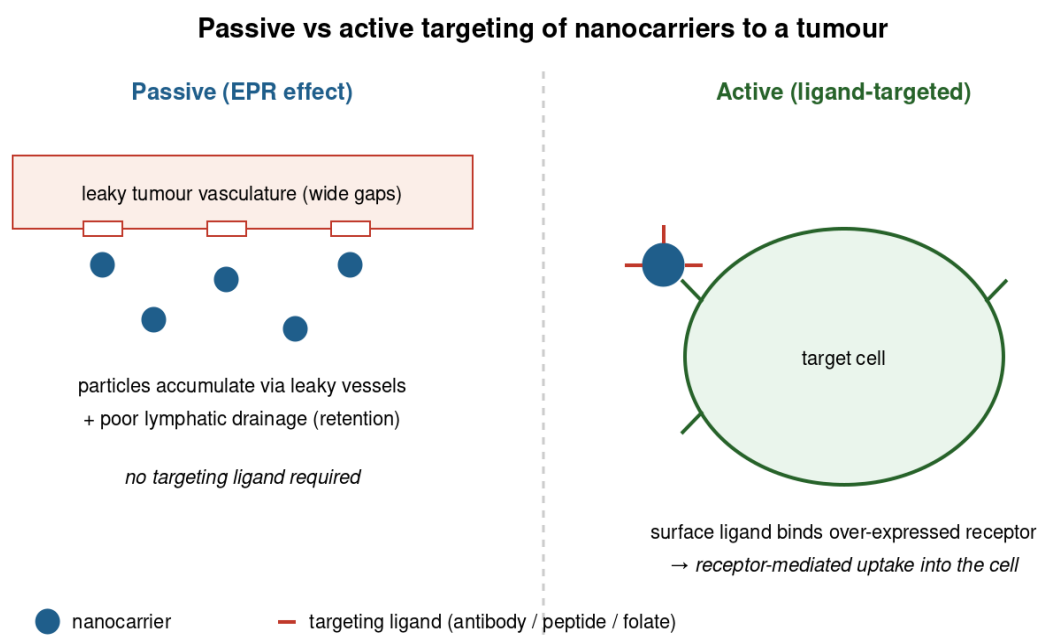


Figure 2.3. *Passive versus active targeting. Left: passive accumulation through the EPR effect, where long-circulating nanoparticles extravasate through leaky tumour vessels and are retained by poor lymphatic drainage. Right: active targeting, where a ligand on the particle surface binds a receptor over-expressed on the target cell and drives receptor-mediated uptake. Active targeting usually relies on first reaching the tissue passively.*

Active targeting adds a homing device. The carrier surface is decorated with a ligand, an antibody, an antibody fragment, a peptide, folate, or a sugar, that binds a receptor over-expressed on the target cell, promoting receptor-mediated internalisation (Figure 2.3, right). It is important to understand that active targeting does not replace passive targeting; in most cases the particle must first reach the tissue by circulating and extravasating, and the ligand

then improves cellular uptake and retention once it is there. Targeting can also be defined by level: **first-order** targeting to an organ or tissue, **second-order** to a specific cell type, and **third-order** to an intracellular compartment.

2.4 Classification by trigger and timing

The third axis concerns when release happens, and what sets it off.

Stimuli-responsive (smart) systems are built so that release is switched on by a signal, either a condition found at the diseased site or one applied from outside. Internal triggers include the lower pH of a tumour microenvironment or an endosome, the raised level of particular enzymes at a site of inflammation or in the colon, or the reducing environment inside a cell. External triggers include heat, light, ultrasound, and magnetic fields. A thermosensitive liposome, for example, is engineered to hold its cargo at body temperature and to release it sharply when the local tissue is warmed a few degrees by an applied source, concentrating release in the heated region. A pH-sensitive colon-targeted coating, examined in Unit II, withholds release until the rising pH of the distal gut dissolves it.

Chronotherapeutic systems match drug delivery to the body's biological rhythms. Many diseases are not constant through the day. Blood pressure and the risk of cardiac events rise in the early morning; the symptoms of rheumatoid arthritis and bronchial asthma are often worst in the pre-dawn hours; gastric acid secretion peaks at night. A drug delivered at a constant rate ignores these rhythms. A chronotherapeutic system instead times the peak drug level to the peak need, typically by combining a delayed-release lag with a subsequent rapid (pulsatile) release, so that a bedtime dose delivers its peak in the early morning when it is needed. Antihypertensive chronotherapy is the most developed example and is taken up as a case study in Unit IV.

2.5 Classification by route and carrier

The fourth axis is the most concrete and is the organising principle for much of this book. Systems can be grouped by the **route** they use, transdermal and mucosal, ocular, nasal, pulmonary, oral, parenteral and implantable, each of which is a response to one of the limitations of Chapter 1, and by the **carrier** that does the work.

The carriers fall into recognisable families, each the subject of a later chapter:

- **Vesicular carriers:** liposomes, niosomes, transferosomes, and related lipid vesicles that enclose drug within one or more bilayer membranes (Chapter 4, Chapter 19).
- **Polymeric systems:** nanoparticles, microspheres, micelles, dendrimers, and matrices built from biodegradable polymers such as PLGA (Chapter 3, Chapter 19).
- **Lipid particulate systems:** solid lipid nanoparticles (SLN) and nanostructured lipid carriers (NLC), and the lipid nanoparticles (LNP) that carry nucleic acids (Chapter 4).
- **Inorganic carriers:** mesoporous silica, gold, iron oxide, and metal–organic frameworks (Chapter 5).

The same product is read on every axis. Doxorubicin in a PEGylated liposome (the marketed product Doxil/Caelyx) is, by route, an intravenous injectable; by carrier, a vesicular liposome; by rate, a controlled-release reservoir; and by targeting, a passively (EPR) targeted system.² No single label captures it, which is exactly the point of using independent axes.

Table 2.2. *A working classification of novel delivery systems across the four axes.*

Axis	Categories	Representative examples
Rate of release	Immediate, sustained, controlled (zero-order), extended, delayed, pulsatile	Osmotic pump tablets; enteric-coated delayed release; matrix sustained release
Spatial targeting	Passive (EPR), active (ligand); organ / cell / intracellular	PEGylated liposomal doxorubicin (passive); antibody-decorated carriers (active)
Trigger / timing	Stimuli-responsive (pH, temperature, enzyme, light, magnetic); chronotherapeutic / pulsatile	Thermosensitive liposomes; bedtime chronotherapeutic antihypertensives
Route / carrier	Transdermal, mucosal, ocular, nasal, pulmonary, oral, parenteral/implant; vesicular, polymeric, lipid, inorganic carriers	Transdermal patches; DPIs; liposomes; PLGA microspheres; SLN; MSN

Quick-recall box — read a product on all four axes For any named NDDS, answer four questions: How is the rate controlled? Where is it targeted, and how? What triggers or times release? What route and carrier are used? If a single classification term feels insufficient, that is correct; good systems are multi-axis by design.

2.6 A note on overlap and on naming discipline

Students often lose marks by treating these categories as a checklist where a product fits one box. Examiners reward the opposite: the ability to take a product apart along several axes and to justify each placement. The categories overlap because real engineering combines strategies. A long-acting injectable antipsychotic is a depot (rate) delivered parenterally (route) using a PLGA microsphere (carrier). An inhaled insulin is a pulmonary-route product (route) whose entire reason for existing is to bypass first-pass and gastric degradation while delivering systemically. When asked to classify, name the axis first, then place the product on it, and acknowledge the overlaps rather than hiding them.

Chapter summary

Novel delivery systems are best classified along four independent axes rather than forced into a single hierarchy. By rate of release they range from immediate through sustained, controlled, extended, delayed, and pulsatile, with zero-order release the controlled-release ideal and a small set of physical mechanisms, diffusion, dissolution/erosion, osmosis, and ion exchange, doing the work. By spatial targeting they are passive, exploiting the EPR effect of leaky tumour vasculature and long circulation, or active, using surface ligands for receptor-mediated uptake, with active targeting usually building on passive accumulation. By trigger and timing they are stimuli-responsive, switched on by internal cues such as pH or enzymes or by external cues such as heat or light, or chronotherapeutic, timed to a circadian disease rhythm. By route and carrier they span transdermal, mucosal, ocular, nasal, pulmonary, oral and parenteral routes, and vesicular, polymeric, lipid and inorganic carriers. Crucially, a real product belongs to several categories at once, as PEGylated liposomal doxorubicin demonstrates, so the skill being trained is to describe a system on every axis rather than label it with one word.

Self-assessment

Short-answer (2 marks each)

1. Differentiate sustained release from controlled release.
2. What is zero-order release, and why is it the ideal for a controlled-release system?
3. Define the EPR effect and state the approximate particle size range that exploits it.
4. Why is a nanoparticle often coated with PEG?
5. Give one internal and one external trigger used in stimuli-responsive delivery.

Long-answer (5–10 marks each)

6. Classify novel delivery systems along the four axes described in this chapter, giving one example product for each axis.
7. Compare passive and active targeting with a labelled diagram, and explain why active targeting usually depends on passive accumulation first.
8. Explain the principle of chronotherapeutic delivery, giving a disease example, and describe the release profile such a system aims to produce.

Multiple-choice

9. A release profile in which the same amount of drug appears per unit time is: (a) first-order (b) Higuchi (c) zero-order (d) pulsatile
10. The EPR effect depends on: (a) active ligand binding (b) leaky tumour vessels and poor lymphatic drainage (c) gastric pH (d) hepatic metabolism
11. An enteric-coated tablet is an example of: (a) sustained release (b) delayed release (c) zero-order release (d) active targeting

Answer pointers. 1. Sustained = prolonged but not necessarily constant; controlled = predetermined, ideally constant rate. 2. Constant rate independent of amount remaining; matches input to elimination, holds steady plasma level. 3. Accumulation of nanocarriers in tumour via leaky vessels + poor drainage; ~tens to ~200 nm. 4. To prolong circulation by reducing recognition/clearance by the mononuclear phagocyte system. 5. Internal: pH/enzyme/redox; external: heat/light/ultrasound/magnetic. 6. Expect the four axes with examples per Table 2.1. 7. Figure 2.3 reproduced with explanation. 8. Match peak delivery to peak disease need (e.g. early-morning hypertension), lag then pulse. 9. (c). 10. (b). 11. (b).

References

1. Maeda H, Nakamura H, Fang J. The EPR effect for macromolecular drug delivery to solid tumors: improvement of tumor uptake, lowering of systemic toxicity, and distinct tumor imaging in vivo. *Adv Drug Deliv Rev.* 2013;65(1):71–9.
2. Barenholz Y. Doxil — the first FDA-approved nano-drug: lessons learned. *J Control Release.* 2012;160(2):117–34.
3. Jain KK, editor. *Drug Delivery Systems*. 3rd ed. New York: Humana Press; 2020.
4. Aulton ME, Taylor KMG, editors. *Aulton's Pharmaceuticals: The Design and Manufacture of Medicines*. 6th ed. Edinburgh: Elsevier; 2022.
5. Siepmann J, Siepmann F. Modeling of drug release from delivery systems based on hydroxypropyl methylcellulose (HPMC). *Adv Drug Deliv Rev.* 2012;64(Suppl):163–74.

Chapter 3 - Biodegradable and Biocompatible Polymers

Learning outcomes

After working through this chapter the student will be able to:

1. Distinguish biodegradable from biocompatible, and bulk erosion from surface erosion, and explain how each governs drug release.
2. Describe the chemistry, degradation, and delivery applications of PLA, PLGA and PCL, and explain how the lactide:glycolide ratio and molecular weight tune PLGA degradation rate.
3. Explain the origin, polycationic behaviour, and mucoadhesive/permeation-enhancing properties of chitosan.
4. Describe gelatin, its types A and B, and its uses in delivery and as a capsule and microsphere material.
5. Select a polymer for a stated delivery requirement and justify the choice on degradation rate, processing, and safety grounds.

3.1 Two words that are not the same: biodegradable and biocompatible

A polymer used inside the body must satisfy two separate requirements, and students routinely confuse them. **Biocompatible** means the material, and anything it releases as it breaks down, does not provoke an unacceptable biological response: no toxicity, no chronic inflammation, no immune reaction. **Biodegradable** means the material is broken down in the body into smaller fragments by chemical or enzymatic processes, ideally into substances the body already handles, so that nothing has to be surgically removed. A polymer can be biocompatible without being biodegradable (medical-grade silicone), and a material that degrades but releases toxic fragments is biodegradable yet not biocompatible. For a drug delivery carrier intended to release its dose and then disappear, both properties are required at once.

The polymers in this chapter are valued precisely because they degrade into endogenous or readily cleared products. That single feature, degradation into something the body already knows what to do with, is what makes them safe to inject as microspheres or implant as depots.

3.2 How a biodegradable matrix releases its drug: erosion mechanisms

Before looking at individual polymers, it helps to understand the two ways a solid polymer matrix can erode, because the erosion mode shapes the release curve (Figure 3.1).

In **bulk erosion**, water penetrates the whole matrix faster than the polymer can degrade. Hydrolysis then proceeds throughout the entire volume more or less simultaneously. The matrix keeps roughly its original size for much of its life and then loses mechanical integrity and collapses fairly suddenly. PLA and PLGA erode this way. A practical consequence is that bulk-eroding systems often show a multi-phase release: an initial burst of surface-associated drug, a slower diffusion-controlled middle phase, and a final accelerated phase as the matrix breaks down.

In **surface erosion**, the polymer degrades faster than water can penetrate, so breakdown is confined to the outer layer. The matrix shrinks steadily from the surface inward, like a bar of soap, while the core stays dry and intact. Polyanhydrides and poly(ortho esters) behave this way. Surface erosion is attractive because, for a constant surface area, it can give release close to zero-order, and the drug in the core is protected from water until its turn comes.

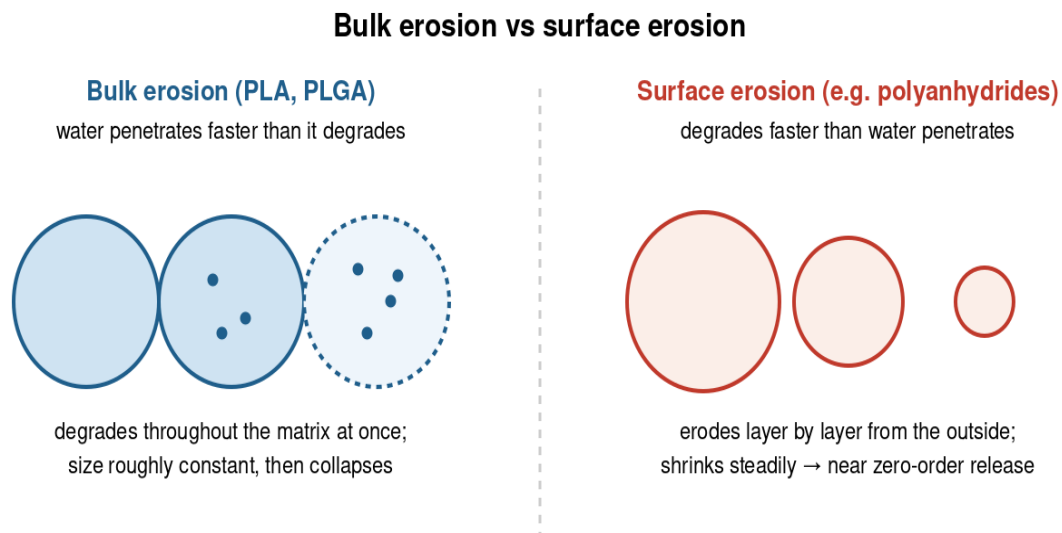


Figure 3.1. Bulk erosion versus surface erosion. In bulk erosion (left) water saturates the matrix and degradation occurs throughout, so the device keeps its size then collapses. In surface erosion (right) degradation outruns water penetration, so the device shrinks layer by layer, which favours near zero-order release.

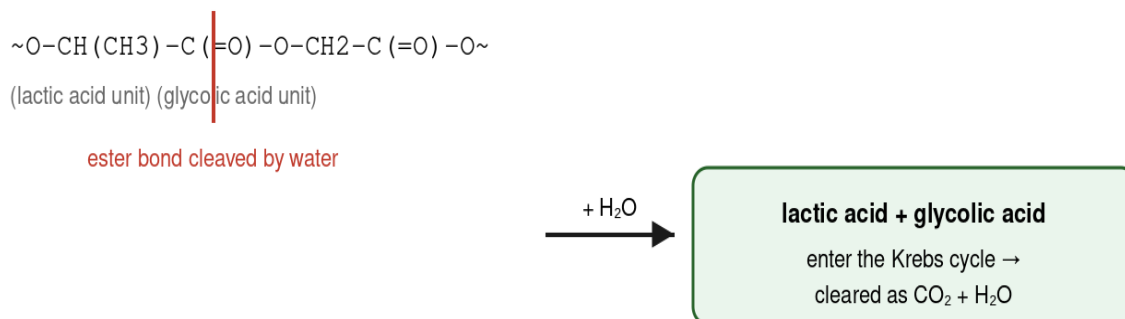
3.3 The polyesters: PLA, PGA and PLGA

The aliphatic polyesters are the workhorses of biodegradable delivery, and PLGA in particular is the most clinically established synthetic biodegradable polymer, present in many approved products.

3.3.1 Chemistry and degradation

Poly(lactic acid) or polylactide (PLA) is a polyester of lactic acid; **poly(glycolic acid) or polyglycolide (PGA)** is the polyester of glycolic acid; and **poly(lactic-co-glycolic acid) (PLGA)** is the copolymer of the two. All three share the same fundamental degradation chemistry: the ester bonds of the backbone are cleaved by water in a hydrolysis reaction, splitting the chain into shorter fragments and ultimately into the monomeric acids (Figure 3.2). The reaction needs no enzyme; it is simple hydrolytic chain scission, which is why these polymers degrade at a predictable rate set by their chemistry rather than by the variable enzyme levels of a given patient.

Hydrolysis of the ester backbone of PLA / PLGA



Degradation is bulk hydrolysis; the monomers are endogenous, hence biocompatible.

Figure 3.2. *Hydrolytic cleavage of the ester backbone of PLA/PLGA. Water attacks the ester linkages, releasing lactic acid and glycolic acid. Both monomers are normal metabolic intermediates that enter the citric-acid (Krebs) cycle and are eliminated as carbon dioxide and water, which is the basis of the polymers' biocompatibility.*

The biocompatibility follows directly from the products. Lactic acid is a normal metabolite, and glycolic acid is readily converted and cleared; both ultimately leave the body as carbon dioxide and water through the citric-acid cycle. There is no foreign residue to accumulate. The only caution is local: as the ester bonds hydrolyse they release acidic monomers, and inside a large bulk-eroding implant this can lower the local pH (autocatalysis), which both accelerates degradation in the core and can, for acid-sensitive drugs such as some peptides and proteins, threaten stability. Formulators manage this with buffering excipients and by controlling device size.

3.3.2 Tuning PLGA: ratio, molecular weight, and crystallinity

The reason PLGA dominates is that its degradation rate can be dialled in over a wide range, from a few weeks to several months, by changing the copolymer composition and molecular weight. Three levers matter.

First, the **lactide:glycolide ratio**. Glycolide units are more hydrophilic and the polymer takes up more water as glycolide content rises, but the relationship with degradation rate is not linear, because crystallinity also changes. The homopolymers PLA and PGA are relatively crystalline and resist water; the copolymer is most amorphous, and therefore most permeable to water and fastest to hydrolyse, near the middle of the range. A 50:50 lactide:glycolide PLGA degrades fastest, on the order of weeks, while compositions richer in lactide (such as 75:25 or 85:15) degrade more slowly, over months (Figure 3.3).¹ This is a frequent exam point and a frequent student error: 50:50 is fastest, not slowest.

Effect of lactide:glycolide ratio on PLGA degradation

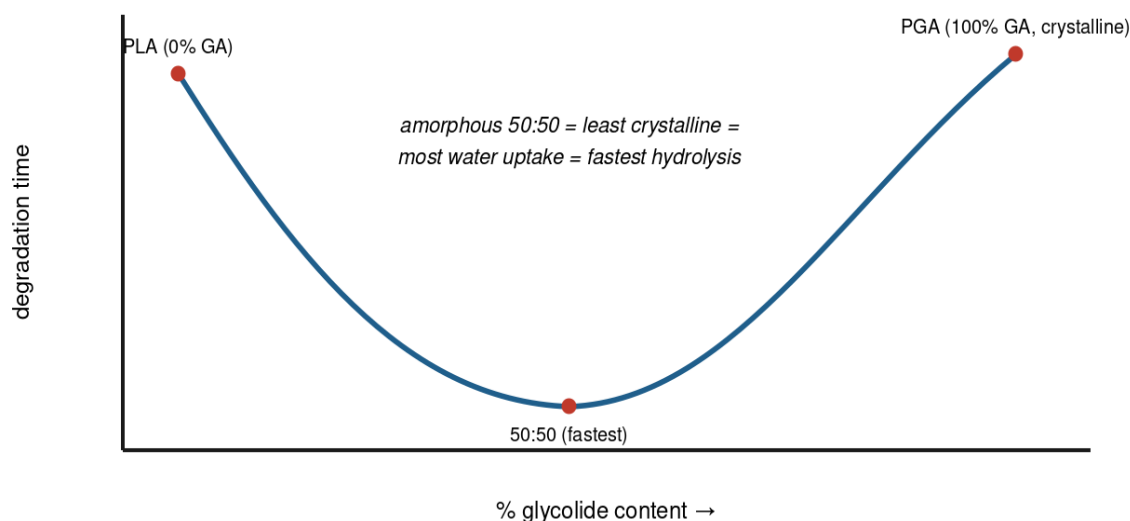


Figure 3.3. Schematic dependence of PLGA degradation time on copolymer composition. Degradation is fastest near the 50:50 lactide:glycolide ratio, where the copolymer is least crystalline and takes up the most water. The more crystalline homopolymers PLA and PGA, and lactide-rich copolymers, degrade more slowly.

Second, **molecular weight**. Higher-molecular-weight PLGA has longer chains, so more bonds must be cleaved before the fragments become water-soluble and the matrix loses integrity; higher molecular weight therefore means slower degradation and slower release.

Third, the **end-group chemistry**. PLGA with free carboxylic-acid end groups (uncapped) is more hydrophilic and degrades faster than ester-capped grades. Commercial PLGAs are sold specified by ratio, molecular weight (or inherent viscosity), and end group precisely so that a formulator can choose a grade to hit a target release duration.

A second mechanism, also under the formulator's control, is the glass-transition temperature. The glass-transition temperature (T_g) of most drug-loaded PLGA grades lies above body temperature, broadly in the region of 40–60 °C, so the matrix is glassy at 37 °C; absorbed water and hydrophilic drug act as plasticisers that lower T_g and can speed release.²

3.3.3 Applications

PLGA is the basis of the long-acting injectable depot, one of the great commercial successes of controlled delivery. The drug is dispersed in PLGA microspheres a few tens of micrometres across, injected intramuscularly or subcutaneously as a suspension, and released over weeks to months as the microspheres slowly hydrolyse. Marketed examples include leuprolide-loaded microspheres (Lupron Depot) for hormone-responsive cancers and endometriosis, the risperidone microsphere depot (Risperdal Consta) for schizophrenia, octreotide (Sandostatin LAR), triptorelin (Trelstar) and naltrexone (Vivitrol).^{3,4} These products are a recurring case study throughout the book because they show every principle of this chapter working at once: bulk erosion, ratio-tuned duration, and clearance to endogenous acids. Selected products are listed in Table 3.1.

Table 3.1. Selected FDA-approved PLGA-based depot products and their drugs.

Product	Drug	Therapeutic use
Lupron Depot	Leuprolide acetate	Prostate cancer, endometriosis
Risperdal Consta	Risperidone	Schizophrenia
Sandostatin LAR	Octreotide acetate	Acromegaly, neuroendocrine tumours
Trelstar	Triptorelin	Prostate cancer
Vivitrol	Naltrexone	Alcohol and opioid dependence
Zoladex (goserelin)*	Goserelin	Prostate and breast cancer

*Zoladex uses a PLGA implant rather than microspheres.³

PLA, being slower-degrading and more hydrophobic, is used where longer duration or greater mechanical strength is needed, including resorbable sutures and longer-acting implants. PGA alone is too fast-degrading and too crystalline for most controlled-release work and is used mainly in resorbable sutures.

3.4 Poly(ϵ -caprolactone) (PCL)

Poly(ϵ -caprolactone) (PCL) is another aliphatic polyester, but with very different physical behaviour from PLGA. It is semi-crystalline and notably hydrophobic, and its glass-transition temperature is far below body temperature, around $-60\text{ }^{\circ}\text{C}$, so at $37\text{ }^{\circ}\text{C}$ PCL is a soft, rubbery, ductile material rather than a glassy one.⁵ Its melting point is low, near $60\text{ }^{\circ}\text{C}$, which makes it easy to process by melt methods.

Its defining feature for delivery is slow degradation. PCL hydrolyses much more slowly than PLGA, and complete resorption can take on the order of two to three years. This makes PCL the polymer of choice for very long-acting systems. The most prominent example is a subdermal contraceptive implant that releases over multiple years, where the slow, steady erosion of PCL matches the multi-year dosing requirement. Its rubbery flexibility and low melting point also make it valuable in tissue-engineering scaffolds and in three-dimensional printing of resorbable devices. The trade-off is that for a system meant to release over weeks, PCL is far too slow, which is why it and PLGA are complementary rather than competing materials.

3.5 Chitosan: a natural polycation

Chitosan is a natural polysaccharide obtained by partial deacetylation of chitin, the structural polymer of crustacean shells and fungal cell walls. Chemically it is a copolymer of glucosamine and N-acetylglucosamine; the degree of deacetylation sets how many free amino groups it carries, and that number governs most of its useful behaviour.

The signature property is that chitosan is a **polycation** in acidic solution. Its many primary amino groups become protonated and positively charged below a pH of about 6.5, which gives chitosan three properties prized in delivery. It becomes water-soluble in mild acid, allowing easy processing. Its positive charge makes it **mucoadhesive**: it binds electrostatically to the negatively charged sialic-acid residues of mucus, so a chitosan-based formulation lingers at a mucosal surface instead of being swept away, which prolongs contact time at nasal, ocular, buccal and intestinal sites. And the same cationic charge lets chitosan act as a **permeation enhancer**: it transiently opens the tight junctions between epithelial cells, widening the paracellular route and improving the absorption of poorly permeable

drugs, peptides among them. Chitosan is also biodegradable, broken down by lysozyme and other enzymes, and it carries mild intrinsic antimicrobial and haemostatic activity.

These features make chitosan a favourite for mucosal and nasal delivery, for nanoparticles formed by ionic gelation with a polyanion such as tripolyphosphate (a gentle, organic-solvent-free method well suited to fragile biologics), and for wound dressings. Its main limitation is that its solubility and charge depend on pH; above its pKa it loses charge and precipitates, so formulations must control pH carefully.

3.6 Gelatin

Gelatin is a protein obtained by partial hydrolysis of collagen, the most abundant structural protein of animal connective tissue, skin and bone. Because it comes from a natural protein it is biodegradable, biocompatible, and very widely used; its long history in food and pharmaceuticals gives it an excellent safety record.

Gelatin exists in two types defined by how the collagen was processed. **Type A** is produced by acid pretreatment and has a higher isoelectric point (around pH 7 to 9), so it carries a net positive charge over much of the physiological range. **Type B** is produced by alkaline pretreatment and has a lower isoelectric point (around pH 4.5 to 5), carrying a net negative charge near neutral pH. The choice of type therefore lets a formulator set the sign of the polymer's charge, which matters for binding charged drugs and for forming complexes.

A key physical property is its thermoreversible sol–gel behaviour: a warm gelatin solution sets to a gel on cooling and melts again on warming, which is exploited in soft- and hard-capsule shells and in forming microspheres. In delivery, gelatin is used to make nanoparticles and microspheres (often hardened by crosslinking, classically with glutaraldehyde, though gentler crosslinkers are now preferred to avoid toxicity), as a coating and matrix material, and as the classical shell of capsules. Its protein nature is also its main drawback: there is potential for immunogenicity and batch variability, and the animal origin raises sourcing and acceptability concerns, which has driven interest in recombinant and plant-derived alternatives.

Table 3.2 brings the five polymers together for comparison.

Table 3.2. *Comparative properties of the principal biodegradable delivery polymers.*

Polymer	Origin	Degradation product / mode	Approx. degradation time	Key delivery role
PLA	Synthetic polyester	Lactic acid; bulk erosion	Months to >1 year	Longer-acting implants, sutures
PLGA	Synthetic copolymer	Lactic + glycolic acid; bulk erosion	Weeks to months (ratio-tunable)	Depot microspheres, nanoparticles
PCL	Synthetic polyester	6-hydroxycaproic acid; bulk erosion	~2–3 years	Very long-acting implants, scaffolds
Chitosan	Deacetylated chitin (natural)	Enzymatic (lysozyme)	Variable	Mucoadhesive/nasal, ionic-gelation NPs, permeation enhancer
Gelatin	Hydrolysed	Enzymatic	Variable	Capsules,

	collagen (natural)	(proteases)		micro/nanospheres, coatings
--	-----------------------	-------------	--	--------------------------------

Quick-recall box — picking a polyester Need release over weeks and clearance to endogenous acids? PLGA, and choose the ratio: 50:50 for fastest (weeks), lactide-rich for slower (months). Need years? PCL. Need mucoadhesion and gentle, solvent-free nanoparticles for a peptide? Chitosan by ionic gelation. Remember: for PLGA, 50:50 degrades fastest because it is least crystalline.

Chapter summary

Biocompatibility and biodegradability are distinct requirements, and a delivery carrier intended to release and then vanish must meet both. How a solid matrix releases its drug depends on whether it undergoes bulk erosion, degrading throughout its volume as PLA and PLGA do, or surface erosion, shrinking from the outside as polyanhydrides do, the latter favouring zero-order release. The aliphatic polyesters dominate the field. PLA, PGA and PLGA all degrade by simple hydrolysis of their ester backbones into lactic and glycolic acid, normal metabolites cleared through the citric-acid cycle, which is the source of their safety. PLGA is the most used because its degradation can be tuned from weeks to months through three levers: the lactide:glycolide ratio (50:50 is fastest because it is least crystalline and most water-permeable), the molecular weight (higher is slower), and the end-group chemistry. This tunability underlies the long-acting injectable depots such as Lupron Depot and Risperdal Consta. PCL is a slow, rubbery polyester used for multi-year implants. Among natural polymers, chitosan is a polycation that is mucoadhesive and a tight-junction permeation enhancer and forms gentle ionic-gelation nanoparticles, while gelatin, a hydrolysed collagen available as charge-distinct Type A and Type B, serves in capsules, microspheres and coatings. Polymer choice is always a match between the required release duration and the degradation behaviour of the material.

Self-assessment

Short-answer (2 marks each)

1. Differentiate biodegradable and biocompatible.
2. Why does a 50:50 PLGA degrade faster than an 85:15 PLGA?
3. State the degradation products of PLGA and explain why they are non-toxic.
4. Why is chitosan mucoadhesive and a permeation enhancer?
5. Distinguish Type A and Type B gelatin.

Long-answer (5–10 marks each)

6. Compare bulk erosion and surface erosion with a labelled diagram, and explain the release profile each tends to produce.
7. Explain, with the three controlling variables, how the degradation rate of PLGA is tuned to a target release duration, and relate this to a marketed depot product.
8. Write a comparative account of PLA, PLGA and PCL covering chemistry, degradation time, and a suitable application for each.

Multiple-choice

9. Which PLGA composition degrades most rapidly? (a) 85:15 (b) 75:25 (c) 50:50 (d) 100:0 (PLA)

10. PCL is best suited to: (a) release over 24 hours (b) release over a few days (c) multi-year implants (d) immediate release
11. Chitosan becomes cationic and soluble because, in mild acid, it: (a) loses amino groups (b) protonates its amino groups (c) becomes crosslinked (d) crystallises

Answer pointers. 1. Biodegradable = broken down in body; biocompatible = no unacceptable biological response. 2. 50:50 is least crystalline, takes up most water, hydrolyses fastest. 3. Lactic + glycolic acid; endogenous metabolites cleared via Krebs cycle to $\text{CO}_2 + \text{H}_2\text{O}$. 4. Polycationic amino groups bind anionic mucus (adhesion) and transiently open tight junctions (enhancement). 5. Type A acid-treated, high pI, net positive; Type B alkali-treated, low pI, net negative near neutral pH. 6. Figure 3.2 reproduced + zero-order favoured by surface erosion. 7. Ratio (50:50 fastest), MW (higher slower), end group (uncapped faster); link to Lupron Depot or Risperdal Consta. 8. Comparative per Table 3.1. 9. (c). 10. (c). 11. (b).

References

1. Makadia HK, Siegel SJ. Poly lactic-co-glycolic acid (PLGA) as biodegradable controlled drug delivery carrier. *Polymers (Basel)*. 2011;3(3):1377–97.
2. Bohr A, Water JJ, Wang Y, et al. Glass transition temperature of PLGA particles and the influence on drug delivery applications. *Polymers (Basel)*. 2022;14(5):993.
3. Park K, Skidmore S, Hadar J, et al. Injectable, long-acting PLGA formulations: analyzing PLGA and understanding microparticle formation. *J Control Release*. 2019;304:125–34.
4. Wang Y, Qin B, Xia G, Choi SH. FDA's poly (lactic-co-glycolic acid) research program and regulatory outcomes. *AAPS J*. 2021;23(4):92.
5. Woodruff MA, Hutmacher DW. The return of a forgotten polymer — polycaprolactone in the 21st century. *Prog Polym Sci*. 2010;35(10):1217–56.
6. Bernkop-Schnürch A, Dünnhaupt S. Chitosan-based drug delivery systems. *Eur J Pharm Biopharm*. 2012;81(3):463–9.

Chapter 4 - Lipids and Surfactants

Learning outcomes

After working through this chapter the student will be able to:

1. Explain the amphiphilic structure of phospholipids and how the hydrophobic effect drives their self-assembly into bilayers, micelles and liposomes.
2. Describe lecithin and the principal phospholipids used in delivery, and state where a drug of given polarity sits within a liposome.
3. Define the hydrophilic–lipophilic balance (HLB) and use it to select a surfactant for a given role, distinguishing the Span and Tween series.
4. Describe the structure, advantages, and limitations of solid lipid nanoparticles (SLN) and explain how nanostructured lipid carriers (NLC) overcome those limitations.
5. Name verified lipid-based products, including liposomal and lipid-nanoparticle medicines, and relate their design to the principles of this chapter.

4.1 Why lipids are central to delivery

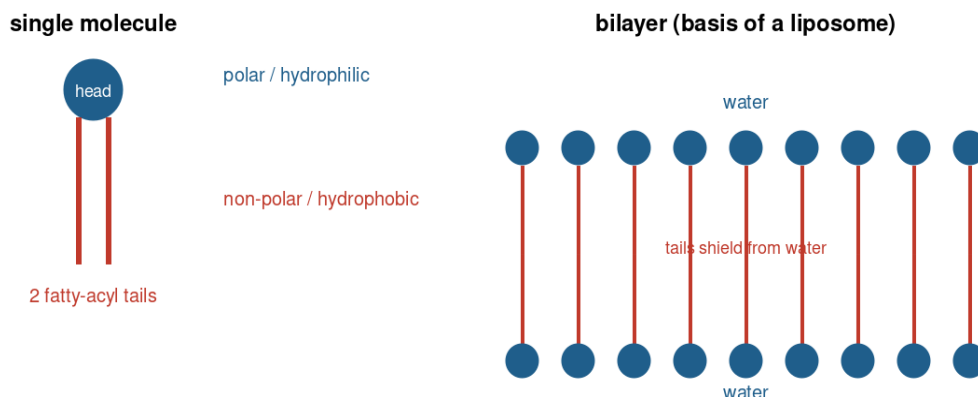
Cell membranes are made of lipid. It is no accident that lipid-based carriers are among the most biocompatible and clinically successful delivery systems: they are built from the same kind of molecules the body uses for its own membranes, and they interact gently with biological barriers. Lipids serve delivery in two complementary ways. As **structural building blocks**, amphiphilic lipids self-assemble into ordered structures, bilayers, vesicles, micelles, that can carry drug. As **functional excipients**, surfactants stabilise emulsions and nanoparticles, solubilise poorly soluble drugs, and modify how a drug interacts with membranes. This chapter covers both: first the phospholipids that build vesicles, then the surfactants graded by HLB, and finally the solid-lipid nanoparticle systems that combine the two.

4.2 Amphiphiles and the hydrophobic effect

The single idea behind every structure in this chapter is **amphiphilicity**: a molecule with a water-loving (hydrophilic) part and an oil-loving (lipophilic) part on the same skeleton. Such a molecule cannot satisfy both parts in a simple solution. Placed in water, the hydrophilic head is comfortable but the lipophilic tail is not, and the system minimises the unfavourable contact between the tails and water by hiding the tails away from it. This driving force, the tendency of non-polar groups to cluster together and exclude water, is the **hydrophobic effect**, and it does the assembly work for free, without any input of energy by the formulator.

What structure forms depends on the molecular shape. A single-tailed surfactant, roughly cone-shaped, packs into a small sphere with tails inward and heads out, a **micelle**, which can carry a lipophilic drug dissolved in its oily core. A double-tailed phospholipid, roughly cylindrical, packs instead into a flat sheet two molecules thick, a **bilayer**, with the tails sandwiched in the middle and the heads facing the water on both sides (Figure 4.1). A bilayer that closes on itself to remove its exposed edges becomes a hollow sphere, a **liposome**.

Phospholipid: an amphiphile that self-assembles into a bilayer



Hydrophobic effect drives tails inward; a closed bilayer vesicle is a liposome, with an aqueous core (holds water-soluble drug) and a bilayer (holds lipophilic drug).

Figure 4.1. A phospholipid is an amphiphile with a polar head and two non-polar acyl tails. Driven by the hydrophobic effect, such molecules arrange tail-to-tail into a bilayer that shields the tails from water. A closed bilayer vesicle is a liposome, with an aqueous core able to carry water-soluble drug and a lipid bilayer able to carry lipophilic drug.

This last point is worth fixing because it is examined often. A liposome has two distinct compartments. Its **aqueous core** dissolves and carries hydrophilic (water-soluble) drug; its **lipid bilayer** dissolves and carries lipophilic drug within the membrane itself. A liposome can therefore, in principle, carry drugs of either polarity, and where a given drug sits is decided by its own solubility.

4.3 Lecithin and the phospholipids

The phospholipids are the natural amphiphiles that build biological membranes and most liposomes. Each has a glycerol (or sphingosine) backbone bearing two fatty-acyl chains (the hydrophobic tails) and a phosphate-containing head group (the hydrophilic head). The identity of the head group names the phospholipid: **phosphatidylcholine**, **phosphatidylethanolamine**, **phosphatidylserine**, **phosphatidylglycerol**, and so on.

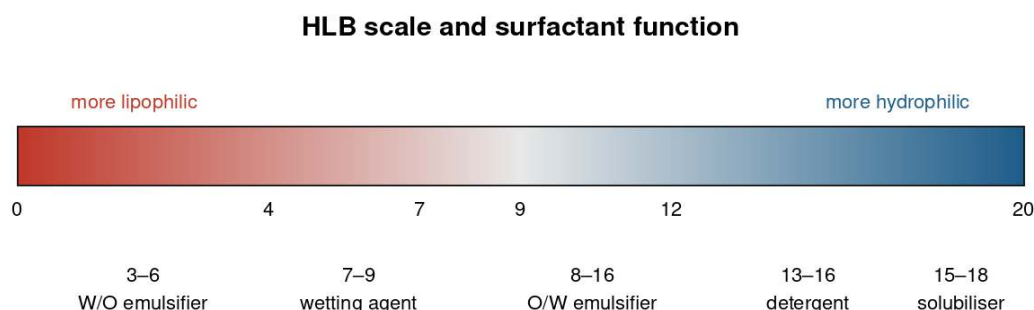
Lecithin is the common name for a natural mixture of phospholipids, rich in phosphatidylcholine, obtained from sources such as egg yolk and soybean. It is inexpensive, biocompatible, and amphiphilic, which makes it a workhorse emulsifier and the classic raw material for liposomes. Because it is a natural mixture, its exact composition and the saturation of its acyl chains vary with source, and that variation matters: the **transition temperature**, the temperature at which the bilayer changes from an ordered gel phase to a fluid liquid-crystalline phase, depends on chain length and saturation. Saturated, long-chain phospholipids give rigid, well-sealed, leak-resistant bilayers; unsaturated chains give fluid, more permeable membranes. Formulators choose phospholipids by transition temperature to set how tightly the liposome holds its cargo at body temperature. **Cholesterol** is almost always added to liposome bilayers because it sits between the phospholipid molecules and modulates membrane fluidity, increasing rigidity and reducing leakage of entrapped drug, which improves stability in the blood.

Liposomes are the most clinically validated lipid carrier. Verified marketed examples include liposomal amphotericin B (AmBisome, introduced in Chapter 1), which reduces nephrotoxicity, and PEGylated liposomal doxorubicin (Doxil/Caelyx), where a PEG coating prolongs circulation and the liposome reduces cardiotoxicity while exploiting EPR-mediated tumour accumulation.^{1,2} These show the two great gifts of the liposome: changing a drug's biodistribution to spare a vulnerable organ, and controlling its release.

4.4 Surfactants and the HLB system

Surfactants (surface-active agents) are amphiphiles used as functional excipients rather than as the main structural lipid. They lower interfacial tension, which lets them stabilise emulsions, wet hydrophobic powders, and solubilise poorly soluble drugs into micelles. Choosing the right surfactant for a job is made systematic by the **hydrophilic–lipophilic balance (HLB)**, a number, originally devised by Griffin, that expresses the relative weight of the hydrophilic and lipophilic portions of a surfactant on a scale running from 0 (entirely lipophilic) to 20 (entirely hydrophilic) for non-ionic surfactants.³

The value predicts the surfactant's natural role (Figure 4.2). A low-HLB surfactant is predominantly oil-loving and sits mostly in the oil phase, so it stabilises a **water-in-oil (W/O)** emulsion; HLB roughly 3 to 6 is the W/O emulsifier range. Around HLB 7 to 9 the surfactant is balanced enough to act as a wetting agent. A high-HLB surfactant is predominantly water-loving and stabilises an **oil-in-water (O/W)** emulsion; HLB roughly 8 to 16 covers O/W emulsifiers, with the higher values (around 13 to 16) acting as detergents and the highest (around 15 to 18) acting as solubilisers that form micelles to carry lipophilic drug. A simple way to remember the logic: the surfactant stabilises the emulsion in which its larger, more soluble portion forms the continuous phase.



HLB = hydrophilic–lipophilic balance. Low HLB favours water-in-oil; high HLB favours oil-in-water.

Span series = low HLB (lipophilic); Tween series = high HLB (hydrophilic).

Figure 4.2. The HLB scale and the function associated with each range. Low HLB surfactants (lipophilic) stabilise water-in-oil systems; high HLB surfactants (hydrophilic) stabilise oil-in-water systems, act as detergents, and solubilise drugs. The Span series occupies the low end, the Tween series the high end.

A practical formulation point: two surfactants are frequently blended to reach a target HLB, and the HLB of a blend is the weighted average of the components' values by mass fraction. This lets a formulator pair a low-HLB Span with a high-HLB Tween to hit precisely the value an emulsion needs.

4.4.1 The Span and Tween series

Two families dominate pharmaceutical practice and are routinely examined. The **Spans** are sorbitan fatty-acid esters; they are lipophilic, low-HLB, and oil-soluble. The **Tweens** (polysorbates) are the same sorbitan esters made hydrophilic by adding poly(oxyethylene) chains; they are high-HLB and water-soluble. Pairing a Span with the corresponding Tween is a classic way to formulate stable emulsions. Table 4.2 lists the common members with their chemical identity and approximate HLB; these values are worth committing to memory for examinations.

Table 4.1. Common Span and Tween surfactants.

Trade name	Chemical identity	Approx. HLB
Span 80	Sorbitan monooleate	4.3
Span 60	Sorbitan monostearate	4.7
Span 40	Sorbitan monopalmitate	6.7
Span 20	Sorbitan monolaurate	8.6
Tween 80	Polysorbate 80 (POE-sorbitan monooleate)	15.0
Tween 60	Polysorbate 60 (POE-sorbitan monostearate)	14.9
Tween 40	Polysorbate 40 (POE-sorbitan monopalmitate)	15.6
Tween 20	Polysorbate 20 (POE-sorbitan monolaurate)	16.7

Table 4.2. HLB ranges and their applications.

HLB range	Application
3–6	Water-in-oil (W/O) emulsifier
7–9	Wetting agent
8–16	Oil-in-water (O/W) emulsifier
13–16	Detergent
15–18	Solubiliser

Polysorbate 80 (Tween 80) deserves special mention because it appears as a solubiliser and stabiliser in many parenteral and biologic products, including several vaccines, where it helps keep proteins and other hydrophobic components in stable dispersion.

4.5 Solid lipid nanoparticles and nanostructured lipid carriers

Liposomes are excellent but can be physically and chemically unstable on storage, and leakage of drug from the bilayer can be a problem. The **solid lipid nanoparticle (SLN)** was developed to combine the biocompatibility of lipids with the stability of a solid matrix. An SLN is a nanoparticle whose core is a lipid that is solid at body temperature (such as a triglyceride, a fatty acid, or a wax), stabilised by a surfactant shell. Because the core is solid rather than liquid, molecular mobility within it is low, which slows drug diffusion and gives controlled release, and protects sensitive drug from degradation. SLNs can be made from physiologically tolerated lipids by scalable, organic-solvent-free methods such as high-pressure homogenisation, which is a real manufacturing advantage.

SLNs have one structural weakness, illustrated in Figure 4.3. A single, pure solid lipid tends to crystallise into a near-perfect, highly ordered lattice. A perfect crystal has little room between its molecules to accommodate drug, so drug loading is limited; worse, if the lipid continues to crystallise into a more ordered form during storage, it can squeeze out the drug it was carrying, an effect called drug expulsion that shortens shelf life.

The **nanostuctured lipid carrier (NLC)** is the second-generation answer. An NLC is made by blending the solid lipid with a smaller amount of a liquid lipid (an oil). The mismatch between the two lipids prevents the formation of a perfect crystal; the matrix solidifies into an imperfect, disordered structure with many small pockets, including liquid-oil compartments, that can hold more drug and hold it more securely (Figure 4.3). The result is higher drug loading and far less expulsion on storage, while keeping the solid-matrix benefits of controlled release and protection. SLNs and NLCs are studied for oral, topical, ocular and parenteral routes, and are particularly attractive for poorly soluble drugs.

Solid lipid nanoparticle (SLN) vs nanostructured lipid carrier (NLC)

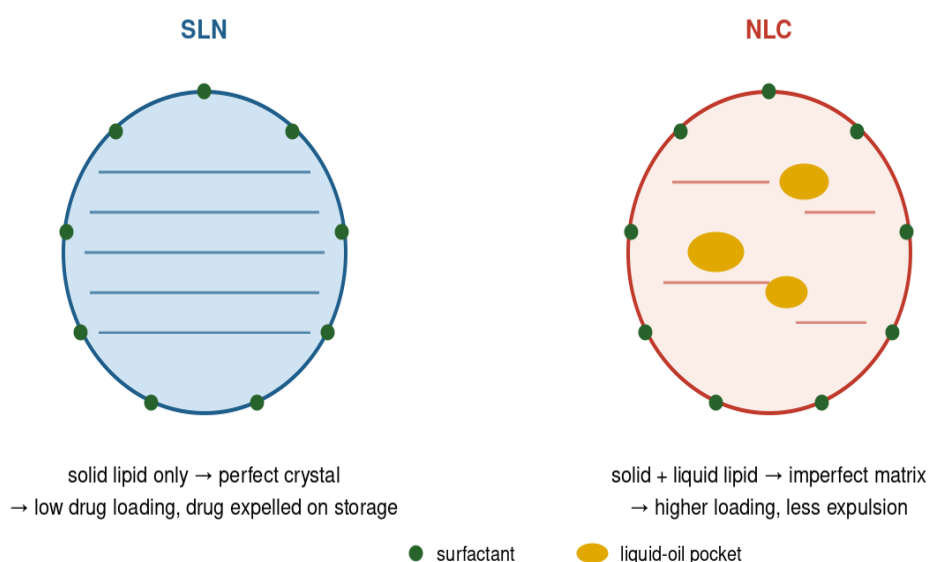


Figure 4.3. Internal structure of an SLN versus an NLC. The SLN core is a single solid lipid that crystallises into a near-perfect lattice, leaving little room for drug and tending to expel it as crystallisation perfects on storage. The NLC core blends solid and liquid lipid, giving a deliberately imperfect matrix with oil pockets that accommodates more drug and retains it better. Both are stabilised by a surfactant shell.

Table 4.3 compares the three lipid nanocarriers covered in this chapter.

Table 4.3. Comparison of the principal lipid nanocarriers.

Feature	Liposome	SLN	NLC
Core	Aqueous, enclosed by bilayer(s)	Single solid lipid	Solid + liquid lipid blend
Drug location	Aqueous core (hydrophilic) + bilayer (lipophilic)	Within solid lipid matrix	Within imperfect matrix + oil pockets

Main strength	Carries either polarity; alters biodistribution	Solid matrix → controlled release, protection	High loading, low expulsion
Main limitation	Storage instability, drug leakage	Limited loading, drug expulsion	More complex composition
Marketed examples	AmBisome, Doxil	Mostly cosmetic/topical and investigational	Mostly cosmetic/topical and investigational

4.6 Lipid nanoparticles for nucleic acids

A development of major recent importance is the **lipid nanoparticle (LNP)** used to deliver nucleic acids, most prominently the messenger-RNA COVID-19 vaccines (Comirnaty and Spikevax).⁴ These are distinct from classical liposomes. An LNP combines four lipid components: an **ionisable lipid** that is near-neutral at physiological pH but becomes positively charged in the acidic environment of the endosome, which is the key to escaping into the cytoplasm; a **helper phospholipid** and **cholesterol** for structure; and a **PEG-lipid** that controls particle size and prolongs circulation. The ionisable-lipid trick solves two problems at once: a permanently cationic lipid would be toxic and rapidly cleared, whereas an ionisable lipid is quiet in the bloodstream and only switches on its membrane-disrupting positive charge once inside the acidic endosome, releasing its RNA cargo into the cell. The mRNA-LNP vaccines are the clearest demonstration to date that delivery, not just the active molecule, can be the decisive enabling technology, and they are taken up again as a case study in Unit IV.

Quick-recall box — lipids at a glance Liposome: bilayer vesicle; aqueous core holds hydrophilic drug, bilayer holds lipophilic drug; add cholesterol to reduce leakage. HLB: low (3–6) = W/O emulsifier (Spans); high (8–18) = O/W emulsifier/solubiliser (Tweens); blend HLB is the mass-weighted average. SLN→NLC: blending in liquid lipid makes the crystal imperfect, raising loading and stopping drug expulsion. LNP: ionisable lipid switches on in the endosome to release nucleic acid.

Chapter summary

Lipid carriers succeed because they are built from the same amphiphilic molecules as biological membranes. Amphiphilicity plus the hydrophobic effect drives self-assembly for free: single-tailed surfactants form micelles, double-tailed phospholipids form bilayers, and a closed bilayer is a liposome with an aqueous core for hydrophilic drug and a lipid membrane for lipophilic drug. Lecithin, a natural phosphatidylcholine-rich mixture, is the classic liposome and emulsifier material, and cholesterol is added to rigidify bilayers and cut leakage; transition temperature, set by chain length and saturation, governs how tightly the membrane holds its cargo. Surfactants used as excipients are graded by HLB, where low values suit water-in-oil emulsification and high values suit oil-in-water emulsification, detergency and solubilisation; the lipophilic Spans and hydrophilic Tweens are the dominant series and are often blended to a target HLB calculated as a mass-weighted average. Solid lipid nanoparticles give a stable solid matrix for controlled release but suffer limited loading and drug expulsion because pure lipids crystallise too perfectly; nanostructured lipid carriers fix this by blending in a liquid lipid to create a deliberately imperfect, higher-capacity matrix. Lipid nanoparticles with an ionisable lipid component have made nucleic-acid medicines

such as the mRNA COVID-19 vaccines possible, and stand as the strongest modern proof that the delivery system can be the enabling technology.

Self-assessment

Short-answer (2 marks each)

1. Explain the hydrophobic effect and how it drives bilayer formation.
2. Where in a liposome does a hydrophilic drug reside, and where a lipophilic drug?
3. Define HLB and state the approximate range for an O/W emulsifier.
4. Why is cholesterol added to liposome bilayers?
5. Why does an SLN made from a single pure lipid tend to expel its drug on storage?

Long-answer (5–10 marks each)

6. With a labelled diagram, explain how phospholipids self-assemble into a liposome and how the liposome can carry drugs of either polarity.
7. Describe the HLB system, the role of the Span and Tween series, and how the HLB of a binary surfactant blend is calculated, with an application example.
8. Compare SLN and NLC in structure, drug loading and storage stability, with a diagram, and explain why the NLC was developed.

Multiple-choice

9. A surfactant of HLB 4.5 is best suited as a: (a) solubiliser (b) O/W emulsifier (c) W/O emulsifier (d) detergent
10. In an mRNA-LNP, the component that becomes positively charged in the endosome is the: (a) PEG-lipid (b) cholesterol (c) ionisable lipid (d) helper phospholipid
11. Tween 80 compared with Span 80 is: (a) more lipophilic, lower HLB (b) more hydrophilic, higher HLB (c) identical in HLB (d) insoluble in water

Answer pointers. 1. Non-polar tails cluster to exclude water; minimising tail–water contact drives ordered assembly. 2. Hydrophilic in aqueous core; lipophilic in the bilayer. 3. Relative weight of hydrophilic vs lipophilic parts (0–20); O/W $\approx$ 8–16. 4. Rigidifies the bilayer, reduces fluidity and drug leakage, improves blood stability. 5. Pure lipid crystallises to a near-perfect lattice with little room for drug, expelling it as ordering perfects. 6. Figure 4.1 reproduced with both compartments. 7. HLB scale + Span/Tween + mass-weighted blend average + example. 8. Per Table 4.3 and Figure 4.3. 9. (c). 10. (c). 11. (b).

References

1. Stone NRH, Bicanic T, Salim R, Hope W. Liposomal amphotericin B (AmBisome): a review. *Drugs*. 2016;76(4):485–500.
2. Barenholz Y. Doxil — the first FDA-approved nano-drug: lessons learned. *J Control Release*. 2012;160(2):117–34.
3. Griffin WC. Classification of surface-active agents by HLB. *J Soc Cosmet Chem*. 1949;1(5):311–26.
4. Schoenmaker L, Witzigmann D, Kulkarni JA, et al. mRNA-lipid nanoparticle COVID-19 vaccines: structure and stability. *Int J Pharm*. 2021;601:120586.
5. Müller RH, Radtke M, Wissing SA. Solid lipid nanoparticles (SLN) and nanostructured lipid carriers (NLC) in cosmetic and dermatological preparations. *Adv Drug Deliv Rev*. 2002;54(Suppl 1):S131–55.

Chapter 5 - Inorganic Systems for Drug Delivery

Learning outcomes

After working through this chapter the student will be able to:

1. Explain why inorganic carriers offer properties, such as high surface area, tunable pores, magnetism and optical activity, that organic carriers cannot match, and state the trade-off in biodegradability.
2. Describe the structure and delivery use of mesoporous silica nanoparticles, including the gatekeeper concept for stimulus-triggered release.
3. Explain surface plasmon resonance in gold nanoparticles and how it is exploited for photothermal therapy, imaging, and triggered release.
4. Describe superparamagnetic iron oxide nanoparticles and their dual role in magnetically guided delivery and as MRI contrast agents, citing a clinically approved example.
5. Outline the architecture of metal–organic frameworks and their promise and limitations as ultra-high-capacity drug hosts.

5.1 Why use an inorganic carrier at all

The polymers and lipids of the previous chapters degrade cleanly into harmless products, which is their great virtue. Inorganic carriers give that virtue up, at least in part, in exchange for physical properties no organic material can provide. A silica particle can be made with an enormous, precisely ordered internal pore network. A gold particle can convert light into heat. An iron oxide particle responds to a magnetic field and sharpens a magnetic-resonance image. A metal–organic framework can hold a larger weight of drug per gram than almost any other material known. These are not properties one can engineer into a polymer; they come from the inorganic solid state itself.

The cost of these properties is the central caution of the chapter: most inorganic carriers do not biodegrade in the way a polyester does. Their long-term fate in the body, how, and how completely, they are cleared, is the principal barrier between the laboratory promise of these systems and routine clinical use. The reader should hold both facts together throughout: striking, useful physics on one side, and an unresolved question of long-term safety and clearance on the other. This is why, with important exceptions, inorganic carriers are at an earlier stage of clinical translation than lipid and polymer systems.

5.2 Mesoporous silica nanoparticles

Silica is silicon dioxide, the same chemistry as sand and quartz, and amorphous silica has a long history as a pharmaceutical excipient. The **mesoporous silica nanoparticle (MSN)** is a more recent and more sophisticated form. “Mesoporous” means it is riddled with pores in the size range of roughly 2 to 50 nanometres, arranged in a regular, ordered array running through the particle (Figure 5.1). The consequence is an extraordinarily large internal surface area, on the order of several hundred to around a thousand square metres per gram, all of it available to adsorb and hold drug. This gives MSNs a high drug-loading capacity, and because the pore size is tunable during synthesis, the carrier can be matched to the size of the drug molecule.

Mesoporous silica nanoparticle (MSN) with gated pores

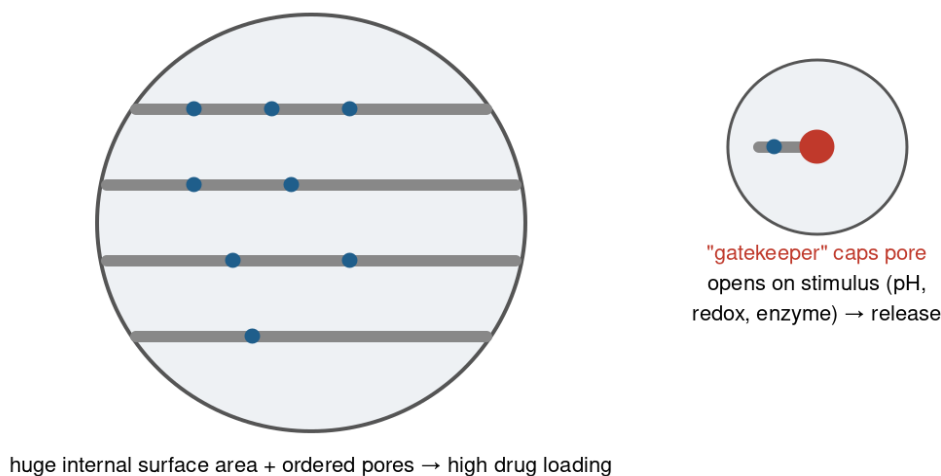


Figure 5.1. A mesoporous silica nanoparticle. Its ordered network of nanometre-scale pores gives a very large internal surface area for high drug loading. A “gatekeeper” molecule can cap each pore mouth and be designed to open only on a specific stimulus, such as a fall in pH, a change in redox state, or the presence of an enzyme, giving stimulus-triggered release.

The feature that makes MSNs genuinely clever is the **gatekeeper** concept. Because the drug sits inside well-defined pores, the pore mouths can be capped with molecules or nanoparticles that act as gates. The gate is designed to stay shut under normal conditions and to open only in response to a chosen trigger: the lower pH of a tumour or endosome, the reducing environment inside a cell, a specific enzyme, or an external signal such as light. Until the trigger arrives, the drug is locked in; when it arrives, the gate opens and release follows. This turns a simple high-capacity sponge into a stimulus-responsive system of the kind described in Chapter 2. Silica is also relatively easy to surface-modify, so MSNs can be PEGylated for circulation and decorated with targeting ligands. Their main unresolved issue remains long-term biodegradation and clearance, though silica does slowly dissolve to silicic acid, which is excretable, and this is an area of active study.

5.3 Gold nanoparticles

Gold is chemically inert and well tolerated, and gold nanoparticles have a distinctive optical property that drives most of their delivery interest. When light of the right wavelength strikes a gold nanoparticle, the free conduction electrons at its surface oscillate collectively in resonance with the light. This phenomenon is **surface plasmon resonance (SPR)**, and the wavelength at which it occurs depends sharply on the particle’s size and, especially, its shape (Figure 5.2).

Gold nanoparticle geometries and surface plasmon resonance

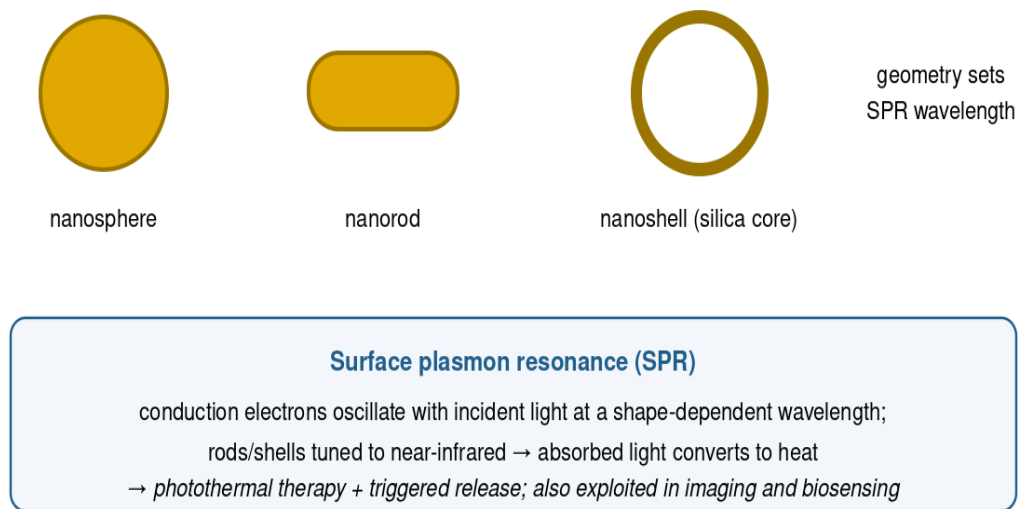


Figure 5.2. Gold nanoparticle geometries and surface plasmon resonance. The resonance wavelength shifts with shape: spheres resonate in the visible, while rods and silica-cored nanoshells can be tuned to the near-infrared. Light absorbed at resonance is converted to heat, which is exploited for photothermal therapy and for triggering drug release, as well as for imaging and biosensing.

The practical payoff comes from tuning the resonance into the **near-infrared**, the wavelength region where biological tissue is relatively transparent so that light penetrates deepest. Gold nanorods and silica-cored gold nanoshells can be engineered to absorb near-infrared light, which they then convert efficiently into heat. Two applications follow. In **photothermal therapy**, gold particles concentrated in a tumour are illuminated, and the local heating destroys the tumour tissue. In **triggered release**, the same heating can be used to release a drug from a heat-sensitive carrier coating the gold, concentrating release where the light is applied. Gold particles are also strong contrast and scattering agents, which makes them useful in imaging and in biosensors. As with silica, the open question for systemic use is the fate of the non-degradable gold core after the therapy is done.

5.4 Superparamagnetic iron oxide nanoparticles

Iron oxide nanoparticles occupy a special place among inorganic carriers because, in iron, the body already has well-developed machinery for handling the element, which gives iron oxide a clearer safety and clearance story than gold or silica. Small enough iron oxide particles are **superparamagnetic**: they become strongly magnetised in an applied magnetic field but retain no magnetism once the field is removed, so they do not clump together magnetically when stored. This single property underlies two distinct uses.

The first is **magnetic targeting and guidance**. Because the particles respond to a magnetic field, an external magnet can in principle concentrate a magnetic drug carrier at a chosen site, or a field can be used to heat the particles (magnetic hyperthermia) for therapy or triggered release. The second, and clinically more established, is as a **magnetic-resonance imaging**

(MRI) contrast agent: superparamagnetic iron oxide alters the local magnetic environment and changes the MRI signal, improving contrast.

The most important real example is **ferumoxytol**, an iron oxide nanoparticle preparation approved by the USFDA for the treatment of iron-deficiency anaemia, particularly in chronic kidney disease, where it is given intravenously as an iron source.¹ Beyond its labelled use it is widely studied off-label and in research as an MRI contrast agent and as a platform for drug delivery and immunomodulation, precisely because it is an already-approved, clinically familiar nanoparticle.¹ Ferumoxytol is therefore a useful teaching example: an inorganic nanoparticle that is genuinely in routine clinical use, made acceptable largely because the body knows how to metabolise iron.

5.5 Metal–organic frameworks

The **metal–organic framework (MOF)** is the newest of the carriers in this chapter and, from a pure capacity standpoint, the most spectacular. A MOF is a crystalline solid built from two kinds of component: **metal ions or metal-ion clusters** that act as connecting nodes, and **organic molecules** (typically rigid, multi-armed carboxylates or similar) that act as linkers bridging the nodes (Figure 5.3). The nodes and linkers self-assemble into a regular, repeating, cage-like lattice that is extraordinarily porous, far more so than mesoporous silica. The internal surface areas of MOFs are the highest of any known material class, which in principle allows very high drug payloads to be hosted in the pores.

Metal–organic framework: metal nodes + organic linkers

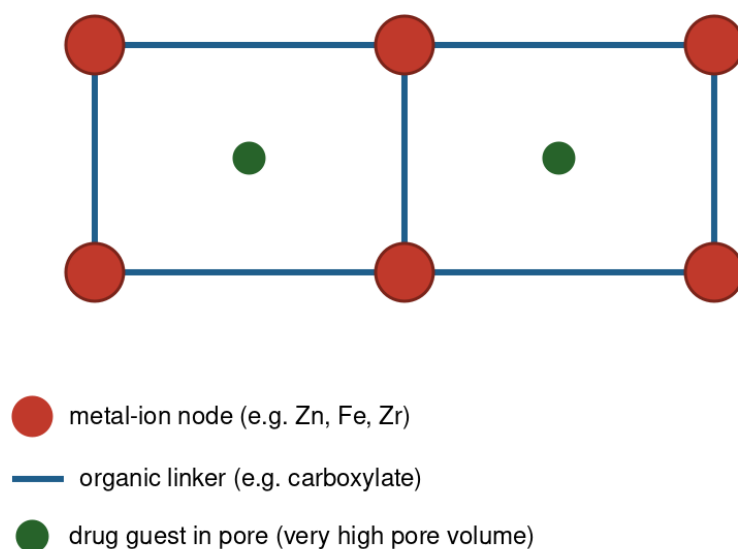


Figure 5.3. The architecture of a metal–organic framework. Metal-ion nodes are bridged by organic linker molecules into a regular, highly porous lattice. The resulting cages can host a large weight of drug as guest molecules. Choosing different metals and linkers tunes pore size, surface chemistry and degradability.

The appeal for delivery is twofold: the immense capacity, and the chemical tunability, since changing the metal or the linker changes pore size, surface chemistry, and stability. Some MOFs built from biologically tolerated metals such as iron and from biocompatible linkers

can degrade in physiological conditions, which addresses the clearance problem better than silica or gold. The honest position is that MOFs for drug delivery remain largely at the research stage. The questions of stability in biological fluids, the toxicity of the chosen metal and linker, reproducible manufacture, and long-term fate are still being worked out, and no MOF drug-delivery product is in routine clinical use. They are included here because the syllabus requires them and because they represent the direction the field is exploring, not because they are established practice.

Table 5.1 compares the four inorganic carriers, and Table 5.2 lists clinically relevant examples.

Table 5.1. Comparison of the principal inorganic drug-delivery carriers.

Carrier	Defining property	Main delivery use	Biodegradability / clearance
Mesoporous silica (MSN)	Ordered nanopores, huge surface area	High loading; gated stimulus-triggered release	Slow dissolution to silicic acid; under study
Gold nanoparticle	Surface plasmon resonance (light→heat)	Photothermal therapy; triggered release; imaging	Non-degradable core; clearance a concern
Iron oxide (SPION)	Superparamagnetism	Magnetic targeting/hyperthermia; MRI contrast	Iron metabolised by normal pathways
Metal–organic framework (MOF)	Highest known porosity	Ultra-high-capacity host (research stage)	Tunable; some biodegradable variants

Table 5.2. Clinically relevant inorganic nanoparticles.

Example	Type	Status
Ferumoxytol	Iron oxide nanoparticle	USFDA-approved for iron-deficiency anaemia; studied as MRI agent and delivery platform
Iron oxide MRI contrast agents	Iron oxide nanoparticle	Clinically used as contrast media
Mesoporous silica carriers	Silica	Investigational / early clinical
Gold nanoshells, nanorods	Gold	Investigational (photothermal)

Quick-recall box — inorganic carriers by their physics Silica → ordered pores + gatekeeper = high loading, triggered release. Gold → SPR = light into heat = photothermal therapy and triggered release. Iron oxide → superparamagnetism = magnetic guidance + MRI contrast (ferumoxytol is approved). MOF → highest porosity = huge capacity, still mostly research. The shared caveat: unlike polymers and lipids, most do not cleanly biodegrade, so clearance is the translation bottleneck, with iron oxide the clearest exception.

Chapter summary

Inorganic carriers trade the clean biodegradability of polymers and lipids for physical properties no organic material can supply, and the recurring caveat across the chapter is the uncertain long-term clearance of a non-degradable solid. Mesoporous silica nanoparticles offer an ordered nanopore network of vast internal surface area for high drug loading, and their pore mouths can be fitted with gatekeeper caps that open only on a chosen stimulus, turning the particle into a stimulus-responsive system. Gold nanoparticles display surface plasmon resonance, a shape-dependent collective oscillation of surface electrons that, when tuned to the near-infrared, converts light efficiently into heat, enabling photothermal therapy and light-triggered release as well as imaging. Superparamagnetic iron oxide nanoparticles respond to magnetic fields without retaining magnetism, allowing magnetic targeting and hyperthermia while serving as MRI contrast agents; ferumoxytol is the standout clinically approved example, made acceptable because the body already handles iron. Metal–organic frameworks, assembled from metal-ion nodes and organic linkers, have the highest porosity of any material class and therefore the highest potential capacity, but remain largely at the research stage with stability, toxicity and clearance still to be resolved. Across all four, the lesson is that the inorganic solid state buys remarkable function at the price of a clearance question that determines how far each can travel toward the clinic.

Self-assessment

Short-answer (2 marks each)

1. State the principal trade-off in using an inorganic carrier instead of a polymer or lipid.
2. What does “mesoporous” mean, and why does it give silica a high drug-loading capacity?
3. Define surface plasmon resonance and state why near-infrared tuning is useful.
4. What does “superparamagnetic” mean, and name one approved iron oxide nanoparticle.
5. Name the two structural components of a metal–organic framework.

Long-answer (5–10 marks each)

6. Describe the structure of a mesoporous silica nanoparticle and explain the gatekeeper concept for triggered release, with a diagram.
7. Explain how the optical properties of gold nanoparticles are exploited for photothermal therapy and triggered drug release.
8. Compare the four inorganic carriers in this chapter on their defining physical property, delivery application, and clearance, and comment on their relative clinical maturity.

Multiple-choice

9. The gatekeeper concept is most associated with: (a) gold nanorods (b) mesoporous silica (c) ferumoxytol (d) liposomes
10. Ferumoxytol is approved primarily for: (a) photothermal cancer therapy (b) iron-deficiency anaemia (c) gene delivery (d) vaccination
11. Metal–organic frameworks are notable for having the highest: (a) magnetism (b) optical absorption (c) porosity/surface area (d) biodegradation rate

Answer pointers. 1. Useful inorganic physics (porosity, optics, magnetism) versus loss of clean biodegradability and uncertain clearance. 2. Ordered ~2–50 nm pores giving very large internal surface area to adsorb drug. 3. Collective oscillation of surface conduction electrons

with light; NIR penetrates tissue best and heats. 4. Magnetised only while field applied, no residual magnetism; ferumoxytol. 5. Metal-ion nodes + organic linkers. 6. Figure 5.1 reproduced with gatekeeper explanation. 7. SPR → NIR absorption → heat → tumour ablation / heat-triggered release. 8. Per Table 5.1, with maturity comment (iron oxide clinical, others earlier). 9. (b). 10. (b). 11. (c).

References

1. Toth GB, Varallyay CG, Horvath A, et al. Current and potential imaging applications of ferumoxytol for magnetic resonance imaging. *Kidney Int.* 2017;92(1):47–66.
2. Vallet-Regí M, Colilla M, Izquierdo-Barba I, Manzano M. Mesoporous silica nanoparticles for drug delivery: current insights. *Molecules.* 2018;23(1):47.
3. Dykman L, Khlebtsov N. Gold nanoparticles in biomedical applications. *Chem Soc Rev.* 2012;41(6):2256–82.
4. Wu MX, Yang YW. Metal–organic framework (MOF)-based drug/cargo delivery and cancer therapy. *Adv Mater.* 2017;29(23):1606134.
5. Torchilin VP, editor. *Nanoparticulates as Drug Carriers.* London: Imperial College Press; 2006.

Chapter 6 - Excipients in NDDS: GRAS, the Inactive Ingredient Database, and Regulatory Constraints

Learning outcomes

After working through this chapter the student will be able to:

1. Explain why the word “inactive” is misleading for excipients in novel delivery systems, and classify excipients by the function they perform.
2. Define GRAS status and explain what it does and does not mean for an excipient.
3. Describe the FDA Inactive Ingredient Database (the database the syllabus calls the “IIG”), how a formulator uses it, and the significance of precedent of use and maximum potency.
4. Explain why a novel excipient faces a much harder regulatory path than an established one, and outline the route to acceptance, including the role of regulatory programmes intended to ease it.
5. Apply these concepts to a formulation decision, choosing between an established and a novel excipient and justifying the regulatory consequences.

6.1 Excipients are not passengers

In a conventional tablet the excipients, the fillers, binders, disintegrants and lubricants, support a drug that does the therapeutic work largely on its own. Calling them “inactive ingredients” is reasonable there. In a novel delivery system that description breaks down. The PLGA that sets a depot’s release duration, the ionisable lipid that lets an mRNA vaccine enter a cell, the PEG chain that keeps a liposome circulating, the chitosan that opens a tight junction, none of these is a passive bystander. They are the components that do the actual engineering. The drug may be unchanged from its conventional form; the excipients are what make the delivery novel.

This reframing has a practical consequence that runs through the whole chapter. Because the excipient in an NDDS performs a function on which efficacy and safety depend, regulators scrutinise it accordingly, and the choice of excipient is one of the most consequential, and most constrained, decisions a formulator makes. Figure 6.1 maps the functional categories an excipient can occupy in a novel system; Table 6.1 gives representative examples.

Functional categories of excipients in NDDS

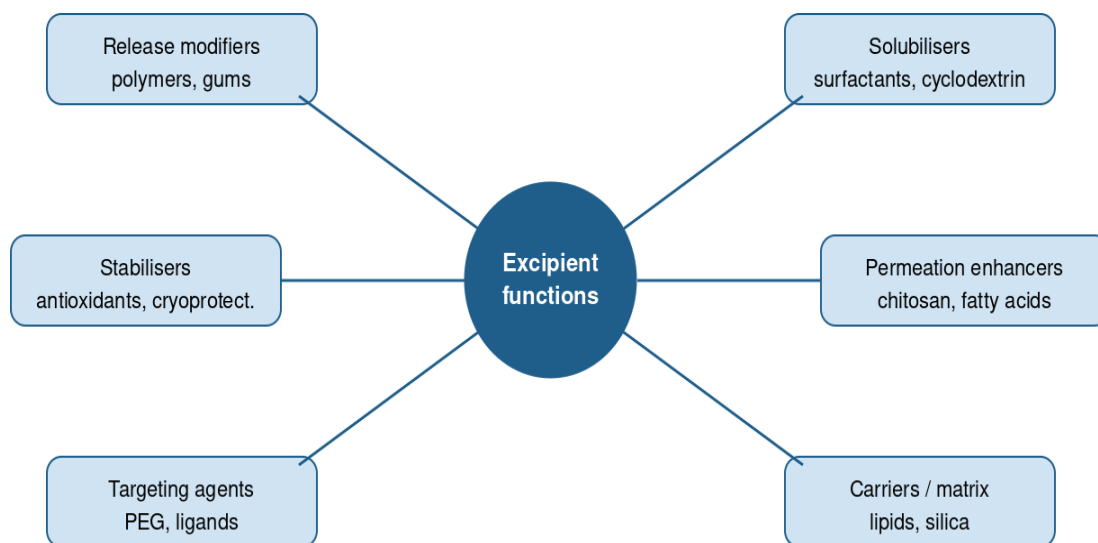


Figure 6.1. Functional categories of excipients in novel delivery systems. Far from being inert, excipients act as release modifiers, solubilisers, stabilisers, permeation enhancers, targeting agents and structural carriers. In an NDDS the excipient often performs the central engineering function.

Table 6.1. Excipient functions and representative examples in NDDS.

Function	What it does	Representative examples
Carrier / matrix	Forms the body of the delivery system	PLGA, lipids, mesoporous silica
Release modifier	Controls rate or timing of release	HPMC, ethylcellulose, PLGA, ion-exchange resins
Solubiliser	Raises apparent solubility of poorly soluble drug	Polysorbates, cyclodextrins, poloxamers
Stabiliser	Protects drug/carrier physically or chemically	Antioxidants, cholesterol, cryo-/lyoprotectants (sugars)
Permeation enhancer	Improves transport across a barrier	Chitosan, fatty acids, bile salts
Targeting / stealth agent	Directs or prolongs circulation	PEG, surface ligands (folate, antibodies)

6.2 GRAS: generally recognised as safe

GRAS stands for **Generally Recognised As Safe**. It is a regulatory status, originating in US food law, applied to substances that qualified experts generally agree are safe for their intended use, based either on a long history of safe use before a defined date or on published scientific evidence. A substance can reach GRAS status either through the regulator's own determination or, in the US system, through a self-determination that the FDA may be notified of and may review.

GRAS status is genuinely useful to a formulator because a substance with a recognised safety record carries far less regulatory burden than an untested material. But two cautions must be understood, and they are commonly tested. First, GRAS is **use- and route-specific**. A substance recognised as safe as a food ingredient taken orally is not automatically safe, or GRAS, when injected, inhaled, or applied to the eye. The safety judgement is tied to a particular use at a particular level by a particular route; change the route and the judgement does not transfer. Second, GRAS is a food-law concept, and while it is a strong supporting argument, it is not by itself a guarantee of acceptability in an injectable or other non-oral drug product. For parenteral and ophthalmic routes especially, the relevant evidence is the substance's history of safe use by that route, not its food status.

6.3 The Inactive Ingredient Database (the syllabus “IIG”)

The single most practical regulatory tool for an excipient decision is the FDA's **Inactive Ingredient Database (IID)**. (The PCI syllabus and many Indian texts refer to it as the “IIG”, from the older name Inactive Ingredient Guide; the current official name is the Inactive Ingredient Database, and the two terms refer to the same resource.)

The database lists every inactive ingredient that appears in an FDA-approved drug product, organised by the **route of administration** and **dosage form** in which it has been used, together with the **maximum potency** or maximum amount of that ingredient per dose unit that has appeared in an approved product.¹ An excipient is added to the database only once it has been used in a product the FDA has approved; the database is built from approved labelling and is updated quarterly.^{1,2}

Its value to a formulator lies in the concept of **precedent of use**. If a developer wants to use an excipient in a new product, and the IID shows that the same excipient has already been used by the same route, in the same kind of dosage form, at or below the amount the developer intends, then a precedent exists. The excipient is, in effect, pre-vetted for that use, and the developer does not have to generate a fresh safety package to justify it; this can be decisive for the speed and cost of development. Conversely, if the developer wants to use more of the excipient than any approved product has used, or to use it by a route for which there is no precedent, then the part above precedent must be justified with new data, as if it were partly novel. The IID is therefore consulted early and constantly: a formulator checks whether a planned excipient, route and level already have precedent before committing to them.

Two limits should be noted. The IID is a US-specific resource; an Indian developer working to CDSCO and the Indian Pharmacopoeia uses it as a strong reference and supporting argument, but the governing standard for an Indian product is set by the IP and CDSCO requirements. And precedent eases the path but does not remove the developer's responsibility to show that the excipient is safe and functions correctly in the specific new formulation.

6.4 The novel excipient problem

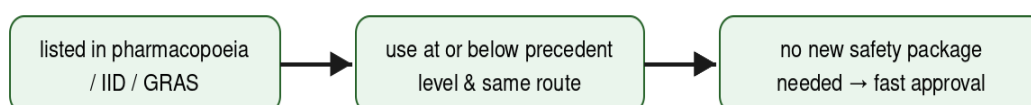
A recurring obstacle in novel delivery is that genuinely new delivery functions often require genuinely new excipients, materials with no precedent of use. Here the regulatory structure works against innovation, and understanding why is important.

The core fact is that **an excipient has no independent regulatory approval of its own**. Unlike a drug, an excipient is not approved as a standalone entity; it is reviewed and accepted

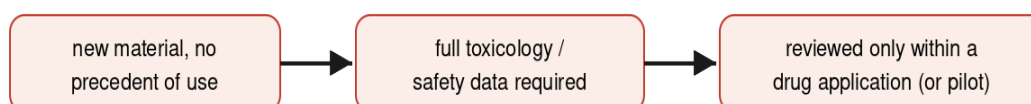
only as a component of a specific drug product application (Figure 6.2). This creates a chicken-and-egg problem. To use a novel excipient, a developer must include it in a drug application and supply the full safety and toxicology data needed to justify a never-before-used material. That data package is expensive and slow to generate, and it loads the entire risk of the new excipient onto a single drug programme. Most developers, understandably, prefer to reach for an established excipient with IID precedent rather than carry the cost and risk of qualifying a new one. The result is conservatism: novel materials that could enable better delivery are passed over because the path to acceptance is so onerous.

How an excipient reaches an approved product

Established excipient (easy route)



Novel excipient (hard route)



Excipients have no independent approval; they are cleared as part of a drug product.

Figure 6.2. The contrasting routes for established and novel excipients. An excipient with pharmacopoeial, IID or GRAS precedent, used at or below the precedented level by the same route, needs no new safety package and clears quickly. A novel excipient with no precedent must carry full toxicology data and is reviewed only inside a drug product application (or through a dedicated pilot programme), a far slower and costlier path.

Recognising that this conservatism holds back useful innovation, regulators have begun to create routes to assess promising novel excipients somewhat ahead of a specific drug application. The USFDA established a **Novel Excipient Review Pilot Program** to allow excipient manufacturers to submit certain novel excipients for FDA review of safety and toxicology data independently of a drug application, so that a successfully reviewed excipient is more attractive for later use in products. The existence of such a programme is itself the lesson worth retaining: the regulatory system has identified the novel-excipient bottleneck as a real barrier to delivery innovation and is experimenting with ways to ease it. Table 6.2 summarises the three regulatory categories a formulator weighs.

Table 6.2. Regulatory standing of excipient categories and what each implies.

Category	What it means	Implication for a new product
Pharmacopoeial (IP/USP/etc.)	Compendial monograph defines identity and quality	Quality standard available; still must suit the route/use
IID precedent	Used in an approved product by a given route/form/level	Use at or below precedent → no new safety package

GRAS	Expert consensus on safety for a defined use	Strong support; route- and use-specific, food-origin
Novel (no precedent)	New material, never used in an approved product	Full toxicology data; reviewed within a drug application or pilot

6.5 Putting the constraints to work

The chapter's practical message is that excipient selection in an NDDS is a negotiation between what the delivery science needs and what the regulatory system will accept without an enormous data burden. A skilled formulator reaches first for materials that combine the required function with strong precedent: a pharmacopoeial polymer used at an IID-precedented level by the intended route, a GRAS surfactant with a history of parenteral use, a PEG grade already present in approved injectables. Only when no precedent material can deliver the needed function does the formulator take on the cost of a novel excipient, and then with eyes open to the toxicology programme that will follow.

This is also why the established materials of the previous chapters, PLGA, lecithin, the polysorbates, cholesterol, dominate marketed products: they are not merely effective, they are accepted. A novel system built entirely on established, precedent excipients can reach patients far faster than an equally clever one that depends on a material no regulator has seen before. The science of delivery and the regulation of excipients are, in the end, the same conversation viewed from two sides.

Quick-recall box — the excipient decision Ask three questions of any planned excipient. Is there a pharmacopoeial monograph for its quality? Does the IID (the syllabus "IIG") show precedent for this route, form and level? Is its safety supported by GRAS status or a history of use by this route? If all three are yes, the path is short. If the material is novel, expect a full toxicology package, because an excipient is only ever approved as part of a drug product, not on its own.

Chapter summary

In a novel delivery system the excipient is rarely inactive; it is frequently the component doing the central engineering, which is why regulators scrutinise it and why excipient choice is one of the most constrained decisions in formulation. Excipients can be grouped by function into carriers, release modifiers, solubilisers, stabilisers, permeation enhancers and targeting agents. GRAS status records an expert consensus that a substance is safe for a defined use, but it is route- and use-specific and rooted in food law, so it does not transfer automatically to injection or inhalation. The most practical tool is the FDA Inactive Ingredient Database, the resource the syllabus calls the IIG, which lists each excipient by the route, dosage form and maximum level at which it has appeared in an approved product; its central idea is precedent of use, since an excipient used at or below precedent by the same route needs no fresh safety package. The hard case is the novel excipient, which has no precedent and, because an excipient is never approved independently but only as part of a drug application, must carry a full and costly toxicology package, a structure that pushes developers toward established materials and slows the adoption of new ones. Regulators have responded with programmes such as the FDA novel-excipient review pilot to ease this bottleneck. The working conclusion is that the best formulator matches the delivery science to materials that already carry regulatory precedent, reserving novel excipients for the rare case where nothing precedent will do.

Self-assessment

Short-answer (2 marks each)

1. Why is the term “inactive ingredient” misleading in the context of NDDS?
2. Define GRAS and state one important limitation of relying on it for a parenteral product.
3. What information does the FDA Inactive Ingredient Database provide about each excipient?
4. Explain “precedent of use” and why it matters to a developer.
5. Why does a novel excipient face a harder regulatory path than an established one?

Long-answer (5–10 marks each)

6. Classify excipients used in novel delivery systems by function, giving an example of each and explaining its role.
7. Describe the FDA Inactive Ingredient Database, the meaning of maximum potency and precedent of use, and how a formulator uses it when selecting an excipient, noting its relationship to Indian regulatory practice.
8. Explain the novel-excipient problem: why excipients lack independent approval, the consequences for delivery innovation, and the regulatory response.

Multiple-choice

9. The IID lists an excipient by all of the following except: (a) route of administration (b) dosage form (c) maximum potency used (d) the patent holder
10. GRAS status is: (a) universal across all routes (b) specific to a defined use and route (c) a guarantee of injectable safety (d) independent of evidence
11. An excipient receives regulatory approval: (a) as a standalone entity (b) only as part of a drug product application (c) automatically once GRAS (d) from the pharmacopoeia alone

Answer pointers. 1. The excipient often performs the central delivery function (release, uptake, circulation), not a passive role. 2. Expert consensus on safety for a defined use; it is food-derived and route-specific, so oral GRAS does not justify injection. 3. Route, dosage form, and maximum amount per dose unit used in an approved product. 4. Prior approved use at a given route/form/level; using at or below it avoids a new safety package. 5. No precedent → full toxicology data, and excipients are only approved within a drug application. 6. Per Table 6.1 and Figure 6.1. 7. Per Section 6.3, including IP/CDSCO note. 8. Per Section 6.4 and Figure 6.2. 9. (d). 10. (b). 11. (b).

References

1. US Food and Drug Administration. Inactive Ingredient Database (IID) — overview and download. Silver Spring (MD): FDA; updated quarterly.
2. US Food and Drug Administration. Using the Inactive Ingredient Database: Guidance for Industry. Silver Spring (MD): FDA; 2021.
3. US Food and Drug Administration. Novel Excipient Review Pilot Program. Federal Register notice and programme description. Silver Spring (MD): FDA; 2021.
4. Indian Pharmacopoeia Commission. Indian Pharmacopoeia. Ghaziabad: IPC; current edition.
5. Aulton ME, Taylor KMG, editors. Aulton’s *Pharmaceutics: The Design and Manufacture of Medicines*. 6th ed. Edinburgh: Elsevier; 2022.

Unit II - Oral and Mucosal Delivery Systems

Unit overview. The oral route is the one patients prefer and the one industry returns to whenever it can, yet Chapter 1 showed how many ways it can fail. This unit takes the oral and mucosal routes and shows how delivery design rescues them. The first problem is timing and position: a dose moves through the gut on a schedule the formulator does not control, so Chapter 7 keeps a dose in the stomach when it needs to stay there, and Chapter 8 holds a dose intact until it reaches the colon. The second problem is that some drugs cannot use the gut at all, so Chapter 9 turns to the buccal, nasal and pulmonary surfaces, mucosal routes that bypass the gut and, often, the liver. Chapter 10 then asks the question that decides whether any of these systems can be developed efficiently: how well does a dissolution test in a beaker predict what the body will do, and how is a meaningful in-vitro–in-vivo correlation built. Throughout, the governing facts are the pH and the transit times of the gastrointestinal tract, so those numbers anchor the unit.

Chapter 7 - Gastro-retentive Drug Delivery Systems

Learning outcomes

After working through this chapter the student will be able to:

1. State the gastrointestinal pH and transit-time values that make gastric retention both necessary and difficult, and identify which drugs benefit from it.
2. Define the absorption-window problem and explain why a conventional sustained-release form can fail for a narrow-window drug.
3. Describe the four principal gastro-retentive strategies, floating, bioadhesive, expandable and high-density, and the mechanism of each.
4. Distinguish effervescent from non-effervescent (hydrodynamically balanced) floating systems and explain how each achieves buoyancy.
5. Evaluate the limitations of gastro-retentive systems, including the strong influence of fed and fasted states.

7.1 The problem: the gut does not wait

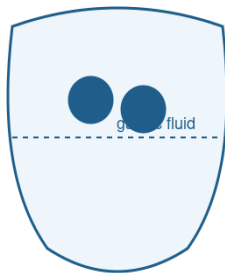
A dose taken by mouth is carried through the gastrointestinal tract by motility the patient cannot control, and the time it spends in each region is both short and variable. In the fasted state the stomach empties on a median of roughly half an hour, the small intestine passes its contents in something like three to four hours, and only the colon offers a long residence, well over a day.¹ The pH climbs along the way, from a fasted gastric pH near 1.7, through the proximal small intestine around pH 6, to the distal small intestine approaching pH 7.4 or higher.¹ These numbers, used here and again in Chapters 8 and 10, are the physiological constants every oral delivery system must design around.

The short, variable transit time creates a specific difficulty. Many drugs are not absorbed uniformly along the whole gut; they have an **absorption window**, a region, usually the upper small intestine, where absorption is efficient, with little or no absorption beyond it. For such a drug, a conventional sustained-release tablet is self-defeating: it is designed to release slowly over many hours, but it is swept past its absorption window in three or four, so the drug released after that point lands where it cannot be absorbed and is wasted. The plasma levels fall short and become erratic. The solution is not to release faster but to slow the dosage form's journey, holding it in the stomach so that drug is released steadily above its absorption window for as long as the dose lasts. That is the purpose of a **gastro-retentive drug delivery system (GRDDS)**.

Three kinds of drug benefit. Drugs with a narrow upper-GI absorption window, such as several antivirals, some antibiotics, levodopa and certain vitamins, gain because release is kept above the window. Drugs that act locally in the stomach, such as antibiotics against *Helicobacter pylori* or antacids, gain because the drug is held where it is needed. And drugs that are unstable or poorly soluble at the higher pH of the intestine gain because they dissolve and are delivered in the acidic stomach. Figure 7.1 shows the four engineering strategies used to achieve retention.

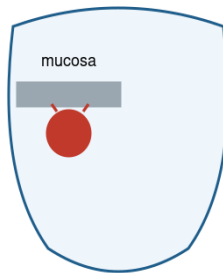
Strategies for gastric retention

1. Floating (low density)



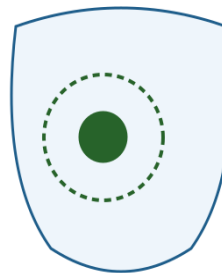
density < 1 → floats on fluid

2. Bioadhesive



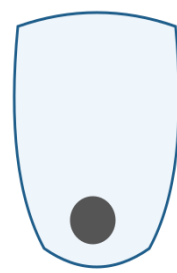
adheres to stomach lining

3. Expandable / swelling



swells > pylorus → cannot pass

4. High density



density > 2.5 → sinks to antrum

Shared aim: prolong residence in the stomach

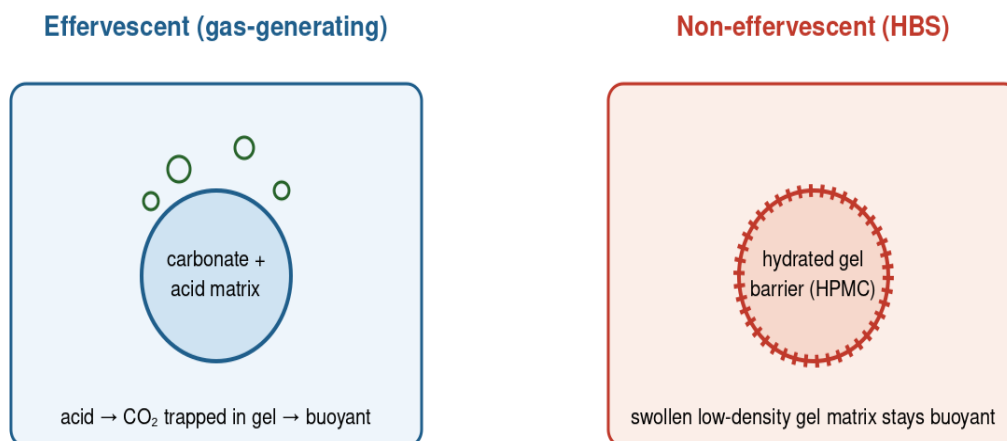
for drugs with a narrow upper-GI absorption window, local gastric action, or acid solubility

Figure 7.1. The four principal gastro-retentive strategies. A floating system has a bulk density below that of gastric fluid and rides on its surface; a bioadhesive system sticks to the gastric mucosa; an expandable system swells beyond the diameter of the pylorus so it cannot pass; and a high-density system sinks to the antrum and resists being swept out. All share the single aim of prolonging gastric residence.

7.2 Floating systems

Floating systems are the most developed approach. The principle is buoyancy: if the bulk density of the dosage form is kept below that of gastric contents, which is close to 1 g/mL, the form floats on the gastric fluid rather than settling near the pylorus, and so resists being emptied. While it floats, the drug is released slowly, and the spent system eventually empties once it has lost its buoyancy or its drug. Two mechanisms produce the low density (Figure 7.2).

Two mechanisms of floating systems



HBS = hydrodynamically balanced system. Both keep bulk density below that of gastric fluid (~1 g/mL).

Figure 7.2. The two mechanisms of floating. An effervescent system contains a carbonate (and sometimes an acid source) that reacts with gastric acid to generate carbon dioxide, which is trapped in a gel layer and lowers the system's density. A non-effervescent or hydrodynamically balanced system uses a swelling hydrocolloid that hydrates into a low-density gel that stays buoyant. Both keep the bulk density below that of gastric fluid.

Effervescent (gas-generating) systems incorporate a carbonate or bicarbonate, often with an organic acid such as citric or tartaric acid. On contact with gastric acid, or with the built-in acid, the carbonate liberates carbon dioxide. The gas is trapped within a gelled hydrocolloid layer surrounding the dose, swelling it and dropping its density below that of the surrounding fluid, so it rises and floats. The release of drug is controlled by the hydrated gel layer while the trapped gas keeps the system up.

Non-effervescent systems, the classic form being the **hydrodynamically balanced system (HBS)**, achieve buoyancy without gas. They are built from a high proportion of a swellable hydrocolloid, typically a high-viscosity grade of hydroxypropyl methylcellulose (HPMC). On contact with gastric fluid the outer hydrocolloid hydrates into a thick, low-density gel that entraps air and keeps the bulk density of the whole mass below 1 g/mL, so it floats while the gel simultaneously controls drug release by diffusion and erosion. A marketed historical example of the HBS concept is a sustained-release levodopa/benserazide capsule designed to float and release in the stomach, which suits levodopa because of its narrow upper-GI absorption.²

A practical requirement common to floating systems is that they need a minimum volume of gastric fluid to float in. This is why such products are typically taken with a glass of water, and why their performance can be poor if the stomach is nearly empty.

7.3 Bioadhesive (mucoadhesive) systems

A bioadhesive system retains itself by sticking to the mucus layer lining the stomach rather than by floating. It is formulated with a mucoadhesive polymer that binds to the mucus through hydrogen bonding, electrostatic attraction and physical interpenetration of polymer

chains into the mucus gel. Polymers used include the polyacrylic acids (carbomers), chitosan (whose cationic mucoadhesion was described in Chapter 3), and various cellulose derivatives. The appeal is intimate, prolonged contact with the mucosa. The difficulty, and the reason gastric bioadhesion has been less successful in practice than it sounds in theory, is that the gastric mucus layer turns over continuously: the mucus the system adheres to is itself shed and replaced, carrying the dosage form away with it, and the dilution and movement of gastric contents further weaken adhesion. Bioadhesion is therefore generally more reliable at quieter mucosal sites, such as the buccal cavity (Chapter 9), than in the churning stomach.

7.4 Expandable and high-density systems

Expandable (swelling or unfolding) systems work by mechanical size. The dosage form is small enough to be swallowed but, once in the stomach, swells by absorbing fluid or unfolds from a folded configuration to a size larger than the diameter of the pylorus, the narrow opening to the small intestine. Too large to pass, it is retained until it eventually degrades or breaks down into pieces small enough to leave. The engineering challenge is real: the form must stay compact enough to swallow, expand quickly enough to avoid being emptied before it grows, remain robust during the retention period, and finally break down safely so it cannot cause an obstruction. This last safety requirement, that the system must reliably disintegrate before it could block the gut, is the dominant design constraint.

High-density systems take the opposite approach to floating. Instead of rising above the gastric fluid, they are made dense enough, with a density around 2.5 g/mL or more achieved with heavy excipients such as barium sulfate or iron powder, to sink to the lowest part of the stomach, the antrum, where they resist being caught by the peristaltic waves that empty the stomach. The approach is conceptually simple but has seen limited commercial use, partly because achieving and maintaining the required density in a swallowable unit is difficult.

7.5 Limitations and the fed–fasted problem

Gastro-retentive systems are attractive in principle but humbling in practice, and a balanced account must state their limitations plainly. The single largest is the strong dependence on the **fed or fasted state** of the stomach. Gastric emptying is governed in the fasted state by a periodic, powerful contraction known as the migrating motor complex, sometimes called the “housekeeper wave”, which sweeps the stomach clear of residual material roughly every couple of hours. A floating or bioadhesive system caught by this wave can be emptied regardless of its design. In the fed state the picture changes: a meal, especially a fatty or substantial one, delays gastric emptying and raises gastric pH, which generally prolongs retention and improves the performance of these systems. The consequence is that many gastro-retentive products work well only when taken with or after food, and their behaviour can vary considerably between patients and between doses. Floating systems additionally need sufficient gastric fluid, and high-density and expandable systems carry their own difficulties of manufacture and safety. None of this negates their value for the right drug, but it explains why gastric retention is a niche, carefully chosen strategy rather than a routine one.

Quick-recall box — gastric retention essentials Retain the dose when the drug has a narrow upper-GI absorption window, acts locally in the stomach, or is acid-soluble. Four strategies: floating (density < ~1 g/mL), bioadhesive (sticks to mucus), expandable (swells beyond the pylorus), high-density (sinks to the antrum). Floating splits into effervescent (CO₂ from carbonate) and non-

effervescent/HBS (swollen low-density gel). The biggest real-world variable is fed vs fasted state; most systems favour the fed stomach.

Chapter summary

Gastric retention exists because the gut moves a dose along faster than some drugs can be absorbed. With fasted gastric emptying on the order of half an hour and small-intestinal transit only a few hours, a drug with a narrow upper-GI absorption window is swept past the region where it can be absorbed long before a conventional sustained-release tablet has finished releasing, wasting the later fraction of the dose. Holding the dosage form in the stomach solves this and also suits drugs that act locally in the stomach or are acid-soluble. Four strategies achieve retention: floating systems ride on the gastric fluid by keeping their density below it, either by generating trapped carbon dioxide (effervescent) or by swelling into a low-density hydrocolloid gel (the non-effervescent hydrodynamically balanced system); bioadhesive systems stick to the gastric mucus, though continuous mucus turnover limits their reliability; expandable systems swell or unfold beyond the diameter of the pylorus and must be engineered to break down safely afterwards; and high-density systems sink to the antrum to avoid the emptying waves. The decisive practical limitation across all of them is the dependence on the fed or fasted state, since the fasted migrating motor complex can empty even a well-designed system, while a meal prolongs retention. Gastric retention is therefore a deliberately chosen strategy for particular drugs, not a general-purpose tool.

Self-assessment

Short-answer (2 marks each)

1. State approximate fasted gastric emptying time and small-intestinal transit time, and explain why they matter to oral delivery.
2. What is an absorption window, and why does it defeat a conventional sustained-release tablet?
3. How does an effervescent floating system achieve buoyancy?
4. What is a hydrodynamically balanced system?
5. Why is bioadhesion less reliable in the stomach than at the buccal mucosa?

Long-answer (5–10 marks each)

6. Describe the four gastro-retentive strategies with a diagram, naming a suitable drug type for each.
7. Compare effervescent and non-effervescent floating systems in mechanism, composition and limitations.
8. “Gastro-retentive systems are strongly influenced by physiology.” Discuss with reference to the fed–fasted state and the migrating motor complex.

Multiple-choice

9. A floating tablet must have a bulk density: (a) greater than ~ 2.5 g/mL (b) equal to gastric fluid (c) less than gastric fluid (~ 1 g/mL) (d) zero
10. The migrating motor complex is most likely to empty a gastro-retentive system in the: (a) fed state (b) fasted state (c) presence of fatty food (d) presence of water
11. An expandable gastro-retentive system is retained because it: (a) sticks to mucus (b) generates gas (c) swells beyond the pylorus diameter (d) increases density

Answer pointers. 1. ~30 min fasted emptying, ~3–4 h small-intestinal transit; short window for absorption. 2. Region where a drug is absorbed (usually upper small intestine); SR form releases drug after it has passed the window, wasting it. 3. Carbonate + acid generates CO₂ trapped in a gel, lowering density below gastric fluid. 4. Non-effervescent floating system using a swelling hydrocolloid (e.g. HPMC) that hydrates to a low-density buoyant gel. 5. Gastric mucus turns over and churns continuously; buccal mucosa is quieter and more stable. 6. Per Figure 7.1 with drug types. 7. Per Section 7.2. 8. Fasted housekeeper wave empties systems; fed state delays emptying and improves retention. 9. (c). 10. (b). 11. (c).

References

1. Koziolk M, Grimm M, Becker D, et al. Investigation of pH and temperature profiles in the GI tract of fasted human subjects using the Intellicap system. *J Pharm Sci.* 2015;104(9):2855–63.
2. Streubel A, Siepmann J, Bodmeier R. Gastroretentive drug delivery systems. *Expert Opin Drug Deliv.* 2006;3(2):217–33.
3. Pawar VK, Kansal S, Garg G, et al. Gastroretentive dosage forms: a review with special emphasis on floating drug delivery systems. *Drug Deliv.* 2011;18(2):97–110.
4. Aulton ME, Taylor KMG, editors. *Aulton's Pharmaceutics*. 6th ed. Edinburgh: Elsevier; 2022.

Chapter 8 - Colon-Targeted Drug Delivery Systems

Learning outcomes

After working through this chapter the student will be able to:

1. State the therapeutic reasons for targeting drug release to the colon, both for local and for systemic purposes.
2. Use the pH and transit-time profile of the gastrointestinal tract to explain why each colon-targeting trigger works and where each can fail.
3. Describe the pH-triggered approach, including the dissolution thresholds of the relevant methacrylate coatings, and its principal weakness.
4. Explain the microbially triggered approach using azo polymers and polysaccharides, and why it is the most colon-specific strategy.
5. Describe time-dependent systems and justify why combining triggers gives more reliable colonic delivery than any single trigger.

8.1 Why target the colon

There are two quite different reasons to deliver a drug specifically to the colon, and keeping them separate clarifies the whole chapter.

The first is **local treatment of colonic disease**. Conditions such as ulcerative colitis and Crohn's disease (together, inflammatory bowel disease), and colorectal cancer, are diseases of the colon itself. Delivering the drug directly to the colonic mucosa concentrates it where it is needed and limits its absorption higher up, which both improves local efficacy and reduces systemic side effects. A corticosteroid or an aminosalicylate released in the colon treats the inflamed tissue while sparing the rest of the body.

The second is **systemic delivery by exploiting the colon's special features**. The colon has a very long residence time, well over a day, far longer than the few hours available in the small intestine, which makes it attractive for sustained release and for once-daily systems. It also has lower enzymatic activity of the kind that degrades peptides and proteins in the small intestine, which has made the colon a target of interest for the oral delivery of peptide and protein drugs that would otherwise be destroyed higher up. The catch, examined in Chapter 9 and Chapter 10, is that the colon is also a region of relatively low fluid and low absorptive surface area, so systemic colonic absorption is not guaranteed; the colon is a refuge from degradation but not an ideal absorber.

The engineering problem is the same in both cases: the dosage form must travel intact through the stomach and the entire small intestine, releasing nothing, and then release its drug only on reaching the colon. The triggers used to achieve this all read some signal that distinguishes the colon from everything upstream. Three such signals exist, and they map directly onto the GI profile shown in Figure 8.1.

pH and transit along the GI tract — the basis of pH-triggered targeting

Stomach	Small intestine	Colon	transit times
pH ~1.5–3.5	pH ~6 (proximal) → ~7.4 (distal)	pH ~6.4–7 (variable)	stom~0.5h · SI~3–4h · colon>24h

Enteric/colonic coating dissolution thresholds:

- Eudragit L100-55 → dissolves above pH ~5.5 (upper small intestine)
- Eudragit L100 → dissolves above pH ~6.0
- Eudragit S100 → dissolves above pH ~7.0 (distal SI / ileo-colonic) ← colon targeting

A coating that dissolves only above pH 7 holds the dose intact until the distal gut.

Risk: colonic pH is variable and can dip below 7, so pH alone is unreliable → combine triggers.

Figure 8.1. The pH and transit profile of the gastrointestinal tract, with the dissolution thresholds of the common methacrylate (Eudragit) coatings. pH rises from the acidic stomach through the small intestine to around 7.4 distally, then is somewhat lower and variable in the colon. A coating that dissolves only above pH 7 stays intact until the distal gut, but because colonic pH is variable and can fall below 7, pH alone is an unreliable colon-specific signal.

8.2 pH-triggered systems

The pH-triggered approach uses the rise in luminal pH along the gut as its signal. The dosage form is coated with a polymer that is insoluble in acid and in mildly acidic conditions but dissolves above a threshold pH, so the coating survives the stomach and dissolves only once the pH has climbed high enough, which happens in the distal small intestine and the ileo-colonic region.

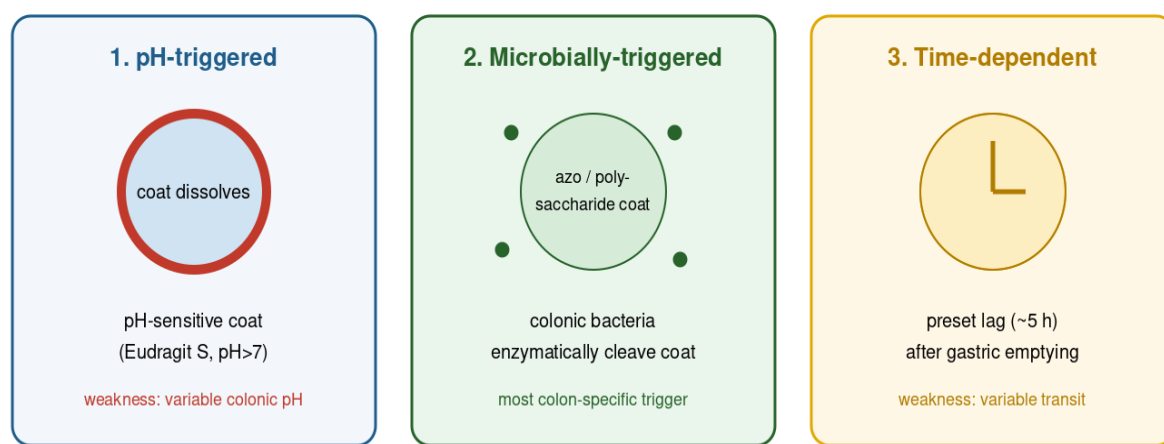
The workhorses are the **methacrylic acid copolymers** sold under the Eudragit name, whose dissolution thresholds are set by their chemistry. **Eudragit L100-55** dissolves above approximately pH 5.5, **Eudragit L100** above approximately pH 6.0, and **Eudragit S100** above approximately pH 7.0.¹ For colon targeting the high-threshold S100 is used, because a coating that requires pH 7 to dissolve will remain intact through most of the small intestine and only open in the distal small intestine and colon. Formulators frequently blend L100 and S100 in chosen ratios to fine-tune the exact pH at which release begins, which lets them place the release point along the distal gut.^{1,2}

The weakness of the pH approach is fundamental and must be stated. Colonic pH is not a fixed, reliable number. It varies between individuals, with diet, and with disease; in inflammatory bowel disease in particular the colonic pH can be substantially lower than in health, sometimes well below 7. A coating that depends on reaching pH 7 may therefore dissolve too early in someone whose distal small-intestinal pH is high, or fail to dissolve at all in a colitis patient whose colonic pH never reaches the threshold, releasing the dose in the wrong place or not releasing it. This unreliability is precisely why pH is rarely used alone for serious colon targeting and is increasingly combined with a second trigger.

8.3 Microbially triggered systems

The most colon-specific signal is biological rather than chemical. The colon hosts a dense population of bacteria, the gut microbiota, that is essentially absent higher up the tract, and these bacteria produce a distinctive set of enzymes, including azoreductases and a range of polysaccharide-degrading enzymes (glycosidases). A delivery system built so that its drug is released only by the action of these colonic enzymes is, by definition, released only in the colon. This is the principle of microbially (or enzymatically) triggered colon targeting, and it is the most genuinely colon-specific of the three approaches because the trigger is unique to the colonic environment (Figure 8.2).

Three triggers for colon-targeted release



Best reliability comes from combining triggers (e.g. pH + microbial).

Figure 8.2. The three triggers for colon-targeted release. A pH-triggered coat dissolves above a threshold pH but is vulnerable to the colon's variable pH. A microbially triggered coat of azo polymer or polysaccharide is cleaved by enzymes that only colonic bacteria produce, making it the most colon-specific. A time-dependent system releases after a fixed lag timed to gastric emptying, but is vulnerable to variable transit. Reliability is greatest when triggers are combined.

Two chemistries dominate. **Azo-polymer systems** use coatings or matrices crosslinked through aromatic azo ($-N=N-$) bonds. These bonds are stable in the upper gut but are reduced and cleaved by the azoreductase enzymes of the colonic bacteria, which breaks down the coating and releases the drug. The same azo principle underlies the classic prodrug **sulfasalazine**, in which 5-aminosalicylic acid (the active anti-inflammatory) is joined to a carrier by an azo bond; colonic bacteria cleave the bond and liberate the active drug in the colon, which is a beautifully economical natural example of microbial triggering and a standard treatment in inflammatory bowel disease.³

Polysaccharide systems use natural polysaccharides, such as pectin, guar gum, chitosan, amylose, inulin and chondroitin sulfate, as coatings or matrices. These polymers resist the host's own digestive enzymes in the stomach and small intestine but are fermented and broken down by the polysaccharide-degrading enzymes of the colonic microbiota, releasing the drug in the colon. Polysaccharides are attractive because they are abundant, inexpensive, biocompatible and biodegradable, though a formulator must guard against premature swelling

or dissolution of these hydrophilic polymers in the upper gut, often by combining them with a pH-resistant component.

8.4 Time-dependent systems

The time-dependent (or time-controlled) approach uses elapsed time as a surrogate for position. The reasoning is that small-intestinal transit time is relatively constant compared with gastric emptying, at roughly three to four hours, so a system that begins a fixed lag period after it leaves the stomach and releases its drug at the end of that lag should, in principle, release in the colon. The typical design combines an enteric coat (so the lag does not start until the stomach is cleared) with a time-controlled barrier that delays release for a programmed interval, often around five hours, after which the drug is released.

The weakness mirrors that of the pH approach: it depends on a physiological value that is in fact variable. Gastric emptying itself is erratic, especially with food, so the moment the lag begins is uncertain; and small-intestinal transit, while less variable than emptying, still differs between people and with disease. A purely time-dependent system can therefore release too early or too late. For this reason time control is most often used as one component of a combined system rather than on its own.

8.5 Combining triggers and marketed approaches

The recurring theme of this chapter is that each single trigger has a real and predictable failure mode: pH is undone by variable colonic pH, time by variable transit, and even the reliable microbial trigger can be compromised if the coating swells prematurely or if the patient's microbiota is altered, for instance by antibiotics or by the dysbiosis of inflammatory bowel disease. The robust engineering answer is to combine triggers so that the strengths of one cover the weaknesses of another. A system that uses a pH-sensitive outer coat to survive the stomach and to begin responding only in the distal gut, combined with a polysaccharide or azo component that requires colonic bacteria to complete release, is far more reliable than either mechanism alone, because release then requires both the right pH region and the right microbial environment.

Commercial and clinical examples of colon targeting are concentrated in inflammatory bowel disease, where local delivery of mesalazine (5-aminosalicylic acid) and of budesonide to the colon is the therapeutic goal; various marketed formulations achieve this with pH-dependent coatings, with multi-matrix systems that delay and prolong release, and with the natural azo prodrug strategy of sulfasalazine. The osmotic approach has also been adapted for the colon, in a colon-targeted osmotic system that delivers drug to the colon by osmotic pumping after an enteric and timed delay.

Quick-recall box — three colon triggers and their failure modes pH (Eudragit S, dissolves above ~7): fails because colonic pH is variable, especially low in colitis. Microbial (azo polymers, polysaccharides cleaved by colonic bacteria): most colon-specific, the basis of sulfasalazine; vulnerable to altered microbiota and premature swelling. Time (enteric + ~5 h lag): fails because gastric emptying and transit vary. Reliable colon targeting combines triggers, classically pH + microbial.

Chapter summary

The colon is targeted either to treat colonic disease locally, concentrating drug at the inflamed mucosa while sparing the rest of the body, or to exploit the colon's long residence and lower

peptidase activity for sustained or peptide delivery, with the caveat that the colon absorbs systemically less readily than the small intestine. In every case the dosage form must pass intact through the stomach and small intestine and release only in the colon, which requires reading a signal unique to the colonic environment. Three triggers do this. pH-triggered systems use methacrylate coatings, chiefly Eudragit S100 dissolving above pH 7, sometimes blended with L100, but are undermined by the colon's variable pH, which is often abnormally low in inflammatory bowel disease. Microbially triggered systems use azo polymers cleaved by bacterial azoreductase, the same chemistry as the sulfasalazine prodrug, or polysaccharides fermented by colonic bacteria, and are the most colon-specific because the trigger exists nowhere else in the gut. Time-dependent systems release after a fixed lag timed to gastric emptying and small-intestinal transit, but suffer from the variability of those very transit times. Because each single trigger has a defined failure mode, the most reliable colon-targeted systems combine triggers, most effectively pH together with a microbial mechanism, so that release requires both the correct region and the correct microbial environment.

Self-assessment

Short-answer (2 marks each)

1. Give one local and one systemic reason for targeting the colon.
2. State the approximate dissolution pH thresholds of Eudragit L100-55, L100 and S100.
3. Why is pH alone an unreliable colon-targeting trigger?
4. Explain how an azo-polymer coating releases drug in the colon.
5. What is the rationale of a time-dependent colon-targeting system, and its main weakness?

Long-answer (5–10 marks each)

6. Using the GI pH and transit profile, compare the three colon-targeting triggers and the failure mode of each, with a diagram.
7. Explain microbially triggered colon delivery using azo polymers and polysaccharides, and discuss why it is the most colon-specific approach, citing sulfasalazine.
8. Justify, with reasoning, why a combined-trigger system is more reliable than any single-trigger system for colon targeting.

Multiple-choice

9. For colon targeting, the most suitable single methacrylate coating is: (a) Eudragit L100-55 (b) Eudragit L100 (c) Eudragit S100 (d) Eudragit E
10. The trigger unique to the colon and therefore most colon-specific is: (a) pH (b) transit time (c) colonic microbial enzymes (d) temperature
11. Sulfasalazine releases 5-ASA in the colon because colonic bacteria: (a) raise the pH (b) cleave its azo bond (c) increase transit time (d) dissolve its coating

Answer pointers. 1. Local: treat IBD/colorectal cancer at the mucosa; systemic: long residence and low peptidase activity (peptide delivery, once-daily). 2. ~5.5, ~6.0, ~7.0. 3. Colonic pH is variable and often low in IBD, so a pH-7 coat may not dissolve or may dissolve too early. 4. Bacterial azoreductase reduces/cleaves the –N=N– bonds, breaking down the coat. 5. SI transit is relatively constant, so a fixed lag after gastric emptying targets the colon; weakness is variable emptying/transit. 6. Per Figure 8.2 with failure modes. 7. Per Section 8.3 with sulfasalazine. 8. Combined triggers require both region and microbial environment, covering each other's failure modes. 9. (c). 10. (c). 11. (b).

References

1. Khan MZ, Prebeg Z, Kurjaković N. A pH-dependent colon-targeted oral drug delivery system using methacrylic acid copolymers. I. Manipulation of drug release using Eudragit L100-55 and Eudragit S100 combinations. *J Control Release*. 1999;58(2):215–22.
2. Ibekwe VC, Fadda HM, McConnell EL, et al. Interplay between intestinal pH, transit time and feed status on the in vivo performance of pH responsive ileo-colonic release systems. *Pharm Res*. 2008;25(8):1828–35.
3. Amidon S, Brown JE, Dave VS. Colon-targeted oral drug delivery systems: design trends and approaches. *AAPS PharmSciTech*. 2015;16(4):731–41.
4. Jain KK, editor. *Drug Delivery Systems*. 3rd ed. New York: Humana Press; 2020.

Chapter 9 - Buccal, Nasal and Pulmonary Delivery

Learning outcomes

After working through this chapter the student will be able to:

1. Explain why the buccal, nasal and pulmonary mucosae are attractive routes, in terms of first-pass avoidance, accessibility and absorptive properties.
2. Describe buccal and sublingual delivery, the dosage forms used, and the role of mucoadhesion.
3. Describe nasal delivery, its advantages including rapid onset and the nose-to-brain possibility, and the constraints of mucociliary clearance and dose volume.
4. Explain pulmonary delivery and the central role of aerodynamic particle size, defining the respirable fraction and how a dry powder inhaler works.
5. Describe how vesicular carriers such as liposomes are used to improve mucosal and pulmonary delivery.

9.1 Why deliver across a mucosa

Chapter 1 established the single strongest argument for these routes: a drug absorbed across the buccal, nasal or pulmonary mucosa enters the systemic veins directly and so escapes the hepatic first pass that destroys many drugs taken orally. For a high-extraction drug, this alone can be decisive. The mucosal routes add further advantages: they are non-invasive and acceptable to patients, they avoid the acid and enzymes of the stomach that degrade fragile drugs, and several of them offer very rapid absorption and onset of action. Figure 9.1 summarises the three routes, their dosage forms, and the trade-offs that distinguish them.

Three mucosal routes that bypass the gut (and first pass)

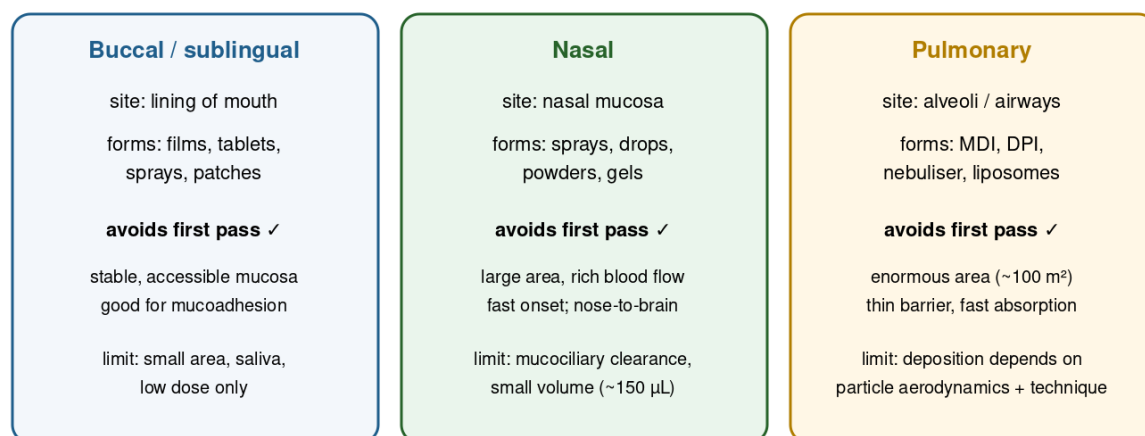


Figure 9.1. The three mucosal routes covered in this chapter. The buccal/sublingual, nasal and pulmonary mucosae all deliver drug into the systemic circulation directly, bypassing hepatic first-pass metabolism. They differ in absorptive area, in the dosage forms suited to each, and in the physiological barriers, salivary clearance, mucociliary clearance and aerodynamic deposition, that limit them.

Each route has a characteristic limitation that shapes how it is used, and the rest of the chapter is organised around matching the route's strengths against those limitations.

9.2 Buccal and sublingual delivery

The lining of the mouth offers a mucosa that is accessible, relatively robust, and richly supplied with blood that drains into the systemic circulation, bypassing the liver. Two sub-routes are distinguished. **Sublingual** delivery, under the tongue, uses a thin, highly permeable mucosa and gives very fast absorption, which suits drugs needed urgently, the classic example being sublingual glyceryl trinitrate for angina, a drug that is useless orally because of near-complete first-pass metabolism (Chapter 1) but acts within minutes sublingually. **Buccal** delivery, against the inside of the cheek, uses a slightly thicker mucosa but offers a more stable surface that can hold a dosage form in place for prolonged, controlled release.

The dosage forms reflect these properties: fast-dissolving sublingual tablets and films for rapid onset, and buccal tablets, films and patches for sustained delivery. The enabling technology for sustained buccal delivery is **mucoadhesion**, introduced in Chapter 3. A buccal patch or film is formulated with a mucoadhesive polymer, such as a carbomer, chitosan or a cellulose derivative, that binds to the buccal mucus and holds the dose against the mucosa for hours, often with a backing layer to direct release toward the mucosa and away from the saliva. Buccal delivery is far more reliable for mucoadhesion than the gastric route of Chapter 7, because the buccal mucosa is not churned and its mucus turns over more slowly. The principal limitations are the small absorptive area, the constant flushing by saliva and the risk of the patient swallowing the dose, and consequently the route suits only relatively low-dose, potent drugs.

9.3 Nasal delivery

The nasal mucosa has several features that make it one of the most useful non-injection systemic routes. It presents a comparatively large surface area, folded over the turbinates, with a thin epithelium and a rich blood supply directly beneath it, so absorption can be rapid and onset fast, again with first-pass avoidance. Nasal sprays, drops, powders and gels are all used. Beyond the familiar local use for decongestants and allergy, the nasal route is used systemically for drugs needing rapid action, and it is the route for several peptide products where a fast, needle-free option is valued.

Two further points distinguish the nasal route. First is the **nose-to-brain** possibility: a fraction of a nasally administered drug may reach the brain directly along the olfactory and trigeminal nerve pathways, partially bypassing the blood–brain barrier. This is an area of considerable research interest for delivering drugs to the central nervous system, though the efficiency and clinical reliability of the pathway in humans remain under active investigation and should not be overstated.

The dominant limitation is **mucoiliary clearance**. The nasal mucosa is lined with cilia that continuously sweep the overlying mucus, and anything deposited on it, toward the throat to be swallowed, clearing deposited material on a timescale of minutes. A nasal formulation therefore has only a short window to be absorbed before it is cleared, which caps bioavailability, particularly for larger and poorly permeable molecules. Formulation answers include mucoadhesive agents and gelling systems to prolong contact, and permeation enhancers to speed absorption within the available window. The second limitation is simply volume: only a small volume, on the order of 100 to 150 microlitres per nostril, can be

administered before the excess drains away, so the route again suits potent, low-volume drugs.

9.4 Pulmonary delivery

The lung is, by absorptive measures, the most capable of the mucosal routes. Its total alveolar surface area is enormous, of the order of a hundred square metres, the alveolar barrier between air and blood is extraordinarily thin, and the blood supply is the entire cardiac output. A drug that reaches the alveoli can be absorbed very rapidly and completely, and again avoids the liver on first pass. The lung is the obvious route for drugs acting locally on the airways, the inhaled bronchodilators and corticosteroids of asthma and COPD, and it is also a route for systemic delivery, most famously attempted for inhaled insulin (taken up as a case study in Unit IV).

What makes pulmonary delivery distinctive, and difficult, is that the drug must be carried to the alveoli as an aerosol, and where an aerosol particle lands depends almost entirely on its **aerodynamic particle size** (Figure 9.2).

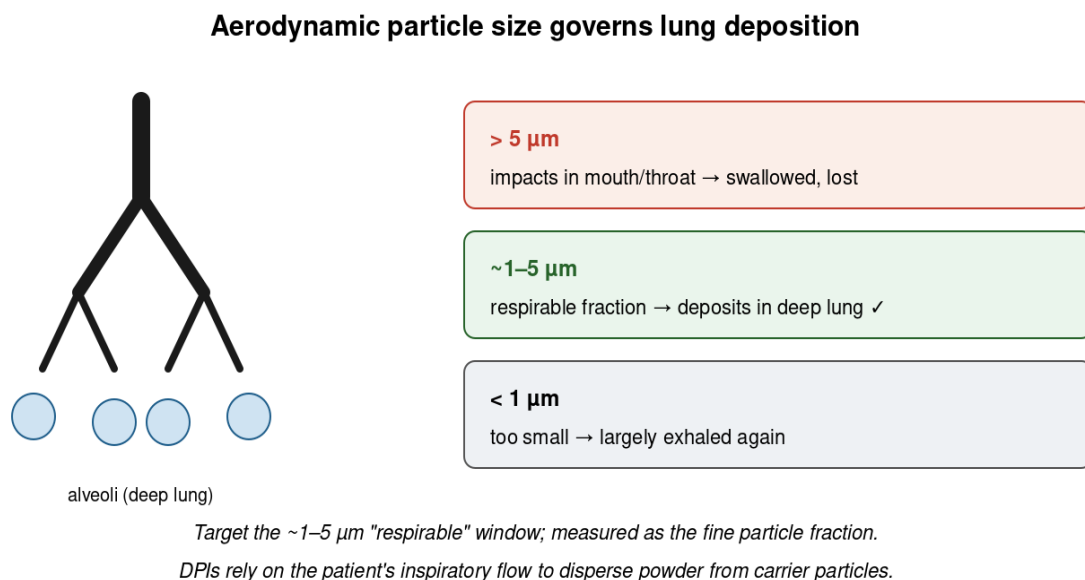


Figure 9.2. Lung deposition as a function of aerodynamic particle size. Particles larger than about 5 micrometres impact in the mouth and throat and are swallowed; particles smaller than about 1 micrometre are largely exhaled before they can deposit; only the intermediate “respirable” range of roughly 1 to 5 micrometres deposits efficiently in the deep lung. The proportion of an aerosol in this range is the **fine particle fraction**.

The physics is unforgiving. Particles larger than roughly 5 micrometres have too much inertia to follow the bending airstream and impact in the mouth and throat, where they are swallowed and wasted. Particles smaller than about 1 micrometre are so light that they stay suspended in the air and are simply breathed out again before they settle. Only the intermediate window, an aerodynamic diameter of roughly 1 to 5 micrometres, deposits efficiently in the deep lung. The fraction of an inhaled dose that falls in this window is the **fine particle fraction**, and maximising it is the central formulation goal of every inhaled product.

Three device classes deliver pulmonary aerosols. The **pressurised metered-dose inhaler (pMDI)** uses a propellant to expel a metered spray. The **nebuliser** converts a liquid into a fine mist for tidal breathing. The **dry powder inhaler (DPI)** delivers the drug as a powder. The DPI is worth understanding in detail because it illustrates a recurring formulation principle. A drug micronised to the 1–5 micrometre respirable size is extremely cohesive and flows poorly, too poorly to meter and disperse reliably. The standard solution is to blend the fine drug particles with much larger carrier particles, usually lactose, of a few tens of micrometres. The drug adheres loosely to the carrier, the blend flows and meters well, and when the patient inhales, the energy of the inspiratory airflow strips the fine drug particles off the carrier so that the drug travels deep into the lung while the large lactose carrier impacts in the throat and is harmlessly swallowed. A consequence is that a DPI is **breath-actuated** and depends on the patient generating sufficient inspiratory flow to achieve that separation, which is a genuine limitation in patients with weak inspiration, such as young children or those in acute respiratory distress.

9.5 Vesicular and particulate carriers for mucosal and pulmonary delivery

The carriers of Unit I reappear here as a way to improve mucosal and pulmonary delivery. **Liposomes** are particularly useful in the lung. Because they are made of phospholipids akin to natural lung surfactant, they are well tolerated in the airways, and they serve two purposes: they prolong the residence and sustain the release of a drug in the lung, reducing dosing frequency, and they can carry poorly soluble or irritant drugs in a better-tolerated form. Liposomal formulations of inhaled antibiotics and antifungals have been developed to sustain high local lung concentrations while limiting systemic exposure.¹ Similarly, micro- and nanoparticles of biodegradable polymers can be engineered to the respirable size while providing sustained release in the lung, and mucoadhesive nanoparticles can prolong contact at the nasal mucosa against mucociliary clearance. In each case the carrier is doing what it did in Unit I, controlling release and protecting the drug, but adapted to the specific clearance mechanism of the mucosal site.

Quick-recall box — mucosal routes at a glance All three bypass first pass. Buccal/sublingual: fast (sublingual) or sustained-mucoadhesive (buccal); low-dose only; limited by saliva and small area. Nasal: large area, fast onset, possible nose-to-brain; limited by mucociliary clearance (minutes) and ~150 μL volume. Pulmonary: huge area, fast and near-complete absorption; deposition needs 1–5 μm aerodynamic size (fine particle fraction); DPIs blend micronised drug with coarse lactose carrier and are breath-actuated.

Chapter summary

The buccal, nasal and pulmonary mucosae share the decisive advantage of delivering drug to the systemic circulation while bypassing hepatic first-pass metabolism, and they avoid gastric acid and enzymes, but each is shaped by a characteristic clearance mechanism. Buccal and sublingual delivery use the accessible, stable lining of the mouth: sublingual forms give very rapid onset, as with glyceryl trinitrate, while buccal patches and films use mucoadhesion to hold the dose for sustained release, the route being limited to potent, low-dose drugs by the small area and salivary washout. Nasal delivery offers a large, well-perfused mucosa with rapid onset and an intriguing nose-to-brain pathway, but is constrained by mucociliary clearance that removes deposited material within minutes and by a small administrable volume, prompting the use of mucoadhesives and permeation enhancers. Pulmonary delivery is the most absorptively capable route thanks to the lung's vast thin-barriered surface, but

success hinges on aerodynamic particle size, since only the 1 to 5 micrometre respirable window deposits in the deep lung; dry powder inhalers solve the poor flow of micronised drug by blending it with coarse lactose carrier and rely on the patient's inspiratory effort to separate the two. Across all three routes, the vesicular and particulate carriers of Unit I are deployed to prolong residence and sustain release against the site's particular clearance mechanism, with liposomes especially suited to the lung.

Self-assessment

Short-answer (2 marks each)

1. Why do buccal, nasal and pulmonary routes avoid first-pass metabolism?
2. Distinguish sublingual and buccal delivery in onset and dosage-form purpose.
3. What is mucociliary clearance and how does it limit nasal delivery?
4. Define the respirable particle-size range and the fine particle fraction.
5. Why is micronised drug blended with lactose carrier in a dry powder inhaler?

Long-answer (5–10 marks each)

6. Compare the buccal, nasal and pulmonary routes in absorptive properties, dosage forms and limiting clearance mechanism, with a diagram.
7. Explain, with a diagram, how aerodynamic particle size governs lung deposition and how a DPI is formulated and actuated to deliver drug to the deep lung.
8. Describe how vesicular and particulate carriers are used to improve delivery at mucosal and pulmonary surfaces, with an example.

Multiple-choice

9. The respirable aerodynamic size range for deep-lung deposition is approximately: (a) $<1\ \mu\text{m}$ (b) $1\text{--}5\ \mu\text{m}$ (c) $5\text{--}10\ \mu\text{m}$ (d) $>10\ \mu\text{m}$
10. The principal factor limiting nasal bioavailability of a peptide is: (a) gastric acid (b) first-pass metabolism (c) mucociliary clearance (d) hepatic extraction
11. A dry powder inhaler is: (a) propellant-driven (b) breath-actuated (c) a liquid mist device (d) an injection

Answer pointers. 1. Drug enters systemic veins directly rather than the hepatic portal vein. 2. Sublingual = rapid onset (thin, permeable, fast-dissolving forms); buccal = stable surface for sustained mucoadhesive release. 3. Cilia sweep mucus and deposited drug to the throat within minutes, capping absorption time. 4. $\sim 1\text{--}5\ \mu\text{m}$ aerodynamic diameter; fine particle fraction = proportion of aerosol in that range. 5. Micronised drug is cohesive and flows poorly; coarse lactose aids flow/metering and releases drug on inhalation. 6. Per Figure 9.1. 7. Per Figure 9.2 and DPI carrier mechanism. 8. Liposomes in lung / mucoadhesive NPs nasally. 9. (b). 10. (c). 11. (b).

References

1. Cipolla D, Blanchard J, Gonda I. Development of liposomal ciprofloxacin to treat lung infections. *Pharmaceutics*. 2016;8(1):6.
2. Patton JS, Byron PR. Inhaling medicines: delivering drugs to the body through the lungs. *Nat Rev Drug Discov*. 2007;6(1):67–74.

3. Sosnik A, das Neves J, Sarmento B. Mucoadhesive polymers in the design of nano-drug delivery systems for administration by non-parenteral routes: a review. *Prog Polym Sci.* 2014;39(12):2030–75.
4. Aulton ME, Taylor KMG, editors. *Aulton's Pharmaceutics*. 6th ed. Edinburgh: Elsevier; 2022.

Chapter 10 - IVIVC and Biorelevant Dissolution Testing

Learning outcomes

After working through this chapter the student will be able to:

1. Explain what an in-vitro–in-vivo correlation (IVIVC) is and why a predictive dissolution test is valuable in developing a modified-release product.
2. Define the three levels of IVIVC (A, B and C), including multiple Level C, and rank them by usefulness.
3. Explain the concept and regulatory value of a biowaiver and how a validated Level A IVIVC supports one.
4. Distinguish simple compendial dissolution media from biorelevant media, and explain why bile salts and lecithin are added.
5. Identify the principal biorelevant fluids (FaSSGF, FaSSIF, FeSSIF) and explain why fed-state media dissolve lipophilic drugs better.

10.1 The problem an IVIVC solves

Developing a modified-release product means running clinical pharmacokinetic studies, which are slow and expensive. Every time a formulation is adjusted, or a manufacturing change is made, the worry is whether the change has altered how the product behaves in patients. Repeating a bioequivalence study for every change would be prohibitive. A dissolution test, by contrast, is fast, cheap and done in a beaker. The obvious wish is to use the cheap in-vitro test as a stand-in for the expensive in-vivo study: if dissolution in the laboratory reliably predicts absorption in people, then dissolution can be used to guide development and, in defined circumstances, to justify a change without a fresh clinical study.

An **in-vitro–in-vivo correlation (IVIVC)** is a predictive mathematical relationship between an in-vitro property of a dosage form, usually the extent or rate of drug dissolution, and a relevant in-vivo response, usually the rate or extent of drug absorption. Establishing one turns the dissolution test from a simple quality-control check into a genuine predictor of clinical performance. The value is greatest for modified-release products, where the formulation, not the drug's own properties, controls the release rate, and where a meaningful correlation is therefore most achievable.

10.2 The levels of IVIVC

Regulatory guidance recognises a hierarchy of correlations, graded by how much of the relationship they capture and therefore how useful they are (Figure 10.1). The levels are conventionally labelled A, B and C, with Level A the most informative.¹

Levels of in-vitro–in-vivo correlation (IVIVC)

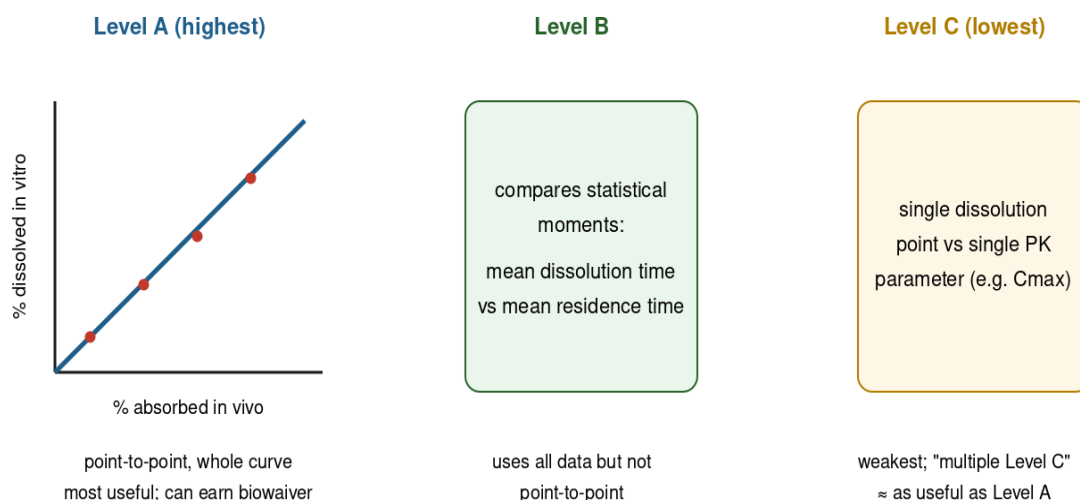


Figure 10.1. The levels of IVIVC. Level A is a point-to-point correlation between the entire in-vitro dissolution curve and the entire in-vivo absorption curve, and is the most useful, capable of supporting a biowaiver. Level B compares statistical moments such as mean dissolution time and mean residence time, using all the data but not point-by-point. Level C relates a single dissolution point to a single pharmacokinetic parameter and is the weakest, though multiple Level C correlations across several time points approach the value of Level A.

Level A is the highest and most useful correlation. It is a point-to-point relationship: the entire in-vitro dissolution curve is correlated with the entire in-vivo absorption curve, so that each point on one corresponds to a point on the other. In the ideal case the two curves are superimposable or related by a simple linear function. A Level A correlation means the dissolution test is, in effect, a surrogate for the in-vivo behaviour across the whole time course, and it is the only level that can reliably support a biowaiver (Section 10.3).

Level B uses the principles of statistical moment analysis. It compares a summary statistic of the dissolution curve, the **mean dissolution time**, with a summary statistic of the in-vivo curve, the **mean residence time** (or mean in-vivo dissolution time). Level B uses all of the data, which is to its credit, but it collapses each curve into a single average value, so it is not point-to-point. Because several different curves can share the same mean, a Level B correlation does not uniquely predict the in-vivo profile and is of limited regulatory use on its own.

Level C is the weakest. It relates a single dissolution time point, for example the percentage dissolved at one hour, to a single pharmacokinetic parameter, such as Cmax or AUC. It is a single-point correlation and captures none of the shape of the curve. A refinement, **multiple Level C**, establishes such a relationship at several dissolution time points across the profile; when this can be done across the clinically relevant range, multiple Level C is considered nearly as informative as Level A, and where multiple Level C is possible a Level A correlation is usually achievable and preferred.

The practical hierarchy to remember is therefore: Level A (point-to-point, whole curve, most useful) > multiple Level C ≈ Level A in value > Level B > single Level C.

10.3 Biowaivers

The chief regulatory prize of a strong IVIVC is the **biowaiver**, a waiver of the requirement to conduct an in-vivo bioequivalence study. If a validated Level A IVIVC exists for a product, then certain post-approval changes, for example a change in the manufacturing site, in equipment, or a moderate change in the formulation or release-controlling excipient, can be supported by showing that the dissolution profile is unchanged, rather than by running a new clinical bioequivalence study. The dissolution test, trusted because the IVIVC has shown it predicts in-vivo behaviour, stands in for the study. This saves time and cost and reduces unnecessary human exposure, and it is the main commercial reason companies invest in establishing an IVIVC during development.

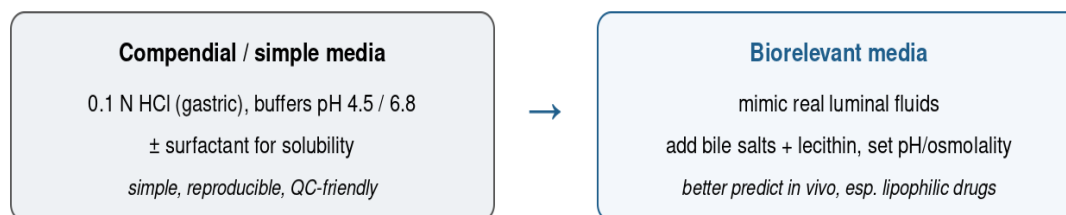
The term biowaiver is also used in a second, related context, the BCS-based biowaiver, in which highly soluble, highly permeable BCS Class I (and, under defined conditions, Class III) immediate-release drugs may be granted bioequivalence on the basis of dissolution data and BCS properties without a clinical study. The two uses share the underlying logic that a well-understood dissolution test can substitute for an in-vivo study when the science justifies it.

10.4 Why dissolution media matter: from compendial to biorelevant

An IVIVC is only as good as the dissolution test that underlies it, and the choice of dissolution medium is decisive. The traditional **compendial media** are simple aqueous solutions chosen for reproducibility and ease of quality control: dilute hydrochloric acid (around 0.1 N HCl, pH ~1.2) to represent the stomach, and aqueous buffers at pH 4.5 and pH 6.8 to represent points along the intestine, sometimes with a synthetic surfactant such as sodium lauryl sulfate added to help dissolve a poorly soluble drug. These media are excellent for routine quality control because they are simple, cheap and reproducible. Their weakness is that they do not resemble the actual fluids in the gut, and for a poorly soluble (BCS Class II) drug whose absorption is dissolution-limited, a test in simple buffer can badly mispredict how the drug will dissolve and be absorbed in vivo.

Biorelevant media were developed to close this gap (Figure 10.2). The key insight is that real intestinal fluid is not simple buffer: it contains **bile salts and lecithin (phospholipid)**, natural surfactants secreted in the bile that form mixed micelles and dramatically increase the solubility of lipophilic drugs, exactly the drugs that dissolve poorly in plain buffer. A biorelevant medium reconstructs the composition of the relevant luminal fluid, setting the pH, the buffer capacity, the osmolality and, crucially, including bile salt and lecithin at physiological concentrations.

From simple to biorelevant dissolution media



Common biorelevant fluids:

FaSSGF — fasted-state simulated gastric fluid (low pH, low bile)

FaSSIF — fasted-state simulated intestinal fluid (pH ~6.5, bile salt + lecithin)

FeSSIF — fed-state simulated intestinal fluid (lower pH ~5, higher bile salt + lecithin)

Fed-state media carry far more bile → dissolve lipophilic drugs much better than QC media.

Figure 10.2. The move from simple compendial media to biorelevant media. Compendial media (dilute acid, pH 4.5 and 6.8 buffers, optionally with a synthetic surfactant) are simple and reproducible, suiting quality control. Biorelevant media reconstruct the real luminal fluids by adding bile salts and lecithin and setting physiological pH and osmolality, which predicts in-vivo dissolution far better, especially for lipophilic drugs. Fed-state media contain much higher bile concentrations than fasted-state media.

10.5 The biorelevant fluids

A small family of biorelevant fluids is in standard use, designed to mimic specific regions and prandial states.²

FaSSGF, fasted-state simulated gastric fluid, mimics the empty stomach: low pH and only a small amount of bile-derived surfactant.

FaSSIF, fasted-state simulated intestinal fluid, mimics the upper small intestine in the fasted state, at a pH around 6.5 and containing bile salt and lecithin at the modest concentrations found between meals.

FeSSIF, fed-state simulated intestinal fluid, mimics the upper small intestine after a meal. Two features distinguish it from FaSSIF: the pH is lower (around 5), reflecting the acid carried through from the fed stomach, and, most importantly, the concentrations of bile salt and lecithin are much higher, because a meal triggers a large release of bile.

The single most useful comparison to carry away is between FaSSIF and FeSSIF. Because FeSSIF contains far more bile-derived surfactant, it dissolves lipophilic, poorly soluble drugs much more effectively than FaSSIF. This is the in-vitro counterpart of the well-known clinical observation that many poorly soluble drugs are absorbed considerably better when taken with food, the **positive food effect**: food triggers bile secretion, bile solubilises the drug, and more dissolves and is absorbed. Testing a Class II drug in both fasted- and fed-state biorelevant media can therefore predict, and help explain, a food effect before it is ever seen in a clinical study, which is a direct, practical payoff of biorelevant dissolution testing and a fitting close to the oral-delivery unit.

Quick-recall box — IVIVC and media IVIVC levels: A = point-to-point, whole curve, supports a biowaiver; B = mean dissolution time vs mean residence time (moments); C = single point vs single PK parameter; multiple Level C $\approx$ Level A. Biowaiver = skip the in-vivo bioequivalence study when a validated correlation lets dissolution stand in. Media: compendial (acid, pH 4.5/6.8) for QC; biorelevant (add bile salt + lecithin) for prediction. FaSSGF/FaSSIF/FeSSIF; FeSSIF has much more bile $\rightarrow$ dissolves lipophilic drugs better $\rightarrow$ predicts the positive food effect.

Chapter summary

A dissolution test is fast and cheap; a clinical pharmacokinetic study is slow and expensive. An in-vitro–in-vivo correlation is the predictive relationship that lets the former stand in for the latter, and it is most valuable for modified-release products whose release the formulation controls. IVIVC comes in graded levels: Level A is a point-to-point correlation of the whole dissolution curve with the whole absorption curve and is the only level that reliably supports a biowaiver; Level B compares statistical moments, the mean dissolution time against the mean residence time, using all the data but not point-by-point; and Level C relates a single dissolution point to a single pharmacokinetic parameter and is weakest, though multiple Level C across several points approaches the value of Level A. The regulatory reward for a validated Level A correlation is the biowaiver, allowing certain manufacturing or formulation changes to be justified by an unchanged dissolution profile rather than a new clinical study. Any IVIVC, however, is only as sound as its dissolution medium. Simple compendial media of dilute acid and pH 4.5 or 6.8 buffer suit quality control but poorly mimic real gut fluids, so for poorly soluble drugs they mispredict. Biorelevant media reconstruct the luminal fluids by adding bile salts and lecithin and setting physiological pH and osmolality. The standard fluids FaSSGF, FaSSIF and FeSSIF represent the fasted stomach, fasted intestine and fed intestine; because FeSSIF carries far more bile, it dissolves lipophilic drugs much better than FaSSIF, predicting the positive food effect seen clinically and demonstrating the practical power of biorelevant testing.

Self-assessment

Short-answer (2 marks each)

1. Define an IVIVC and state for which kind of product it is most useful.
2. What distinguishes a Level A from a Level C correlation?
3. What is a biowaiver, and what level of IVIVC is needed to support one?
4. Why are bile salts and lecithin added to biorelevant media?
5. Name the three principal biorelevant fluids and the region/state each represents.

Long-answer (5–10 marks each)

6. Describe the three levels of IVIVC with a diagram, ranking them by usefulness and explaining the place of multiple Level C.
7. Explain why compendial dissolution media can mispredict the in-vivo behaviour of a poorly soluble drug, and how biorelevant media address this.
8. Explain how testing a Class II drug in FaSSIF and FeSSIF can predict a positive food effect, relating the in-vitro result to the underlying physiology.

Multiple-choice

9. The IVIVC level that is point-to-point across the whole curve is: (a) Level A (b) Level B (c) Level C (d) multiple Level C
10. Compared with FaSSIF, FeSSIF contains: (a) less bile salt (b) more bile salt and lecithin (c) no surfactant (d) higher pH
11. A validated Level A IVIVC can support a: (a) new clinical trial requirement (b) biowaiver (c) higher dose (d) change of indication

Answer pointers. 1. Predictive relationship between in-vitro dissolution and in-vivo absorption; most useful for modified-release products. 2. Level A = point-to-point whole-curve; Level C = single dissolution point vs single PK parameter. 3. Waiver of an in-vivo bioequivalence study; Level A. 4. They are natural surfactants forming mixed micelles that solubilise lipophilic drugs, mimicking real intestinal fluid. 5. FaSSGF (fasted stomach), FaSSIF (fasted intestine), FeSSIF (fed intestine). 6. Per Figure 10.1. 7. Per Section 10.4. 8. FeSSIF's higher bile dissolves more drug → predicts better fed absorption; physiology = meal triggers bile. 9. (a). 10. (b). 11. (b).

References

1. US Food and Drug Administration. Guidance for Industry: Extended Release Oral Dosage Forms — Development, Evaluation, and Application of In Vitro/In Vivo Correlations. Rockville (MD): FDA; 1997.
2. Jantratid E, Janssen N, Reppas C, Dressman JB. Dissolution media simulating conditions in the proximal human gastrointestinal tract: an update. *Pharm Res.* 2008;25(7):1663–76.
3. Dressman JB, Amidon GL, Reppas C, Shah VP. Dissolution testing as a prognostic tool for oral drug absorption: immediate release dosage forms. *Pharm Res.* 1998;15(1):11–22.
4. Aulton ME, Taylor KMG, editors. *Aulton's Pharmaceutics*. 6th ed. Edinburgh: Elsevier; 2022.

Unit III - Transdermal NDDS

Unit overview. The skin is the body's largest organ and, for the formulator, its most stubborn barrier. It evolved precisely to keep the outside world out, which is exactly the opposite of what a drug delivery system wants from it. This unit is about getting drugs across, or around, that barrier, and about the related family of long-acting systems that place a depot at or near the site of action. Chapter 11 sets out why transdermal delivery is worth the trouble, how the skin is built, which drugs can cross it unaided, and how patches and permeation enhancers are designed. Chapter 12 widens the scope to other long-acting systems that share the depot philosophy: in-situ gels, microspheres, implants and ocular inserts. Chapter 13 then takes on the drugs the skin simply will not pass on its own, the large and water-soluble ones, and the physical technologies, microneedles, iontophoresis and sonophoresis, that force the barrier open, closing with the manufacturing realities of sterility, scale-up and process analytical technology that decide whether any of this reaches a patient.

Chapter 11 - Transdermal Drug Delivery: Need, Structure, Components and Permeation Enhancers

Learning outcomes

After working through this chapter the student will be able to:

1. State the advantages and disadvantages of transdermal delivery, and explain the first-pass and steady-state arguments in its favour.
2. Describe the layered structure of the skin and identify the stratum corneum as the rate-limiting barrier, including the brick-and-mortar model.
3. Explain the three routes of permeation and which dominates for most drugs.
4. Apply the physicochemical criteria, the “500-Dalton rule” and its companions, that decide whether a drug is a transdermal candidate.
5. Describe the components and architectures of a transdermal patch and the mechanisms by which chemical permeation enhancers work.

11.1 Why deliver through the skin

The transdermal route shares the great advantage of the mucosal routes of Chapter 9: a drug absorbed through the skin enters the systemic veins directly and so escapes hepatic first-pass metabolism. But the skin adds something the buccal and nasal mucosae cannot easily give, which is duration. A patch can stay on the skin for one to seven days, releasing drug steadily the whole time, so a single application replaces many oral doses and holds the plasma concentration on the flat plateau that Chapter 1 set as the goal. The benefits follow directly: steady plasma levels without peaks and troughs, far less frequent dosing and therefore better adherence, avoidance of the gut and the liver, and a delivery that the patient can start and, crucially, stop simply by removing the patch, which is a real safety advantage for potent drugs.

The disadvantages are equally real and define the route’s narrow scope. The skin is an extraordinarily effective barrier, so only a small set of drugs can cross it in useful amounts. The dose rate is limited, which restricts transdermal delivery to potent drugs needed in small daily amounts, on the order of milligrams. Skin irritation and sensitisation at the application site are common nuisances. And there is wide variation in skin permeability between body sites, between individuals, and with age, which adds variability. The route is therefore powerful but selective: superb for the right drug, impossible for most.

11.2 The structure of the skin and the barrier

To understand why so few drugs cross the skin, one has to understand how it is built (Figure 11.1). The skin has three principal layers. The deepest is the **dermis**, a thick layer of connective tissue carrying the blood vessels; once a drug reaches the dermal capillaries it is taken into the systemic circulation. Above it is the **viable epidermis**, living cells that pose little barrier to most drugs. The barrier, almost entirely, is the thin outermost layer: the **stratum corneum**.

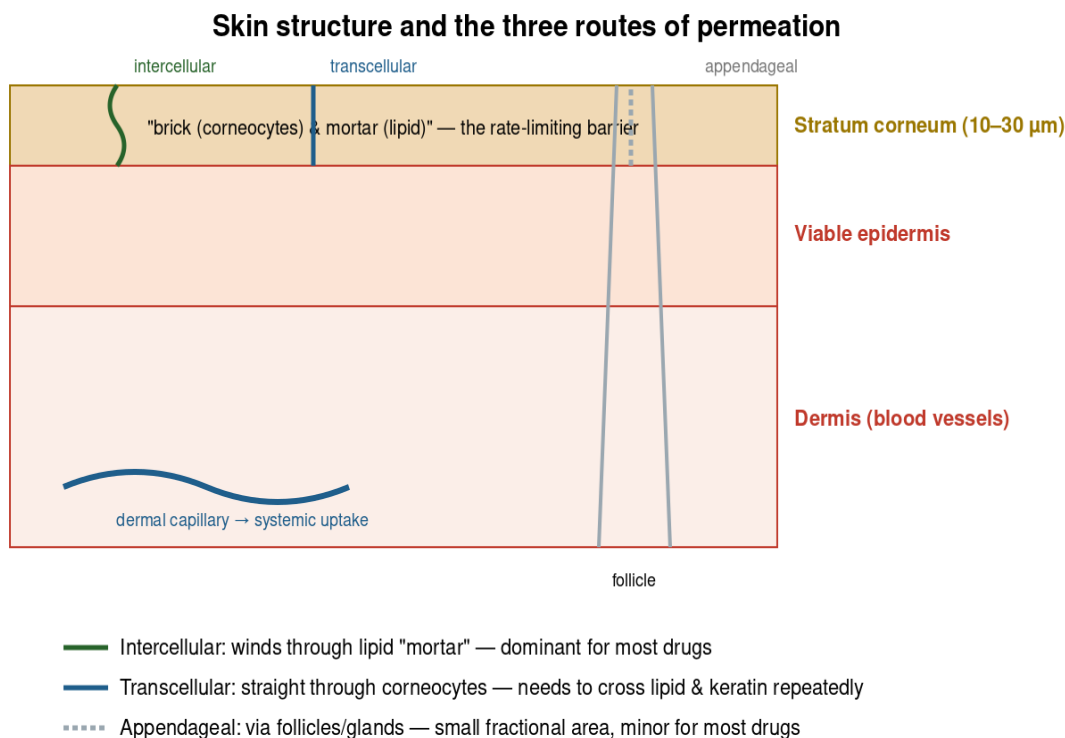


Figure 11.1. The layered structure of the skin and the three routes of permeation. The stratum corneum, only 10–30 μm thick, is the rate-limiting barrier, built as a “brick and mortar” of flattened dead corneocytes (bricks) embedded in lipid lamellae (mortar). Most drugs cross by the tortuous intercellular route through the lipid; the transcellular route passes straight through cells but must repeatedly cross lipid and keratin; the appendageal route through follicles and glands offers only a small fractional area. Below, the dermal capillaries carry absorbed drug into the systemic circulation.

The stratum corneum is only about 10 to 30 micrometres thick, yet it accounts for nearly all of the skin’s resistance to drug penetration.¹ Its structure is captured by the **brick-and-mortar model**. The “bricks” are corneocytes, flattened, dead, keratin-filled cells. The “mortar” filling the spaces between them is a highly ordered, multi-lamellar matrix of lipids, ceramides, cholesterol and free fatty acids. This arrangement is what makes the barrier so formidable: a permeating molecule cannot simply pass through the dead bricks, and the continuous lipid mortar is both tightly packed and chemically demanding to cross.

11.3 The three routes of permeation

A molecule has, in principle, three ways across the stratum corneum (Figure 11.1). The **intercellular route** winds through the lipid mortar between the corneocytes; it is a long, tortuous path, but it is continuous, and for the great majority of drugs it is the dominant route of permeation. The **transcellular route** goes straight through the corneocytes, the shortest geometric path, but it forces the molecule to partition repeatedly into and out of the lipid mortar and the keratin-rich, relatively hydrophilic cell interiors, an alternation that most molecules cannot make efficiently. The **appendageal (transappendageal) route** passes through the hair follicles and sweat glands; these are genuine shunts through the barrier, but they occupy only a tiny fraction of the skin’s surface area, so for most drugs their contribution is minor, though they can matter for the initial lag phase and for certain large or charged molecules and particles.

The practical conclusion is that designing for transdermal delivery means designing for the intercellular lipid route. The drug must be able to dissolve in and diffuse through the lipid mortar, and the strategies for enhancing permeation, both chemical and physical, are largely strategies for making that lipid route easier to traverse.

11.4 What makes a drug a transdermal candidate

Because passive permeation depends on dissolving in and diffusing through a packed lipid layer, the physicochemical properties of the drug decide whether transdermal delivery is even possible. The criteria are well established and frequently examined, and they cluster around the **500-Dalton rule**.²

The rule states that a molecule must have a molecular weight below roughly 500 daltons to cross the stratum corneum by passive diffusion in useful amounts; above this size, passive permeation falls away sharply. This is why no conventional patch delivers a protein or a large peptide, and why those molecules need the physical methods of Chapter 13. Alongside size, the drug should be moderately lipophilic, with a log P (octanol–water partition coefficient) in the range of about 1 to 3.2. The logic is a partitioning compromise: too hydrophilic (low log P) and the drug cannot enter the lipid mortar at all; too lipophilic (high log P) and it enters the lipid happily but will not partition out again into the watery viable epidermis below, so it stalls in the stratum corneum. A middle range crosses best. The drug should also have a relatively low melting point, which correlates with good solubility in the skin lipids, and few hydrogen-bonding groups, since each strong hydrogen-bonding group anchors the molecule and impedes diffusion. Finally, the drug must be potent, deliverable in a few milligrams per day or less, because the barrier limits the achievable flux.

These criteria together explain the actual roster of marketed transdermal drugs, all of which are small, potent and suitably lipophilic: glyceryl trinitrate, nicotine, the hormones estradiol and testosterone, fentanyl, scopolamine, clonidine, and the newer rivastigmine and rotigotine patches. No high-dose or large-molecule drug appears on this list, and the 500-Dalton rule is the reason.

11.5 Components and architecture of a transdermal patch

A transdermal patch is a small laminated device, and although designs vary, a common set of components recurs (Figure 11.2). The **backing layer** is an occlusive, drug-impermeable film that forms the outer face, protects the system and prevents drug loss to the environment; its occlusion also hydrates the skin slightly, which aids permeation. The **drug layer** holds the drug, usually together with a permeation enhancer. A **pressure-sensitive adhesive** holds the patch to the skin. A peelable **release liner** protects the adhesive during storage and is removed before application. Depending on the design a separate **rate-controlling membrane** may also be present.

Two transdermal patch architectures

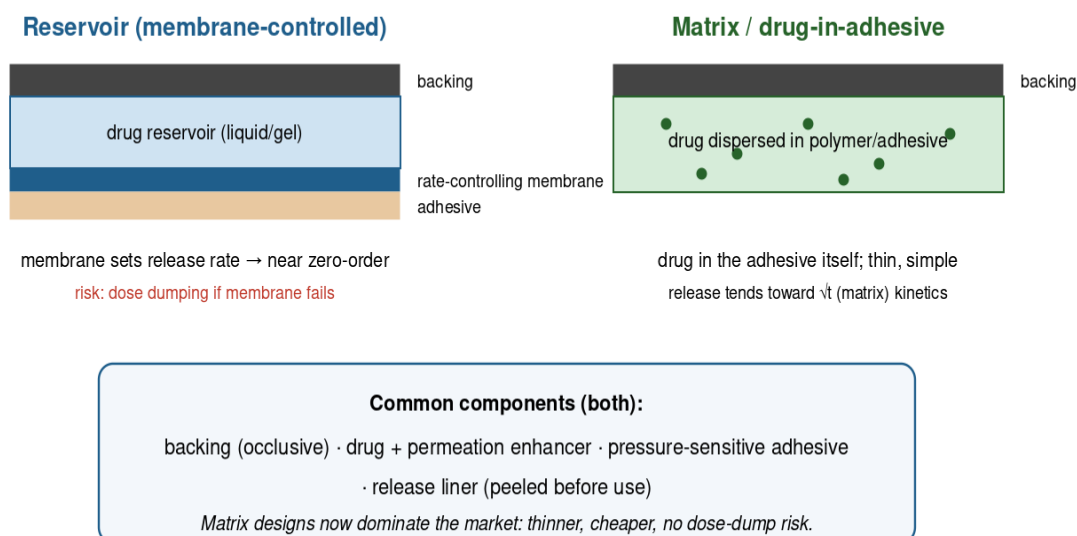


Figure 11.2. The two principal transdermal patch architectures. In the reservoir (membrane-controlled) system the drug sits in a liquid or gel reservoir and its release is governed by a rate-controlling membrane, which can give near zero-order delivery but carries a dose-dumping risk if the membrane is breached. In the matrix or drug-in-adhesive system the drug is dispersed directly in a polymer or in the adhesive itself, giving a thinner, simpler and safer patch; this design now dominates the market.

Two architectures are distinguished. In the **reservoir (membrane-controlled) system**, the drug is held as a liquid or gel in a reservoir, and a rate-controlling polymer membrane between the reservoir and the skin governs the rate of delivery. The advantage is control: the membrane can deliver drug at a near-constant, zero-order rate independent of how much remains in the reservoir. The disadvantage is the **dose-dumping** risk: if the membrane is cut, punctured or otherwise breached, the entire reservoir contents can be released at once, which for a potent drug such as fentanyl is dangerous.

In the **matrix (monolithic) system**, the drug is dispersed uniformly in a polymer matrix; in the popular **drug-in-adhesive** variant the drug is dispersed in the adhesive layer itself, so the adhesive and the drug-carrying matrix are one and the same. These designs are thinner, simpler, cheaper to make, more comfortable, and carry no dose-dumping risk because there is no liquid reservoir to dump. Their release tends to follow matrix (square-root-of-time) kinetics rather than zero-order, but for many drugs this is an acceptable trade. Matrix and drug-in-adhesive patches now dominate the market for these reasons.

11.6 Chemical permeation enhancers

For drugs that meet the 500-Dalton criteria but still permeate too slowly, the formulator's first tool is a **chemical permeation enhancer**, an excipient added to the patch that temporarily and reversibly reduces the barrier resistance of the stratum corneum. The ideal enhancer increases flux while being non-toxic, non-irritating, and reversible in its effect, so the barrier recovers fully after the patch is removed. Enhancers work by a small number of mechanisms, usually acting on the lipid mortar that controls the intercellular route.

Several classes are used. **Lipid-disrupting (fluidising) agents**, including terpenes such as menthol and limonene, and certain fatty acids such as oleic acid, insert themselves into the ordered lipid lamellae and disrupt their packing, making the lipid mortar more fluid and easier for the drug to diffuse through. **Solvents** such as ethanol and propylene glycol act partly by increasing the drug's solubility in the skin and partly by extracting or altering the skin lipids; ethanol also evaporates from the patch and can drag drug into the skin as it does so. **Surfactants** can solubilise and disrupt the skin lipids and proteins, though the stronger ones tend to be irritant, which limits their use. **Azone (laurocapram)** is a classic synthetic enhancer that fluidises the stratum corneum lipids. Water itself is an enhancer: occlusion by the patch hydrates the stratum corneum and a hydrated barrier is more permeable, which is part of why patches are occlusive.

The recurring theme is that chemical enhancers help drugs that are already marginal candidates; they widen the door, but they cannot make the skin pass a molecule that the 500-Dalton rule excludes outright. For those drugs, the physical methods of Chapter 13 are required.

Quick-recall box — transdermal candidacy and design Barrier = stratum corneum (10–30 μm , brick-and-mortar); dominant route = intercellular lipid. Candidate drug: MW < ~500 Da, log P ~1–3, low melting point, few H-bond groups, potent (mg/day). Patch parts: backing, drug + enhancer, adhesive, liner, $\pm$ rate-controlling membrane. Reservoir = zero-order but dose-dump risk; matrix/drug-in-adhesive = simpler, safer, market-dominant. Enhancers fluidise or extract the lipid mortar (terpenes, oleic acid, ethanol, propylene glycol, Azone); they widen the door, they do not break the 500-Da rule.

Chapter summary

Transdermal delivery bypasses the liver like the mucosal routes but adds sustained, multi-day delivery from a single application, giving steady plasma levels, better adherence, and the safety of being able to stop therapy by removing the patch. Its scope is narrow because the skin is a superb barrier. That barrier is almost entirely the stratum corneum, a 10–30 μm brick-and-mortar layer of dead corneocytes in an ordered lipid matrix, and most drugs cross it by the tortuous intercellular lipid route rather than through the cells or the sparse appendageal shunts. Whether a drug can cross at all is decided by physicochemistry summarised in the 500-Dalton rule: a candidate must be smaller than about 500 daltons, moderately lipophilic at log P around 1 to 3, low-melting, sparing in hydrogen-bonding groups, and potent enough to work at a few milligrams a day, which is why the marketed transdermal drugs are all small, potent molecules. Patches are laminated devices of backing, drug-plus-enhancer, adhesive and liner, built either as reservoir systems with a rate-controlling membrane that gives zero-order release but risks dose dumping, or as matrix and drug-in-adhesive systems that are thinner, safer and now dominant. When a marginal candidate still permeates too slowly, chemical enhancers such as terpenes, fatty acids, ethanol and propylene glycol fluidise or extract the stratum-corneum lipids to widen the intercellular route, but they cannot pass a molecule the size rule excludes, which sets up the physical methods of Chapter 13.

Self-assessment

Short-answer (2 marks each)

1. Give two advantages and two disadvantages of transdermal delivery.
2. Identify the rate-limiting barrier of the skin and describe the brick-and-mortar model.

3. Name the three routes of skin permeation and state which dominates for most drugs.
4. State the 500-Dalton rule and the ideal log P range for a transdermal candidate.
5. Differentiate reservoir and matrix patches in one respect each.

Long-answer (5–10 marks each)

6. Describe the structure of the skin and the three permeation routes with a labelled diagram, explaining why the intercellular route dominates.
7. Discuss the physicochemical criteria that determine transdermal candidacy, explaining the reasoning behind the molecular-weight and log P limits.
8. Describe the components and the two architectures of a transdermal patch, and explain the mechanisms of chemical permeation enhancers with examples.

Multiple-choice

9. The rate-limiting barrier to transdermal absorption is the: (a) dermis (b) viable epidermis (c) stratum corneum (d) subcutaneous fat
10. An ideal transdermal candidate has a molecular weight: (a) >1000 Da (b) <500 Da (c) exactly 750 Da (d) irrelevant
11. The dose-dumping risk is associated chiefly with: (a) drug-in-adhesive patches (b) matrix patches (c) reservoir patches (d) the release liner

Answer pointers. 1. Advantages: first-pass avoidance, steady levels, fewer doses, easy termination (any two); disadvantages: only potent small drugs, limited dose rate, irritation, variability (any two). 2. Stratum corneum; corneocyte bricks in lipid mortar. 3. Intercellular, transcellular, appendageal; intercellular dominates. 4. $MW < \sim 500$ Da for useful passive permeation; $\log P \sim 1-3$. 5. Reservoir = membrane-controlled, zero-order, dose-dump risk; matrix = drug dispersed, simpler/safer. 6. Per Figure 11.1. 7. Per Section 11.4 with partitioning logic. 8. Per Figure 11.2 and Section 11.6. 9. (c). 10. (b). 11. (c).

References

1. Menon GK, Cleary GW, Lane ME. The structure and function of the stratum corneum. *Int J Pharm.* 2012;435(1):3–9.
2. Bos JD, Meinardi MMHM. The 500 Dalton rule for the skin penetration of chemical compounds and drugs. *Exp Dermatol.* 2000;9(3):165–9.
3. Prausnitz MR, Langer R. Transdermal drug delivery. *Nat Biotechnol.* 2008;26(11):1261–8.
4. Banga AK. *Transdermal and Intradermal Delivery of Therapeutic Agents: Application of Physical Technologies*. Boca Raton: CRC Press; 2011.
5. Williams AC, Barry BW. Penetration enhancers. *Adv Drug Deliv Rev.* 2012;64(Suppl):128–37.

Chapter 12 - Long-Acting Systems: In-Situ Gels, Microspheres, Implants and Ocular Inserts

Learning outcomes

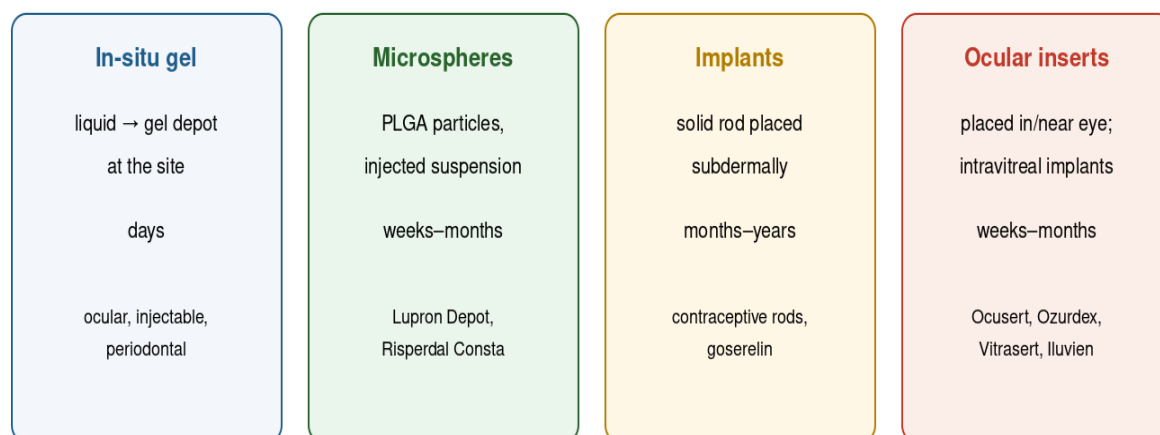
After working through this chapter the student will be able to:

1. Explain the shared philosophy of long-acting depot systems and the trade-off between invasiveness and duration.
2. Describe in-situ gelling systems and the three triggers (temperature, pH, ions) that convert an instilled or injected liquid into a gel at the site.
3. Explain how injectable microspheres provide depot delivery over weeks to months, building on the PLGA chemistry of Chapter 3.
4. Distinguish biodegradable from non-biodegradable implants and state the consequence for removal, with examples.
5. Describe ocular inserts and intravitreal implants, relate them to the physiological problem of the eye, and cite verified products.

12.1 The depot philosophy

The patches of Chapter 11 are one answer to the problem of sustained delivery; the systems in this chapter are another. What they share is the **depot principle**: instead of delivering drug repeatedly from outside, a reservoir of drug is placed at or near the site, once, and then releases slowly over a long period. The depot may be a gel that forms in place, a suspension of polymer microspheres injected into muscle, a solid rod implanted under the skin, or an insert placed in the eye. The durations stretch from days for a gel to years for an implant (Figure 12.1).

Long-acting depot systems: one placement, prolonged release



Shared logic: place a reservoir at/near the target; trade invasiveness for long, steady delivery.

Key choice: biodegradable (no removal) vs non-degradable (removable, but needs explantation).

Figure 12.1. The spectrum of long-acting depot systems, from in-situ gels lasting days, through injectable microspheres lasting weeks to months, to implants lasting months to years and ocular inserts and intravitreal implants. All place a drug reservoir at or near the target and trade some invasiveness for prolonged, steady delivery; a central design choice is whether the system is biodegradable (no removal needed) or non-degradable (removable but requiring explantation).

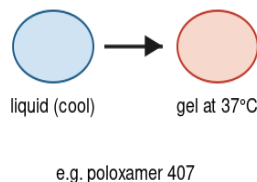
The trade-off running through the chapter is between **invasiveness and duration**. A gel instilled as drops is barely invasive but lasts only days; an implant placed surgically is invasive but lasts months or years. The second recurring decision is **biodegradable versus non-biodegradable**: a biodegradable system disappears on its own and never has to be retrieved, while a non-biodegradable one can be removed to stop therapy but must be physically explanted at the end of its life. Both these tensions appear in every system below.

12.2 In-situ gelling systems

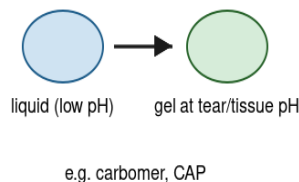
An in-situ gelling system is a clever solution to a packaging dilemma. A gel placed at a site, in the eye, in a body cavity, by injection, gives long residence and sustained release, but a preformed gel is hard to administer accurately: it cannot be instilled as a drop or injected through a fine needle. A liquid, by contrast, is easy to administer accurately but runs away from the site. The in-situ gelling system has it both ways: it is a liquid in the container and during administration, and it transforms into a gel once it reaches the site, triggered by a change in its environment (Figure 12.2). Three triggers are used.

In-situ gelling: liquid in, gel at the site (three triggers)

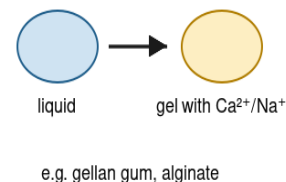
Temperature-triggered



pH-triggered



Ion-triggered



Why it helps (especially the eye):

instilled as easy-to-dose drops → gels on contact → resists drainage/blinking
→ longer residence, sustained release, fewer instillations
solves the <5% ocular bioavailability of conventional eye drops (cf. Ch 1, Ch 25)

Figure 12.2. In-situ gelling and its three triggers. A temperature-triggered system (for example poloxamer 407) is liquid when cool and gels at body temperature. A pH-triggered system (carbomer, cellulose acetate phthalate) is liquid at low pH and gels at the higher pH of the tear film or tissue. An ion-triggered system (gellan gum, sodium alginate) gels on contact with the cations of the physiological fluid. The systems are dosed as easy-to-administer liquids that then resist drainage at the site, which is especially valuable in the eye.

Temperature-triggered systems use a polymer whose solubility changes with temperature so that it is a fluid solution below body temperature and forms a gel at around 37 °C. The classic example is **poloxamer 407**, a block copolymer that gels as it warms. The product is stored and applied cold or cool, then sets when it reaches the body.

pH-triggered systems use a polymer whose ionisation, and therefore whether it stays dissolved or collapses into a gel, depends on pH. Cellulose acetate phthalate and the polyacrylic acids (carbomers) are examples: kept at a low pH in the container, they gel when they meet the higher pH of the tear film or the target tissue.

Ion-triggered systems use a polymer that is crosslinked by the cations present in physiological fluid. **Gellan gum** and **sodium alginate** gel in the presence of calcium or sodium ions; instilled as a liquid, they gel on contact with the ions in the tears or the body fluid.

The application where in-situ gelling matters most is the eye, for the reason set out in Chapter 1 and developed in Chapter 25: a conventional eye drop is drained and blinked away within minutes, so that only a few percent of the dose is absorbed and frequent dosing is needed. An in-situ gelling eye drop is instilled as easily as an ordinary drop but then gels on the eye surface, resisting drainage, prolonging residence, sustaining release and reducing the number of instillations. The same principle is used for injectable in-situ depots and for periodontal and other localised applications.

12.3 Injectable microspheres

Injectable microspheres are the depot system already met in Chapter 3, viewed now as a long-acting system in its own right. The drug is dispersed in microspheres of a biodegradable polymer, almost always PLGA, a few tens of micrometres in diameter, supplied as a powder that is reconstituted into a suspension and injected intramuscularly or subcutaneously through a conventional needle. In the tissue the microspheres form a depot that releases drug over weeks to months as the PLGA undergoes the bulk hydrolysis described in Chapter 3, with the lactide:glycolide ratio and molecular weight tuned to set the duration.

The clinical value is dramatic reductions in dosing frequency for chronic therapies: a daily or more-frequent injection is replaced by one every one to three months. The verified marketed examples are those listed in Chapter 3, including leuprolide microspheres (Lupron Depot) for hormone-responsive conditions and risperidone microspheres (Risperdal Consta) for schizophrenia, where the long interval also helps adherence in a population for whom daily dosing is difficult.¹ Because PLGA is biodegradable, the spent microspheres clear to lactic and glycolic acid and nothing has to be removed, which is a major advantage of the microsphere depot over a non-degradable implant.

12.4 Implants

An implant is a solid depot, usually a small rod or pellet, placed under the skin (subdermally) or at a target site, that releases drug over a long period, typically months to years, the longest durations of any system in this chapter. The defining design decision is whether the implant is biodegradable or not.

A **non-biodegradable implant** is made from an inert, non-degrading polymer such as silicone or ethylene-vinyl acetate, often with a rate-controlling membrane that gives steady, near zero-order release for the life of the device. Its advantage is excellent, constant release control and the ability to remove the implant to stop therapy. Its disadvantage is that, because it does not degrade, it must be surgically removed at the end of its useful life. The long-acting subdermal contraceptive implants are the most familiar example, releasing a progestogen at a controlled rate for several years.

A **biodegradable implant** is made from a polymer such as PLGA or PLA that erodes away over the release period, so no removal is needed. The goserelin implant (Zoladex), a small PLGA rod injected subcutaneously that releases the drug over one to three months for prostate and breast cancer, is a verified example.² The trade-off mirrors the microsphere case: biodegradable implants spare the patient a removal procedure but offer slightly less precise end-of-life control than a removable non-degradable device.

12.5 Ocular inserts and intravitreal implants

The eye is one of the hardest targets in delivery and one of the best illustrations of why long-acting systems exist. The front of the eye clears drops in minutes; the back of the eye, the retina and vitreous, is almost unreachable by drops or systemic dosing because of the blood–retinal barrier, yet many sight-threatening diseases are located there. Long-acting ocular systems answer both problems.

Ocular inserts are small solid or semi-solid devices placed in the conjunctival sac (or elsewhere on the eye surface) that release drug at a controlled rate over an extended period, replacing many drops with a single insert. The historic and classic example is **Ocusert**, a rate-controlled membrane insert that delivered pilocarpine for glaucoma at a steady rate over

about a week, a textbook demonstration of zero-order ocular delivery; it is the standard example of an ocular insert for glaucoma.

Intravitreal implants are placed inside the vitreous to treat diseases of the posterior segment, releasing drug directly where drops cannot reach. The verified examples make the biodegradable/non-degradable distinction concrete. **Ozurdex** is a biodegradable PLGA rod injected into the vitreous that releases dexamethasone over roughly six months and then disappears, used for macular oedema following retinal vein occlusion and for non-infectious posterior uveitis.³ **Vitrasert** was a non-biodegradable implant releasing ganciclovir over several months for cytomegalovirus retinitis, and was the first FDA-approved intravitreal implant; being non-degradable, it had to be surgically removed and replaced.⁴ **Iluvien** is a non-biodegradable fluocinolone acetonide implant that releases over a long period for diabetic macular oedema.⁴ These products show the chapter's themes in their purest form: a reservoir placed exactly where it is needed, releasing for months, with the biodegradable versions sparing a removal procedure and the non-degradable ones offering retrievability at the cost of explantation.

Quick-recall box — long-acting systems Shared logic: place a depot at/near the site, release for long. In-situ gel: liquid in, gel at site; triggers = temperature (poloxamer 407), pH (carbomer), ion (gellan/alginate); great for the eye. Microspheres: PLGA, injected, weeks–months (Lupron Depot, Risperdal Consta), biodegradable so no removal. Implants: solid rod, months–years; non-degradable (silicone/EVA, removable but needs explantation, e.g. contraceptive rods) vs biodegradable (PLGA, e.g. Zoladex). Ocular: inserts (Ocusert/pilocarpine for glaucoma) and intravitreal implants (Ozurdex biodegradable; Vitrasert/Iluvien non-degradable).

Chapter summary

Long-acting systems all follow the depot principle of placing a drug reservoir at or near the site once and releasing it slowly, trading invasiveness against duration and choosing between biodegradable systems that need no removal and non-degradable ones that can be retrieved but must be explanted. In-situ gelling systems resolve the conflict between easy administration and good retention by being liquid in the container and gelling at the site, triggered by body temperature (poloxamer 407), by the higher pH of the tear film or tissue (carbomers, cellulose acetate phthalate), or by physiological cations (gellan gum, alginate); their prime use is the eye, where they overcome the rapid drainage of ordinary drops. Injectable PLGA microspheres, reconstituted and injected through a needle, form depots releasing over weeks to months with duration tuned by polymer composition, as in Lupron Depot and Risperdal Consta, and clear without removal. Implants are solid depots lasting months to years, either non-degradable like the silicone contraceptive rods that give precise removable release but need surgical removal, or biodegradable like the goserelin Zoladex rod that erodes away. Ocular systems make the case most vividly: surface inserts such as Ocusert deliver pilocarpine steadily for glaucoma, while intravitreal implants reach the otherwise inaccessible posterior segment, with the biodegradable Ozurdex releasing dexamethasone over six months and the non-degradable Vitrasert and Iluvien offering long release at the cost of explantation.

Self-assessment

Short-answer (2 marks each)

1. State the depot principle and the invasiveness–duration trade-off.
2. Name the three triggers of in-situ gelling and give a polymer example of each.
3. Why are injectable microspheres advantageous for chronic schizophrenia therapy?
4. Distinguish biodegradable and non-biodegradable implants in terms of removal.
5. Why are conventional eye drops inefficient, and how does an in-situ gel help?

Long-answer (5–10 marks each)

6. Describe in-situ gelling systems and their three triggers with a diagram, and explain why they are particularly useful in ocular delivery.
7. Compare microspheres, implants and ocular inserts as long-acting systems, citing a verified product for each and noting the biodegradable/non-degradable choice.
8. Discuss long-acting ocular delivery, explaining the physiological problems of the anterior and posterior eye and how inserts and intravitreal implants address each, with examples.

Multiple-choice

9. Poloxamer 407 forms an in-situ gel in response to: (a) pH rise (b) calcium ions (c) temperature rise to ~37 °C (d) light
10. Ozurdex is a/an: (a) non-degradable ganciclovir implant (b) biodegradable PLGA dexamethasone intravitreal implant (c) transdermal patch (d) oral tablet
11. The chief disadvantage of a non-biodegradable implant is that it: (a) cannot be removed (b) must be surgically removed at end of life (c) releases too fast (d) cannot control release

Answer pointers. 1. Place a reservoir at the site for prolonged release; more invasive systems last longer. 2. Temperature (poloxamer 407), pH (carbomer/CAP), ion (gellan gum/alginate). 3. Replaces frequent dosing with an injection every 1–3 months, aiding adherence. 4. Biodegradable erodes/no removal; non-degradable must be explanted. 5. Drops drain/blink away in minutes (<5% absorbed); in-situ gel resists drainage, prolongs residence. 6. Per Figure 12.1. 7. Per Sections 12.3–12.5. 8. Per Section 12.5. 9. (c). 10. (b). 11. (b).

References

1. Park K, Skidmore S, Hadar J, et al. Injectable, long-acting PLGA formulations: analyzing PLGA and understanding microparticle formation. *J Control Release*. 2019;304:125–34.
2. Cheung CY, Cheng KK. Therapeutic implants for advanced prostate cancer: a review of goserelin delivery. (Review of LHRH-agonist depot systems.) *Drug Deliv Transl Res*. 2019.
3. Haghjou N, Soheilian M, Abdekhodaie MJ. Sustained release intraocular drug delivery devices for treatment of uveitis. *J Ophthalmic Vis Res*. 2011;6(4):317–29.
4. Yasin MN, Svirskis D, Seyfoddin A, Rupenthal ID. Implants for drug delivery to the posterior segment of the eye: a focus on stimuli-responsive and tunable release systems. *J Control Release*. 2014;196:208–21.

Chapter 13 - Microneedles, Physical Enhancement, Aseptic Manufacturing and PAT

Learning outcomes

After working through this chapter the student will be able to:

1. Explain why physical enhancement methods are needed for drugs the 500-Dalton rule excludes from passive transdermal delivery.
2. Describe the five microneedle types and the size principle that makes them painless yet effective.
3. Explain the mechanisms of iontophoresis and sonophoresis and the kind of drug each suits.
4. Describe the requirements of aseptic manufacturing for skin-piercing and injectable systems, and the difference between aseptic and terminal sterilisation.
5. Explain the scale-up challenges of these systems and the role of process analytical technology (PAT) in controlling their manufacture.

13.1 Beyond the passive limit

Chapter 11 ended with a hard boundary: the 500-Dalton rule excludes large and water-soluble drugs, peptides, proteins, vaccines and oligonucleotides among them, from crossing the intact stratum corneum by passive diffusion, no matter how good the chemical enhancer. These are exactly the drugs that most need a needle-free, patient-friendly delivery route, because they are otherwise confined to injection. **Physical enhancement methods** break the impasse not by coaxing the drug through the lipid mortar but by actively breaching or bypassing the barrier with energy or microscopic mechanical disruption. Three approaches dominate: microneedles, iontophoresis and sonophoresis.

13.2 Microneedles

Microneedles resolve a long-standing contradiction in skin delivery. A hypodermic needle crosses the barrier completely but is painful, frightening and requires trained administration; a patch is painless but cannot pass large molecules. A microneedle array is a patch studded with many tiny needles, each only tens to a few hundred micrometres long. This length is the whole idea (Figure 13.1): the needles are long enough to pierce through the stratum corneum and the viable epidermis, creating micro-channels straight past the barrier, but too short to reach the nerve endings and blood vessels that sit deeper in the dermis. The result is delivery across the barrier that is essentially painless and bloodless, and that can be self-administered like a patch.

Five microneedle types (each pierces the stratum corneum painlessly)



The microneedle insight:

needles are long enough (~tens–hundreds of μm) to cross the stratum corneum,
but too short to reach dermal nerves/vessels $\rightarrow$ painless, no bleeding.
Creates micro-channels $\rightarrow$ delivers large/hydrophilic drugs the 500-Da rule blocks.

Figure 13.1. The five microneedle types. Solid microneedles pierce the skin and are removed before drug is applied to the channels (“poke and patch”). Coated microneedles carry a drug coating that dissolves off once inserted. Dissolving microneedles are made of a soluble matrix that dissolves in the skin, releasing the drug and leaving nothing to remove. Hollow microneedles have a bore through which liquid formulation flows, like a miniature injection. Hydrogel-forming microneedles swell with interstitial fluid and let drug diffuse through the swollen matrix. All exploit the same size principle: long enough to cross the stratum corneum, too short to reach dermal nerves and vessels.

Five types are distinguished by how they carry and release the drug. **Solid microneedles** are inserted to make micro-channels and then withdrawn, after which a drug formulation is applied over the pre-treated skin and diffuses through the channels (the “poke and patch” approach). **Coated microneedles** are solid needles carrying a dried drug coating that dissolves into the skin after insertion. **Dissolving microneedles** are made from a water-soluble or biodegradable matrix in which the drug is embedded; after insertion the entire needle dissolves in the skin, releasing the drug and leaving no sharps waste, which is a major safety and disposal advantage. **Hollow microneedles** contain a bore through which a liquid formulation is actively driven into the skin, functioning as an array of miniature injections suited to larger liquid doses. **Hydrogel-forming microneedles** are made from a crosslinked polymer that does not dissolve but swells by taking up interstitial fluid, opening diffusion pathways through which drug from an attached reservoir passes into the skin. Microneedles are of particular interest for vaccines, both because vaccines are large molecules that passive patches cannot deliver and because the skin is rich in immune cells, which can make intradermal vaccination immunologically efficient.

13.3 Iontophoresis and sonophoresis

Two further physical methods enhance delivery without an array of needles, using an applied physical force instead (Figure 13.2).

Two physical enhancement methods

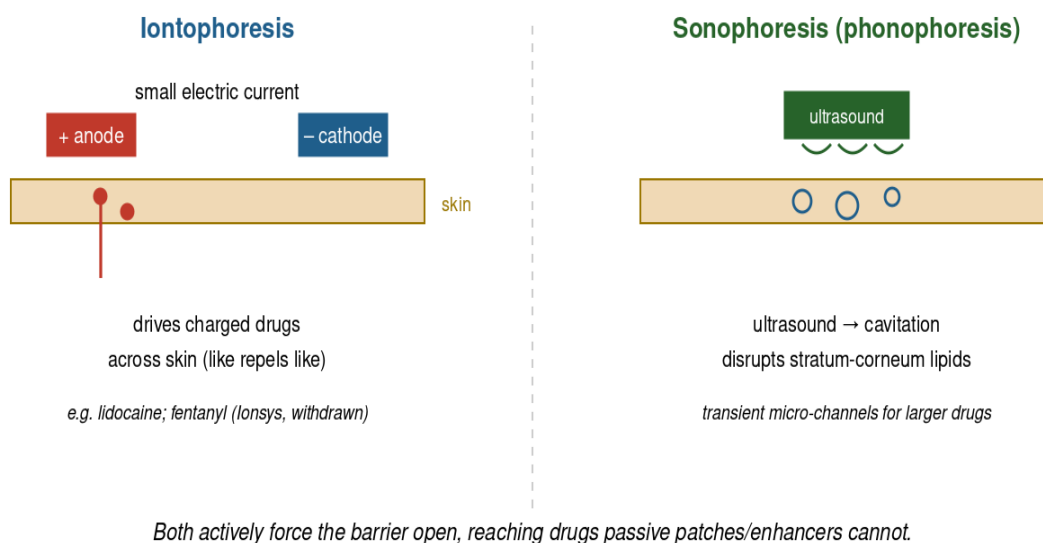


Figure 13.2. Two physical enhancement methods. Iontophoresis uses a small electric current to drive a charged drug across the skin: the drug is loaded under the electrode of the same charge, and electrostatic repulsion pushes it into and through the skin. Sonophoresis uses ultrasound, whose cavitation disrupts the ordered lipids of the stratum corneum to open transient channels for larger molecules. Both actively force the barrier open, reaching drugs that passive patches and chemical enhancers cannot.

Iontophoresis uses a small electric current to drive charged drug molecules across the skin. Two electrodes are placed on the skin; the drug is loaded under the electrode of the same charge, and because like charges repel, the current pushes the ionised drug away from the electrode and into and through the skin. A positively charged drug is driven from the anode, a negatively charged drug from the cathode. The method therefore suits charged, water-soluble molecules, exactly the kind that permeate the lipid barrier poorly on their own, and its rate can be controlled electronically by adjusting the current, which is an attractive feature for on-demand dosing. Iontophoresis has been used for local anaesthetics such as lidocaine and was the basis of a patient-controlled fentanyl system (Ionsys), although that product was later withdrawn from the market for commercial and other reasons.

Sonophoresis (phonophoresis) uses ultrasound to enhance permeation. The dominant mechanism is **cavitation**: the ultrasound waves form and collapse tiny gas bubbles within and near the stratum corneum, and the collapse disrupts the ordered lipid lamellae, transiently increasing their permeability and opening channels through which larger molecules can pass. Low-frequency ultrasound is generally more effective for this than high-frequency ultrasound. Sonophoresis can enhance the delivery of larger and hydrophilic molecules that would otherwise be excluded, and is being studied for proteins and for transdermal delivery of macromolecules.

The shared feature of all three physical methods is that they actively overcome the barrier rather than relying on the drug's own ability to cross it, which is what extends transdermal delivery to the molecules the 500-Dalton rule would otherwise rule out.

13.4 Aseptic manufacturing for skin-piercing and injectable systems

A device that pierces the skin, a microneedle array, a hollow microneedle injection, and certainly an implant or an injectable microsphere depot, breaches the body's primary defence against infection. Such products must be sterile, and that requirement reshapes their manufacture. (Sterilisation methods are treated fully in Unit V; this section gives the principle as it applies to these systems.)

Two routes to sterility exist. **Terminal sterilisation** sterilises the product in its final, sealed container, usually by heat or radiation, and is the preferred and most reliable route because it sterilises the finished article. But many of the systems in this unit cannot withstand it: the polymers and drugs of dissolving microneedles, the PLGA of microspheres and the proteins of biologic payloads are frequently degraded by terminal heat or radiation. When terminal sterilisation is impossible, the product must be made by **aseptic processing**: every component is sterilised separately by a method it can tolerate, and the product is then assembled and filled under conditions designed to exclude all contamination, in a controlled cleanroom environment with sterile components, sterile-filtered solutions, and rigorous environmental control. Aseptic processing is inherently riskier than terminal sterilisation because sterility depends on maintaining control throughout assembly rather than on a final lethal step, and it therefore demands the cleanroom classification, environmental monitoring and validation described in Unit V. The key teaching point is that the heat- and radiation-sensitivity of advanced delivery materials frequently forces the more demanding aseptic route.

13.5 Scale-up challenges

A delivery system that works in the laboratory is not yet a product. Moving from a few hand-made units to reproducible large-scale manufacture, **scale-up**, is where many novel systems falter, and the difficulties are characteristic of the field.

Several recur. The complex multi-step processes used to make nanoparticles, microspheres and microneedles often behave differently at scale: a microsphere's size distribution, or a microneedle's geometry and drug loading, can shift when the process is enlarged, and these attributes directly control release. Sterility and aseptic handling, as above, complicate every step. Reproducibility is demanding because small changes in process parameters can change the critical quality attributes of the product. And the cost and the need for specialised equipment can be prohibitive. The general lesson, developed in Unit IV through Quality by Design, is that a process which is understood and controlled scales far more reliably than one that is merely optimised by trial and error, which is the motivation for the tool described next.

13.6 Process analytical technology (PAT)

Process analytical technology (PAT) is a framework, promoted by regulators as part of the modern, science-based approach to manufacturing, for designing, analysing and controlling manufacturing through timely measurement of critical quality and performance attributes during the process itself, rather than only by testing the finished product.¹ The phrase to hold onto is its underlying philosophy: **quality should be built into the product by understanding and controlling the process, not merely tested at the end.**

In practice PAT means embedding analytical measurements into the manufacturing line, using techniques such as near-infrared and Raman spectroscopy, particle-size monitoring and other in-line or at-line sensors, so that critical attributes are measured continuously or

frequently as the product is made. If a measured attribute drifts, the process can be adjusted in real time to keep it on target, a concept known as real-time control. The benefits are a more consistent product, the possibility of real-time release based on process data rather than slow end-product testing, and a deeper understanding of how process parameters affect quality. For the complex, sensitive systems of this unit, where the critical quality attributes are set during processing and are hard to fix afterward, PAT is especially valuable, and it forms a natural bridge to the Quality by Design framework of Chapter 14, with which it shares its central conviction that quality is designed and controlled rather than inspected in at the finish.

Quick-recall box — physical methods and manufacturing Why physical methods: to deliver the large/charged drugs the 500-Da rule blocks. Microneedles (long enough to cross stratum corneum, too short to hurt): solid (poke-and-patch), coated, dissolving (no sharps), hollow (micro-injection), hydrogel-forming. Iontophoresis = electric current drives charged drugs (like repels like). Sonophoresis = ultrasound cavitation disrupts lipid lamellae. Sterility: terminal (heat/radiation, preferred) vs aseptic (when materials are heat/radiation-sensitive, common here). PAT = measure critical attributes during the process; build quality in, do not test it in.

Chapter summary

Physical enhancement methods deliver the large and water-soluble drugs that the 500-Dalton rule bars from passive transdermal delivery, by actively breaching or bypassing the stratum corneum. Microneedle arrays exploit a size principle, being long enough to pierce the barrier but too short to reach dermal nerves and vessels, so they cross the skin painlessly and bloodlessly; they come in five forms, solid (poke and patch), coated, dissolving (which leave no sharps), hollow (miniature injections) and hydrogel-forming, and are especially promising for vaccines. Iontophoresis uses a small electric current to drive charged drugs across the skin by electrostatic repulsion and offers electronically controllable rates, while sonophoresis uses ultrasound cavitation to disrupt the stratum-corneum lipids and open transient channels for larger molecules. Because these skin-piercing and injectable systems breach the body's defences they must be sterile, and since their heat- and radiation-sensitive materials often cannot survive terminal sterilisation, they are frequently made by the more demanding aseptic route under cleanroom control. Scaling such complex processes is difficult because small parameter changes shift the critical quality attributes that govern release, which is why process analytical technology, the in-line measurement and real-time control of critical attributes during manufacture, is so valuable: it builds quality into the product through process understanding rather than testing it in at the end, carrying the unit directly into the Quality by Design thinking of Unit IV.

Self-assessment

Short-answer (2 marks each)

1. Why are physical enhancement methods needed for some drugs?
2. State the size principle that makes microneedles painless yet effective.
3. Name the five microneedle types.
4. Explain how iontophoresis drives a positively charged drug into the skin.
5. Distinguish terminal sterilisation from aseptic processing, and say why advanced delivery systems often need the latter.

Long-answer (5–10 marks each)

6. Describe the five microneedle types with a diagram, and explain the advantages of dissolving and hollow types.
7. Compare iontophoresis and sonophoresis in mechanism and the type of drug each suits.
8. Explain the concept of process analytical technology, its guiding philosophy, and why it is valuable for the manufacture of complex novel delivery systems; relate it to scale-up challenges.

Multiple-choice

9. A dissolving microneedle is advantageous because it: (a) must be removed after use (b) leaves no sharps waste (c) only delivers small molecules (d) requires electric current
10. Iontophoresis is best suited to: (a) neutral lipophilic drugs (b) charged water-soluble drugs (c) gases (d) only proteins
11. The guiding philosophy of PAT is that quality should be: (a) tested only at the end (b) built in through process understanding and control (c) ignored during processing (d) assured by terminal sterilisation alone

Answer pointers. 1. To deliver large/charged drugs excluded by the 500-Da rule. 2. Long enough to cross the stratum corneum, too short to reach dermal nerves/vessels. 3. Solid, coated, dissolving, hollow, hydrogel-forming. 4. Loaded under the anode; like charges repel, so current pushes the cation into/through the skin. 5. Terminal = sterilise final sealed product (heat/radiation, preferred); aseptic = sterilise components separately and assemble under sterile conditions; needed because materials are heat/radiation-sensitive. 6. Per Figure 13.1. 7. Per Section 13.3. 8. Per Section 13.6 with scale-up link. 9. (b). 10. (b). 11. (b).

References

1. US Food and Drug Administration. Guidance for Industry: PAT — A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance. Rockville (MD): FDA; 2004.
2. Prausnitz MR. Engineering microneedle patches for vaccination and drug delivery to skin. *Annu Rev Chem Biomol Eng.* 2017;8:177–200.
3. Banga AK. Transdermal and Intradermal Delivery of Therapeutic Agents: Application of Physical Technologies. Boca Raton: CRC Press; 2011.
4. Polat BE, Hart D, Langer R, Blankschtein D. Ultrasound-mediated transdermal drug delivery: mechanisms, scope, and emerging trends. *J Control Release.* 2011;152(3):330–48.
5. ICH. Q8(R2) Pharmaceutical Development. Geneva: International Council for Harmonisation; 2009.

Unit IV - Evaluation, Quality and Translation

Unit overview. The first three units built delivery systems. This unit asks the harder questions that decide whether a clever system becomes a medicine: how it is designed for quality from the start, how it is characterised in the laboratory, how it is tested in animals before a human ever receives it, how it is brought to market through a regulatory system, and what the best examples teach. Chapter 14 introduces Quality by Design, the modern philosophy that quality is engineered rather than inspected. Chapter 15 surveys the analytical methods that measure a nanosystem's critical attributes. Chapter 16 covers the non-clinical pharmacokinetic, biodistribution and toxicology studies that gate the move to humans. Chapter 17 maps the regulatory routes, particularly the 505(b)(2) pathway built for modified versions of known drugs, and the ICH framework. Chapter 18 works through the field's defining case studies, and Chapter 19 closes with the nanosystems themselves and the precision-medicine vision toward which the whole discipline is moving. The thread tying the unit together is that translation is a discipline of evidence and control, not of invention alone.

Chapter 14 - Preformulation and Quality by Design

Learning outcomes

After working through this chapter the student will be able to:

1. Explain the purpose of preformulation studies for a novel delivery system and identify the key properties evaluated.
2. Define Quality by Design (QbD) and contrast it with the traditional quality-by-testing approach.
3. Define and correctly sequence the QbD elements: QTPP, CQAs, critical material attributes and CPPs.
4. Explain the concepts of risk assessment, design of experiments, design space and control strategy, and the meaning of the edge of failure.
5. Relate QbD to the ICH quality guidelines (Q8–Q12) and to PAT.

14.1 Preformulation for novel systems

Preformulation is the systematic characterisation of a drug's physicochemical properties before a formulation is designed, so that the formulation is built on knowledge rather than guesswork. For conventional dosage forms the standard preformulation set, solubility across pH, partition coefficient, pKa, solid-state form and polymorphism, melting point, stability to heat, light, oxygen and hydrolysis, particle size and hygroscopicity, remains the foundation. For a novel delivery system, preformulation extends to the properties that govern the delivery strategy specifically. The log P and molecular weight decide transdermal candidacy through the 500-Dalton rule of Chapter 11. The solubility and the BCS class decide whether a solubility-enhancing carrier is needed. The drug's compatibility with the carrier matters acutely: an acid-sensitive peptide may be degraded by the acidic microenvironment of an eroding PLGA depot, as Chapter 3 noted, and a lipophilic drug must actually partition into a lipid carrier to be loaded. Preformulation for an NDSS therefore asks not only “what is this molecule?” but “what will this molecule do inside the intended carrier?”, and the answers feed directly into the QbD process that organises the rest of development.

14.2 From quality-by-testing to Quality by Design

For much of pharmaceutical history, quality was assured by testing: a product was manufactured to a fixed recipe and the finished batch was tested against specifications, with batches that failed being rejected. This works, but it has a weakness. It treats the process as a black box, learns of problems only after a batch is made, and leaves little room to understand or improve. If a batch fails, the manufacturer often does not know why.

Quality by Design (QbD) is the modern alternative, formalised in the ICH Q8 guideline.¹ Its central conviction, shared with the PAT framework of Chapter 13, is that **quality cannot be tested into a product; it must be built in by design**, through a thorough understanding of how the product and its manufacturing process behave. Instead of fixing a recipe and testing the result, QbD begins with the desired outcome, works backward to identify what controls that outcome, and builds a process designed to deliver it reliably. The workflow has a definite sequence (Figure 14.1), and getting the sequence right is essential, because each step is defined by reference to the one before.

The Quality by Design (QbD) workflow (ICH Q8)

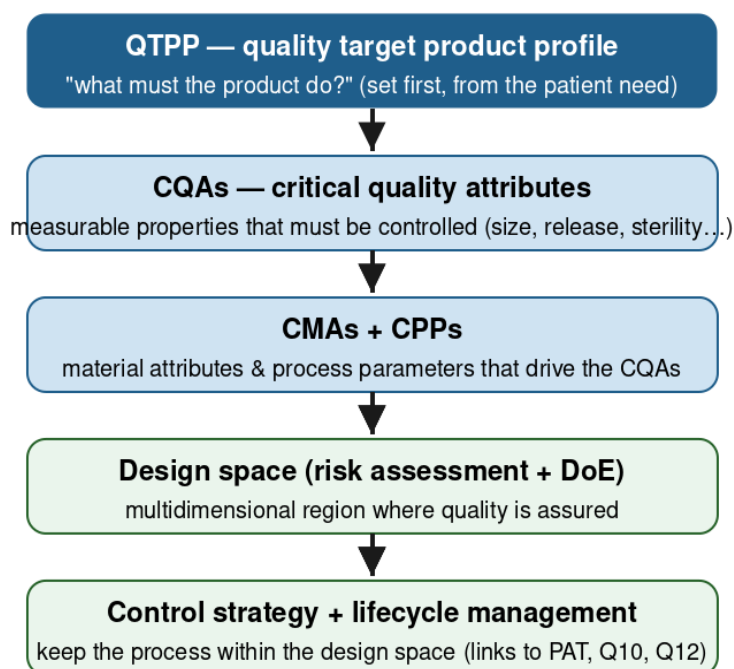


Figure 14.1. The Quality by Design workflow of ICH Q8. Development begins with the QTPP, the definition of what the product must do for the patient. From it are derived the critical quality attributes that must be controlled, then the material attributes and process parameters that drive those attributes, then a design space within which quality is assured, and finally a control strategy and lifecycle plan that keep the process inside that space. The defining feature is that design begins with the end, the patient need, and works backward.

14.3 The QbD elements in sequence

The Quality Target Product Profile (QTPP) comes first. It is a prospective summary of the quality characteristics the product should possess to deliver the intended therapeutic benefit, in effect a written definition of what the finished product must do: the route of administration, the dosage form, the dose, the desired release profile, the required stability and shelf life, and the target patient population. The QTPP is set before development begins and is derived from the patient's clinical need. Everything else in QbD is judged against it.

Critical Quality Attributes (CQAs) come next. A CQA is a physical, chemical, biological or microbiological property of the product that must be within an appropriate limit, range or distribution to ensure the desired product quality. CQAs are the measurable properties that, if they drift, would compromise the QTPP. For a nanoparticle injectable, the CQAs typically include particle size and size distribution, zeta potential, drug loading and encapsulation efficiency, the in-vitro release profile, sterility and endotoxin level. Identifying the CQAs focuses development on the handful of attributes that actually matter for performance and safety, rather than measuring everything indiscriminately.

Critical Material Attributes (CMAs) and **Critical Process Parameters (CPPs)** come third. Having defined which product attributes are critical, the formulator asks what controls them.

A CMA is a property of an input material, a raw material, an excipient or the drug substance, that affects a CQA, for example the lactide:glycolide ratio and molecular weight of the PLGA, which control the release-rate CQA. A CPP is a process parameter whose variability affects a CQA, for example the homogenisation pressure and the surfactant concentration in making a nanoparticle, which control the particle-size CQA. The CMAs and CPPs are the knobs the manufacturer can turn; the CQAs are the outcomes those knobs control.

The relationship between the knobs (CMAs, CPPs) and the outcomes (CQAs) is established by **risk assessment** followed by **design of experiments (DoE)**. Risk assessment, drawing on the quality risk management of ICH Q9, identifies which of the many possible material attributes and process parameters are likely to be critical, so that experimental effort is concentrated where it matters. DoE then varies those parameters systematically, rather than one at a time, to map quantitatively how they affect the CQAs, including how they interact with one another. The product of this work is the design space.

14.4 Design space, control strategy and the edge of failure

The **design space** is the multidimensional combination of input material attributes and process parameters that has been demonstrated to provide assurance of quality (Figure 14.2). In plain terms, it is the region of the process settings within which the product reliably meets its CQAs. The boundary of the design space is the **edge of failure**: cross it, and the product can fall out of specification.

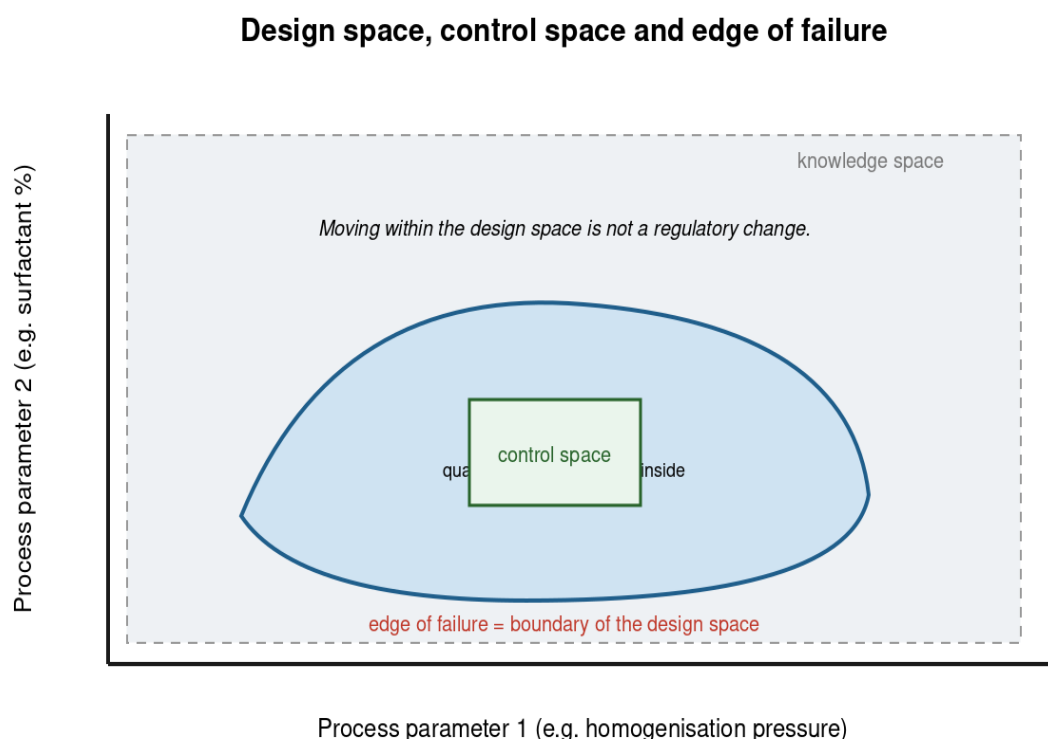


Figure 14.2. Design space, control space and edge of failure, shown for two process parameters. Within the knowledge space (everything explored) lies the design space, the region demonstrated to assure quality, bounded by the edge of failure. The manufacturer typically operates within a smaller control space well inside the design space, leaving a safety margin. A defining regulatory benefit is that moving within an approved design space is not considered a change requiring prior regulatory approval.

The design space carries a significant regulatory advantage. Because quality has been demonstrated everywhere within it, moving from one point in the design space to another, adjusting a parameter in response to a raw-material variation, for instance, is not regarded as a change requiring prior regulatory approval. This gives the manufacturer real operational flexibility, the reward for the front-loaded effort of building the design space.

The **control strategy** is the planned set of controls, derived from product and process understanding, that ensures the process stays within the design space and the product meets its CQAs. It includes in-process controls and the kind of in-line monitoring that PAT (Chapter 13) provides, specifications on input materials and on the finished product, and the procedures that hold the operation within a chosen **control space**, a margin safely inside the design space. Finally, **lifecycle management** maintains and improves this understanding across the product's commercial life, the subject of ICH Q12.

14.5 QbD and the ICH framework

QbD is not a single document but a connected family of ICH quality guidelines, and knowing how they fit together is expected at this level. **ICH Q8 (Pharmaceutical Development)** introduces the QbD concepts, the QTPP, CQAs, design space and control strategy described above. **ICH Q9 (Quality Risk Management)** provides the formal risk-assessment methods used to decide what is critical. **ICH Q10 (Pharmaceutical Quality System)** describes the company-wide quality system that operates QbD across the whole product lifecycle. **ICH Q11** covers the development and manufacture of the drug substance, applying the same principles upstream. **ICH Q12** addresses technical and regulatory tools for managing changes across the lifecycle. Together, and reinforced by the in-line measurement philosophy of PAT, these guidelines express one coherent idea: understand the product and the process deeply enough that quality is designed in, controlled continuously, and managed across the product's life, rather than inspected at the end. For the complex novel systems of this book, whose critical attributes are fixed during sensitive multi-step manufacture and are hard to correct afterward, this approach is not a bureaucratic luxury but a practical necessity.

Quick-recall box — QbD in order Sequence (never scramble it): QTPP (what the product must do, set first from patient need) → CQAs (measurable properties to control: size, zeta, loading, release, sterility) → CMAs + CPPs (the input/process knobs that drive the CQAs) → risk assessment + DoE → design space (region assuring quality; its boundary is the edge of failure; moving inside it is not a regulatory change) → control strategy + lifecycle. ICH map: Q8 development/QbD, Q9 risk, Q10 quality system, Q11 drug substance, Q12 lifecycle. Philosophy: build quality in, do not test it in.

Chapter summary

Preformulation characterises a drug's physicochemical properties before a formulation is designed, and for a novel system it extends beyond the classical set to the properties that govern the chosen delivery strategy and the drug's behaviour inside the intended carrier. Quality by Design then replaces the old quality-by-testing model with the conviction that quality must be engineered through deep understanding of product and process. Its workflow runs in a strict order: the QTPP defines what the product must do for the patient and is set first; the critical quality attributes are the measurable properties that must be controlled to meet it; the critical material attributes and critical process parameters are the input and process variables that drive those attributes; risk assessment and design of experiments establish the quantitative links between them; the design space is the demonstrated region of

settings that assures quality, bounded by the edge of failure and offering the regulatory freedom to move within it without prior approval; and the control strategy with lifecycle management keeps the process inside that space over the product's life. QbD is realised through the connected ICH guidelines Q8 to Q12 and the in-line measurement of PAT, all expressing the single principle that quality is designed and controlled, not inspected in at the finish, which for complex novel delivery systems is a practical necessity rather than an option.

Self-assessment

Short-answer (2 marks each)

1. What is the purpose of preformulation, and name two properties critical for a transdermal candidate.
2. Contrast quality-by-testing with Quality by Design in one sentence.
3. Define QTPP and state when in development it is set.
4. Distinguish a CQA from a CPP, with an example of each for a nanoparticle.
5. What is the edge of failure, and what is the regulatory benefit of operating within a design space?

Long-answer (5–10 marks each)

6. Describe the QbD workflow in correct sequence with a diagram, defining each element.
7. Explain risk assessment, design of experiments, design space and control strategy, and how they connect the process knobs to product quality.
8. Map the ICH Q8–Q12 guidelines and explain how, with PAT, they together express the QbD philosophy.

Multiple-choice

9. In QbD, the element set first, from the patient need, is the: (a) CQA (b) CPP (c) QTPP (d) design space
10. Particle size of a nanoparticle injectable is best classified as a: (a) QTPP (b) CQA (c) CPP (d) CMA
11. Moving from one point to another within an approved design space is: (a) a major change needing approval (b) not considered a change requiring prior approval (c) prohibited (d) a batch failure

Answer pointers. 1. Characterise the drug to design rationally; MW and log P (also pKa, solubility). 2. Quality-by-testing inspects the finished batch; QbD builds quality in by understanding product/process. 3. Prospective summary of required quality characteristics; set before development begins. 4. CQA = controlled product property (size); CPP = process variable affecting it (homogenisation pressure). 5. Boundary of the design space; operating inside the space allows changes without prior regulatory approval. 6. Per Figure 14.1. 7. Per Section 14.3–14.4. 8. Per Section 14.5. 9. (c). 10. (b). 11. (b).

References

1. ICH. Q8(R2) Pharmaceutical Development. Geneva: International Council for Harmonisation; 2009.
2. ICH. Q9(R1) Quality Risk Management. Geneva: ICH; 2023.
3. ICH. Q10 Pharmaceutical Quality System. Geneva: ICH; 2008.

4. Yu LX, Amidon G, Khan MA, et al. Understanding pharmaceutical quality by design. *AAPS J.* 2014;16(4):771–83.
5. Aulton ME, Taylor KMG, editors. *Aulton's Pharmaceutics*. 6th ed. Edinburgh: Elsevier; 2022.

Chapter 15 - Characterization Methods for Novel Delivery Systems

Learning outcomes

After working through this chapter the student will be able to:

1. Explain why particle size, size distribution and surface charge are critical quality attributes of a nanosystem and how dynamic light scattering measures them.
2. Interpret zeta potential as an indicator of colloidal stability.
3. Compare scanning and transmission electron microscopy for characterising particle morphology.
4. Describe how drug loading, encapsulation efficiency and in-vitro release are measured.
5. Distinguish the PAMPA and Caco-2 permeability assays and state what each does and does not capture.

15.1 Why characterization matters

Chapter 14 named particle size, zeta potential, drug loading and release profile as the critical quality attributes of a typical nanosystem. Naming them is not enough; they must be measured, reliably and reproducibly, because they determine how the system behaves in the body. The size of a nanoparticle decides whether it circulates or is cleared and whether it can exploit the EPR effect of Chapter 2. Its surface charge decides whether it stays dispersed or aggregates on the shelf. Its drug loading decides how much dose a given volume delivers, and its release profile decides the plasma curve. This chapter surveys the methods that measure these attributes, which are the analytical backbone of the QbD control strategy.

15.2 Particle size and size distribution by dynamic light scattering

For particles in the nanometre range, the standard sizing method is **dynamic light scattering (DLS)**, also called photon correlation spectroscopy (Figure 15.1). The principle rests on Brownian motion: particles suspended in a liquid move randomly because they are constantly buffeted by solvent molecules, and small particles move faster than large ones. When a laser is shone through the suspension, the particles scatter the light, and because they are moving, the intensity of the scattered light fluctuates over time. DLS measures how fast these fluctuations occur: rapid fluctuations mean fast-moving, small particles, while slow fluctuations mean large ones. From the rate of fluctuation the instrument calculates the diffusion coefficient and, through the Stokes–Einstein relationship, the **hydrodynamic diameter**, the effective size of the particle together with the layer of solvent that moves with it.

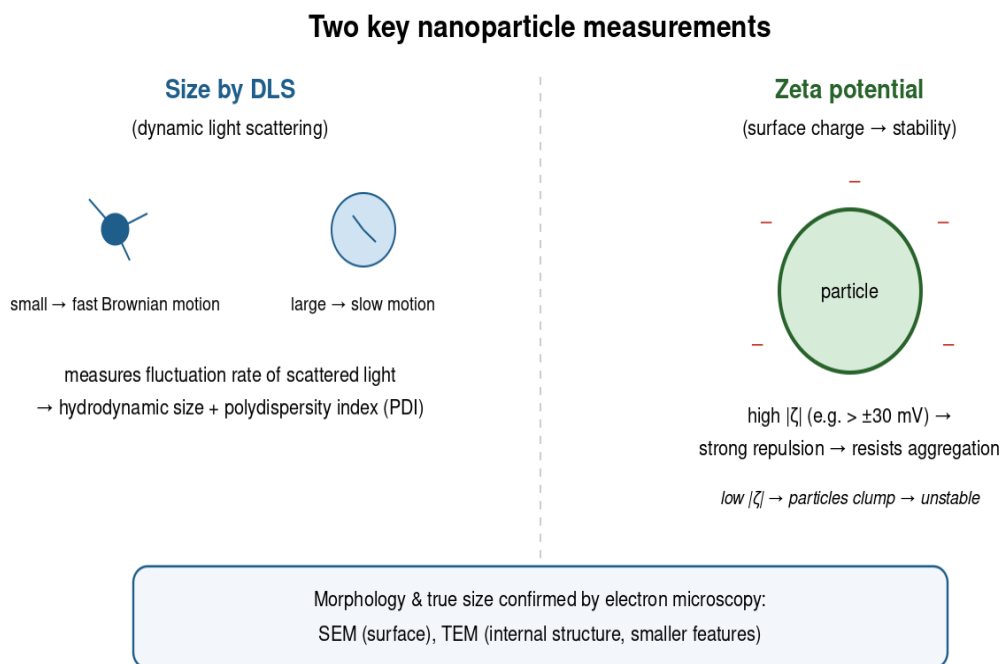


Figure 15.1. Two key nanoparticle measurements. Dynamic light scattering infers size from the rate at which scattered-light intensity fluctuates as particles undergo Brownian motion, giving the hydrodynamic diameter and the polydispersity index. Zeta potential reports the effective surface charge: a large magnitude (commonly taken as beyond about ± 30 mV) produces strong interparticle repulsion that resists aggregation, while a small magnitude allows particles to clump. Electron microscopy confirms true size and morphology.

DLS also reports the **polydispersity index (PDI)**, a dimensionless number describing the width of the size distribution. A low PDI (broadly, values approaching and below around 0.1 to 0.2) indicates a narrow, uniform, near-monodisperse population, while a high PDI indicates a broad spread of sizes. A uniform size distribution is usually desirable for reproducible behaviour, so the PDI is reported alongside the mean size as a routine pair. DLS is fast, requires little sample and is non-destructive, but it gives an intensity-weighted average that is biased toward larger particles and assumes spherical shape, which is why it is complemented by microscopy.

15.3 Zeta potential and colloidal stability

A nanoparticle in suspension usually carries a surface charge, and that charge attracts a cloud of oppositely charged ions from the surrounding medium. The **zeta potential** is the electrical potential at the boundary of the tightly bound part of this ion cloud, the slipping plane, and it serves as a practical measure of the effective surface charge of the particle. It is measured by applying an electric field across the suspension and tracking how fast the particles move (electrophoretic mobility).

The value matters for stability. Particles with the same sign of charge repel one another, and if that repulsion is strong enough it prevents them from approaching closely enough to aggregate. A large zeta-potential magnitude therefore signals a stable suspension, while a value near zero signals a suspension prone to aggregation, caking and loss of the carefully engineered size. A common rule of thumb treats a magnitude beyond roughly ± 30 millivolts as indicating good electrostatic stabilisation, though this is a guide rather than a hard

threshold, and sterically stabilised systems (for instance PEGylated particles) can be stable at lower zeta potentials because their stabilisation is physical rather than electrostatic.¹ Zeta potential also reports useful information about the surface itself, for example confirming that a cationic coating or an anionic drug has been successfully placed on the surface.

15.4 Morphology by electron microscopy

DLS gives a size but not a picture, and it assumes sphericity. To see the actual shape, surface and internal structure of a particle, electron microscopy is used. **Scanning electron microscopy (SEM)** scans a focused electron beam across the specimen surface and builds an image from the scattered and secondary electrons; it gives detailed three-dimensional-looking images of the surface topography and is excellent for examining the shape, surface texture and aggregation state of particles. **Transmission electron microscopy (TEM)** passes the electron beam through an ultrathin specimen and forms an image from the electrons transmitted through it; it offers higher resolution than SEM and, crucially, reveals internal structure, so it can show the bilayer of a liposome, the core-shell architecture of a nanoparticle, or the porous interior of a silica particle. In short, SEM shows the outside in three dimensions, TEM shows the inside at higher resolution. Both require special sample preparation and operate under vacuum, which means the sample is dried and not in its native hydrated state, a limitation to bear in mind when comparing electron-microscope sizes with the hydrodynamic sizes from DLS, which are typically somewhat larger because they include the solvation layer.

15.5 Drug loading, encapsulation efficiency and in-vitro release

Two related quantities describe how much drug a carrier holds. **Drug loading** (or loading capacity) is the amount of drug per unit weight of the carrier system, expressed as a percentage, and tells the formulator how concentrated the product is. **Encapsulation efficiency** is the fraction of the drug used in manufacture that actually ends up entrapped in the carrier rather than lost, expressed as a percentage of the input drug; it is a measure of process efficiency and of how much expensive drug is wasted. Both are measured by separating the free (unentrapped) drug from the carrier, commonly by centrifugation, ultrafiltration or dialysis, and then assaying the drug content by a validated method such as HPLC or UV spectroscopy.

The **in-vitro release** profile is measured by placing the carrier in a release medium at body temperature and sampling the medium over time to follow how much drug has been released, usually plotting cumulative percentage released against time. The practical difficulty unique to nanosystems is separating released drug from still-encapsulated drug without disturbing the system, which is handled by dialysis-membrane methods, sample-and-separate methods, or specialised flow-through cells. The resulting curve is fitted to the kinetic models of Chapter 2 (zero-order, first-order, Higuchi, Korsmeyer-Peppas) to diagnose the release mechanism, and it is the in-vitro half of any IVIVC built as in Chapter 10.

15.6 Permeability assays: PAMPA and Caco-2

For oral and other absorbed products, a key question is how well the drug will cross a biological barrier, and two in-vitro assays answer it at different levels of fidelity (Figure 15.2).

Two in-vitro permeability assays

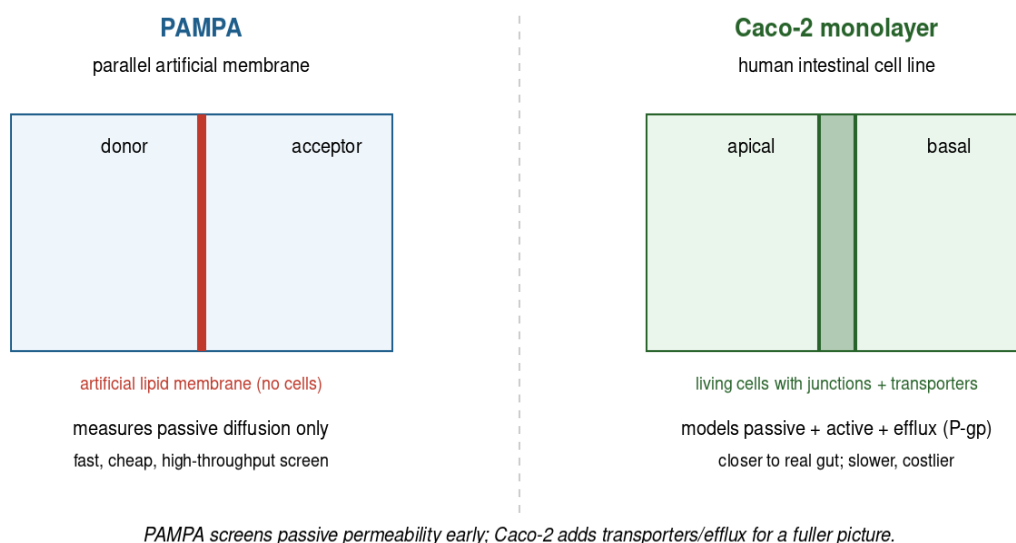


Figure 15.2. The PAMPA and Caco-2 permeability assays. PAMPA uses an artificial lipid membrane between a donor and an acceptor compartment and measures only passive transcellular diffusion; it is fast, cheap and suited to early high-throughput screening. The Caco-2 assay uses a monolayer of human intestinal-type cells with real tight junctions and transporters, so it captures passive, carrier-mediated and efflux transport, giving a more physiologically faithful but slower and more expensive result.

The **PAMPA (Parallel Artificial Membrane Permeability Assay)** uses an artificial lipid membrane separating a donor compartment from an acceptor compartment. The drug is placed in the donor side and its appearance in the acceptor side is measured over time. Because the membrane is a simple lipid film with no cells, PAMPA measures only **passive transcellular diffusion**, the ability of the drug to dissolve in and cross a lipid layer. It is fast, inexpensive and well suited to high-throughput screening of many candidate molecules early in development, but it cannot report anything about active transport or efflux because it has no transporters.

The **Caco-2 assay** uses a monolayer of Caco-2 cells, a human colorectal carcinoma cell line that, when grown to confluence, differentiates into cells resembling the intestinal epithelium, complete with tight junctions between the cells and the carrier and efflux transporters (including P-glycoprotein) found in the real gut. A drug is applied to the apical side and its appearance on the basolateral side is measured. Because the cells are alive and carry real transporters and junctions, the Caco-2 monolayer captures passive diffusion, carrier-mediated uptake, paracellular transport and active efflux all at once, making it a far more physiologically faithful model of intestinal absorption, and a standard tool in BCS-related permeability classification. Its cost is that it is slower, more expensive and more variable, requiring weeks to grow the monolayer and careful standardisation. The practical division of labour is therefore that PAMPA screens passive permeability quickly and cheaply across many compounds, while Caco-2 gives the fuller, transporter-aware picture for the candidates that progress.

Quick-recall box — characterization toolkit Size + PDI: DLS (Brownian motion → hydrodynamic diameter; low PDI = uniform). Surface charge/stability: zeta potential (large magnitude, ~beyond ± 30 mV, resists aggregation; PEGylated

systems stable even at low ζ). Morphology: SEM (surface, 3D) vs TEM (internal structure, higher resolution). Loading: drug loading (% of carrier) and encapsulation efficiency (% of input drug entrapped). Release: cumulative % vs time, fitted to kinetic models. Permeability: PAMPA (passive only, fast screen) vs Caco-2 (passive + transporters + efflux, physiological).

Chapter summary

The critical quality attributes named in Quality by Design must be measured by reliable methods, because they govern how a nanosystem behaves in the body. Dynamic light scattering infers particle size from the rate at which scattered-light intensity fluctuates due to Brownian motion, reporting the hydrodynamic diameter and, through the polydispersity index, the uniformity of the population. Zeta potential measures the effective surface charge at the slipping plane and predicts colloidal stability, with a large magnitude indicating strong repulsion and resistance to aggregation, though sterically stabilised PEGylated systems remain stable even at low values. Electron microscopy supplies the picture that light scattering cannot: SEM reveals surface topography in three dimensions while TEM resolves internal structure at higher resolution, both on dried samples. Drug loading and encapsulation efficiency quantify how much drug the carrier holds and how efficiently it was entrapped, measured after separating free from bound drug, while the in-vitro release profile follows cumulative release over time and is fitted to kinetic models and used to build IVIVCs. For absorption, the PAMPA assay screens passive lipid permeability quickly and cheaply but ignores transporters, whereas the Caco-2 monolayer of human intestinal-type cells captures passive, carrier-mediated and efflux transport for a physiologically faithful, if slower and costlier, assessment. Together these methods form the analytical backbone that turns the abstract critical quality attributes into measured, controllable numbers.

Self-assessment

Short-answer (2 marks each)

1. What does DLS measure, and what physical phenomenon underlies it?
2. Define the polydispersity index and state what a low value indicates.
3. How does zeta potential predict colloidal stability, and why can a PEGylated particle be stable at low zeta potential?
4. Differentiate SEM and TEM in what each reveals.
5. Distinguish drug loading from encapsulation efficiency.

Long-answer (5–10 marks each)

6. Describe how particle size and zeta potential are measured and why each is a critical quality attribute, with a diagram.
7. Explain how drug loading, encapsulation efficiency and in-vitro release are determined for a nanoparticle, and how the release data are interpreted.
8. Compare the PAMPA and Caco-2 permeability assays in principle, what each captures, and their respective roles in development.

Multiple-choice

9. A zeta potential of +35 mV suggests the suspension is: (a) prone to rapid aggregation (b) electrostatically well stabilised (c) uncharged (d) degraded

10. The assay that captures active efflux by P-glycoprotein is: (a) PAMPA (b) DLS (c) Caco-2 (d) zeta potential
11. Internal structure of a liposome bilayer is best seen by: (a) DLS (b) SEM (c) TEM (d) PAMPA

Answer pointers. 1. Hydrodynamic size (and PDI) from fluctuation of scattered light; Brownian motion. 2. Width of size distribution; low = narrow/uniform/near-monodisperse. 3. Like-charged particles repel and resist aggregation; PEG gives steric stabilisation independent of charge. 4. SEM = surface topography (3D); TEM = internal structure at higher resolution. 5. Loading = drug per weight of carrier (%); encapsulation efficiency = fraction of input drug entrapped (%). 6. Per Figure 15.1. 7. Per Section 15.5. 8. Per Figure 15.2 and Section 15.6. 9. (b). 10. (c). 11. (c).

References

1. Clogston JD, Patri AK. Zeta potential measurement. *Methods Mol Biol.* 2011;697:63–70.
2. Bhattacharjee S. DLS and zeta potential — what they are and what they are not? *J Control Release.* 2016;235:337–51.
3. Hubatsch I, Ragnarsson EGE, Artursson P. Determination of drug permeability and prediction of drug absorption in Caco-2 monolayers. *Nat Protoc.* 2007;2(9):2111–9.
4. Torchilin VP, editor. *Nanoparticulates as Drug Carriers.* London: Imperial College Press; 2006.

Chapter 16 - Non-Clinical Evaluation: PK/PD, Biodistribution and Toxicology

Learning outcomes

After working through this chapter the student will be able to:

1. Explain the purpose of non-clinical evaluation and how it differs for a novel delivery system from a conventional drug.
2. Describe how a delivery system alters the pharmacokinetics of its drug, and the basis of PK/PD modelling.
3. Explain biodistribution, the role of the mononuclear phagocyte system, opsonisation and the stealth effect of PEGylation.
4. Describe the principal toxicology studies required before first-in-human use and the carrier-specific toxicities of nanosystems.
5. Explain why the delivery system itself, not only the drug, must be evaluated for safety.

16.1 What non-clinical evaluation is for

Before any new medicine is given to a human, it must be studied in the laboratory and in animals to establish, as far as possible, that it is reasonably safe and to predict how it will behave. This is **non-clinical (preclinical) evaluation**, and it has three connected aims (Figure 16.1): to describe the **pharmacokinetics and pharmacodynamics** (what the body does to the drug and what the drug does to the body), to determine the **biodistribution** (where the drug and its carrier actually go), and to characterise the **toxicology** (what harm the product can do and at what dose). The combined result supports the choice of a safe starting dose for the first clinical trial.

Non-clinical evaluation before first-in-human

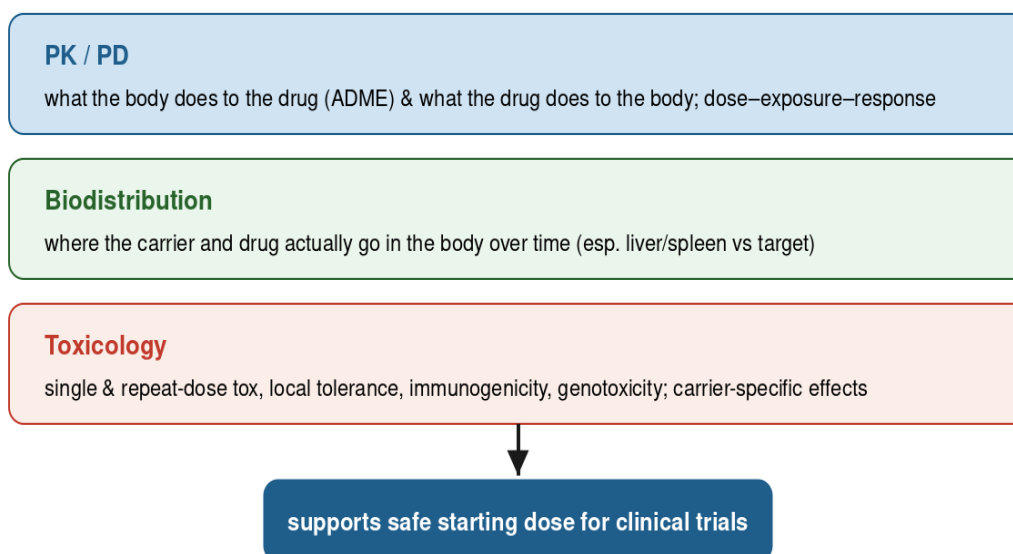


Figure 16.1. The three connected domains of non-clinical evaluation. Pharmacokinetics and pharmacodynamics describe what the body does to the drug and what the drug does to the body; biodistribution describes where the carrier and drug actually travel over time; and toxicology characterises the harm the product can cause and at what exposure. Together they support the selection of a safe starting dose for the first human study.

For a novel delivery system there is a crucial twist that runs through the whole chapter: the carrier changes everything. The same drug molecule, packaged in a liposome or a nanoparticle, has different pharmacokinetics, a different biodistribution and a different toxicity profile from the free drug, because the carrier, not the molecule, now controls where the drug goes and how fast it is released. Non-clinical data generated on the free drug therefore cannot simply be carried over; the delivery system must be evaluated as the new entity it is. This is the central principle of the chapter.

16.2 How a delivery system reshapes pharmacokinetics

The four classical pharmacokinetic processes, absorption, distribution, metabolism and excretion (ADME), are all altered when a drug is placed in a carrier. A long-circulating liposome changes distribution by keeping the drug in the central blood compartment for far longer than the free drug, which is cleared rapidly. A depot changes the absorption profile by releasing slowly over weeks. A nanoparticle may protect its cargo from the metabolising enzymes that would quickly destroy the free drug. The practical effect is that the plasma concentration–time curve of the encapsulated drug looks nothing like that of the free drug, and it is the curve of the encapsulated form that must be measured.

A subtlety unique to carriers is that one must often distinguish **encapsulated drug from released (free) drug** in the plasma, because only the released drug is usually pharmacologically active and only the released drug distributes freely to tissues. Total drug measured in plasma may be mostly still inside circulating carriers. Sound pharmacokinetic analysis of a nanomedicine therefore tries to measure both the total and the free fractions, which is analytically demanding.

PK/PD modelling links the measured pharmacokinetics (the exposure, typically the plasma concentration over time) to the pharmacodynamics (the resulting effect, whether therapeutic or toxic) through a mathematical model. The aim is to understand the dose–exposure–response relationship: how a given dose produces a given exposure, and how that exposure produces a given effect. A good PK/PD model lets investigators predict the effect of doses not yet tested, choose rational doses and schedules for clinical trials, and understand the therapeutic window in quantitative terms. For a delivery system whose whole purpose is to reshape the exposure curve, PK/PD modelling is how one confirms that the reshaped exposure actually translates into the intended improvement in effect or safety.

16.3 Biodistribution and the fate of the carrier

Biodistribution is the study of where the drug and its carrier travel in the body after administration and how those amounts change over time, usually expressed as the percentage of the injected dose found in each organ at successive time points (Figure 16.2). It is measured by sampling tissues, or non-invasively by imaging a radiolabelled or fluorescently labelled carrier. For a nanomedicine, biodistribution is often the single most informative non-clinical study, because the entire rationale of targeting (Chapter 2) is a claim about biodistribution, namely that the carrier will go to the diseased site and not elsewhere.

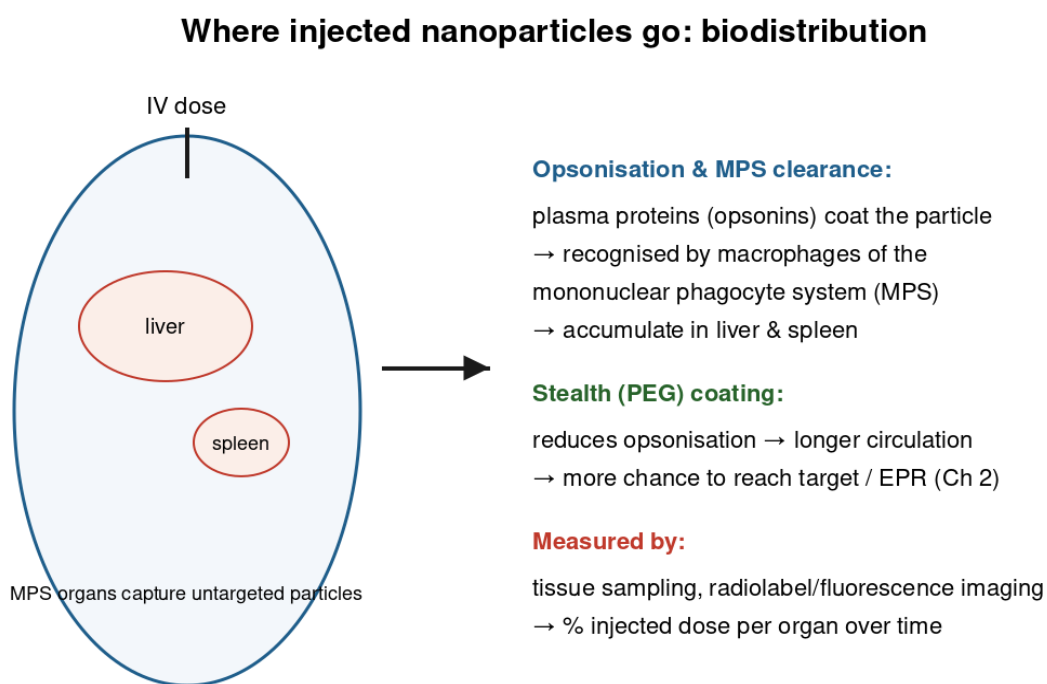


Figure 16.2. The biodistribution of intravenously injected nanoparticles. Plasma proteins called opsonins adsorb onto an uncoated particle (opsonisation), marking it for recognition by the macrophages of the mononuclear phagocyte system, so that the particle accumulates in the liver and spleen. A stealth coating of poly(ethylene glycol) reduces opsonisation, prolonging circulation and giving the carrier more opportunity to reach its target, including by the EPR effect. Biodistribution is quantified by tissue sampling or by imaging a labelled carrier as percentage of injected dose per organ over time.

The dominant fact of nanoparticle biodistribution is capture by the **mononuclear phagocyte system (MPS)**, the network of macrophages concentrated in the liver and spleen. The

mechanism is **opsonisation**: when an uncoated particle enters the blood, plasma proteins called opsonins rapidly adsorb onto its surface, and these adsorbed proteins are recognised by the MPS macrophages, which engulf the particle and remove it to the liver and spleen. The consequence is that untargeted nanoparticles tend to accumulate in these organs rather than reaching their intended site, which is both a barrier to targeting and a source of potential liver and spleen toxicity.

The standard countermeasure, introduced in Chapter 2 and recurring throughout the book, is the **stealth coating** of **poly(ethylene glycol) (PEG)**. The hydrophilic PEG layer forms a barrier that reduces opsonin adsorption, so the particle is recognised more slowly by the MPS and circulates for longer. Longer circulation is what gives a passively targeted particle the time it needs to accumulate at a tumour by the EPR effect. Biodistribution studies are the way this stealth effect is demonstrated: the PEGylated carrier shows a longer blood half-life and lower liver and spleen uptake than its uncoated counterpart. The non-clinical assessment of any targeted nanomedicine therefore lives or dies on its biodistribution data.

16.4 Toxicology

Toxicology establishes what harm the product can cause and at what exposure. The standard non-clinical toxicology programme includes **single-dose (acute) toxicity** to find the immediately harmful dose, **repeat-dose toxicity** over the intended duration of use to find effects that build up, **local tolerance** at the site of administration (important for injected depots and skin-piercing systems), **genotoxicity** to detect damage to DNA, and, for longer-term products, **carcinogenicity** and **reproductive toxicity**. These studies define the no-observed-adverse-effect level that anchors the safe starting dose for humans.

For novel delivery systems, toxicology must look beyond the drug to the carrier and its breakdown products, and several **carrier-specific toxicities** are distinctive. The carrier material itself may be toxic or may accumulate: this is exactly the unresolved clearance concern raised for the inorganic carriers of Chapter 5, where a non-degradable particle that lodges in the liver or spleen indefinitely poses a long-term risk that a biodegradable PLGA system does not. Nanoparticles can provoke **immunogenicity and immunotoxicity**, including complement activation, because the immune system is built to detect particles of this size; PEGylated systems have, somewhat unexpectedly, been associated with anti-PEG immune responses and infusion reactions in some cases, which is an active area of study. The carrier can alter the toxicity of the drug by changing where it goes, sometimes reducing it, as liposomal amphotericin B reduces kidney toxicity (Chapter 1), but in principle sometimes concentrating drug in an organ where the free drug never reached. Degradation products must also be assessed; the acidic microenvironment of an eroding PLGA depot, noted in Chapter 3, is a local toxicity and stability consideration. The lesson is consistent: the delivery system is a new entity with its own safety profile, and evaluating only the drug would miss the risks that the carrier introduces.

16.5 Why the whole system, and not the drug alone, is evaluated

Pulling these threads together, the recurring message of non-clinical evaluation for NDDS is that the delivery system transforms the drug into something pharmacokinetically, distributionally and toxicologically new. A drug with decades of safe clinical use, repackaged into a nanoparticle, can circulate longer, concentrate in different organs, engage the immune system differently and degrade into different products. None of these behaviours can be predicted from the free-drug data alone. This is why regulators require the delivery system to be characterised and tested as a distinct product, and why the analytical methods of Chapter

15 and the non-clinical studies of this chapter precede any human exposure. It is also why this evaluation feeds directly into the regulatory pathways of Chapter 17, which decide what evidence package a modified or novel delivery product must present to reach the market.

Quick-recall box — non-clinical evaluation Three domains: PK/PD (ADME + dose–exposure–response; distinguish encapsulated vs free drug), biodistribution (where it goes, % injected dose/organ), toxicology (acute, repeat-dose, local tolerance, genotoxicity, immunotoxicity). Key biodistribution fact: opsonisation → MPS capture in liver/spleen; PEG stealth coating reduces opsonisation → longer circulation → EPR targeting. Carrier-specific toxicities: non-degradable carrier accumulation (cf. inorganic carriers), immunogenicity/complement, anti-PEG responses, altered drug distribution. Core principle: the system is a new entity; evaluate the carrier, not only the drug.

Chapter summary

Non-clinical evaluation establishes a product's likely safety and behaviour before first human use through three connected investigations: pharmacokinetics and pharmacodynamics, biodistribution, and toxicology, which together justify a safe starting clinical dose. For a delivery system the governing principle is that the carrier reshapes the drug into a new entity, so free-drug data cannot be carried over. The carrier alters all of ADME, prolonging circulation, slowing absorption from a depot, or protecting the drug from metabolism, so the plasma curve of the encapsulated drug differs entirely from the free drug, and analysis must often distinguish encapsulated from released drug since only the free fraction is active; PK/PD modelling then links the reshaped exposure to effect to confirm the intended benefit. Biodistribution is frequently the most informative study for a nanomedicine because targeting is a claim about where the carrier goes; its dominant feature is capture by the mononuclear phagocyte system after opsonins adsorb onto the particle, sending untargeted carriers to the liver and spleen, a fate countered by PEG stealth coating that reduces opsonisation and prolongs circulation to enable EPR targeting. Toxicology extends beyond the drug to the carrier, assessing acute and repeat-dose toxicity, local tolerance, genotoxicity and immunotoxicity, and confronting carrier-specific risks such as accumulation of non-degradable materials, complement activation and anti-PEG responses, and carrier-altered drug distribution. The consistent conclusion, which drives the regulatory requirements of the next chapter, is that the delivery system must be evaluated as a distinct product, not merely as a familiar drug in a new wrapper.

Self-assessment

Short-answer (2 marks each)

1. Name the three domains of non-clinical evaluation.
2. Why can free-drug pharmacokinetic data not be carried over to its encapsulated form?
3. Define opsonisation and state its consequence for nanoparticle biodistribution.
4. How does PEGylation produce a stealth effect?
5. Give two carrier-specific toxicities that must be assessed for a nanomedicine.

Long-answer (5–10 marks each)

6. Explain how a delivery system alters the pharmacokinetics of its drug, and the purpose of PK/PD modelling.

7. Describe biodistribution, the role of the mononuclear phagocyte system and opsonisation, and how PEGylation and biodistribution studies relate, with a diagram.
8. Discuss why the delivery system, not only the drug, must undergo toxicological evaluation, with examples of carrier-specific risks.

Multiple-choice

9. Untargeted intravenous nanoparticles tend to accumulate in the: (a) brain (b) liver and spleen (c) skin (d) bone
10. PEG coating prolongs circulation primarily by reducing: (a) renal filtration (b) opsonisation and MPS uptake (c) drug release (d) particle size
11. PK/PD modelling relates: (a) two doses (b) exposure to effect (c) size to charge (d) two organs

Answer pointers. 1. PK/PD, biodistribution, toxicology. 2. The carrier changes ADME, distribution and release, making the encapsulated form a new entity. 3. Plasma opsonins adsorb onto the particle, marking it for MPS macrophage uptake into liver/spleen. 4. Hydrophilic PEG layer reduces opsonin adsorption, slowing MPS recognition and prolonging circulation. 5. Carrier accumulation/non-degradability, immunogenicity/complement, anti-PEG responses, altered drug distribution (any two). 6. Per Section 16.2. 7. Per Figure 16.1. 8. Per Section 16.4. 9. (b). 10. (b). 11. (b).

References

1. Wilhelm S, Tavares AJ, Dai Q, et al. Analysis of nanoparticle delivery to tumours. *Nat Rev Mater.* 2016;1:16014.
2. Moghimi SM, Hunter AC, Andresen TL. Factors controlling nanoparticle pharmacokinetics: an integrated analysis and perspective. *Annu Rev Pharmacol Toxicol.* 2012;52:481–503.
3. Dobrovolskaia MA, McNeil SE. Immunological properties of engineered nanomaterials. *Nat Nanotechnol.* 2007;2(8):469–78.
4. Dash AK, Cudworth GC. *Therapeutic Proteins and Peptides.* (Reference text for protein/peptide delivery evaluation.) 2014.
5. ICH. M3(R2) Nonclinical Safety Studies for the Conduct of Human Clinical Trials. Geneva: ICH; 2009.

Chapter 17 - Regulatory Pathways for Novel Delivery Systems

Learning outcomes

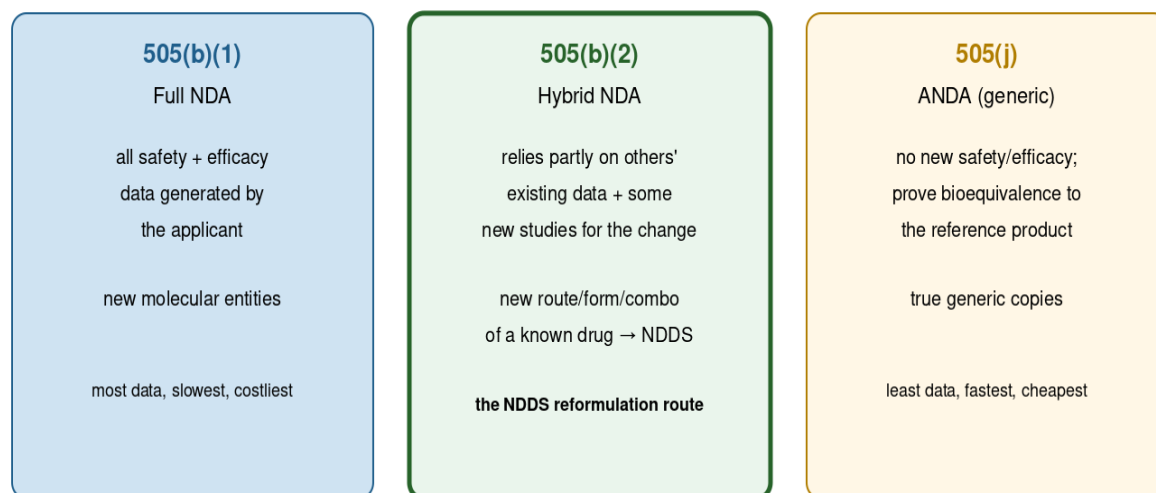
After working through this chapter the student will be able to:

1. Distinguish the three US marketing-application pathways, 505(b)(1), 505(b)(2) and 505(j) (ANDA), by the data each requires.
2. Explain why the 505(b)(2) pathway is the natural route for many novel delivery reformulations of known drugs.
3. Explain what makes a product a “complex generic” and why bioequivalence is difficult to establish for such products.
4. Summarise the ICH quality framework (Q8–Q12) and its role in modern regulatory submissions.
5. Outline the EMA’s role and note the Indian (CDSCO) regulatory context.

17.1 Three doors to the US market

A delivery system is only a product once a regulator approves it for marketing, and the route chosen shapes the entire development programme: what studies must be done, how long approval takes and how much it costs. In the United States the framework is built around three pathways, named after sections of the Federal Food, Drug, and Cosmetic Act (Figure 17.1).¹

Three US marketing-application pathways



505(b)(2) sits between a full NDA and a generic — built for modified versions of known drugs.

Figure 17.1. The three US marketing-application pathways. A 505(b)(1) full New Drug Application requires the applicant to generate all safety and efficacy data and is used for new molecular entities. A 505(j) Abbreviated New Drug Application (ANDA) requires no new safety or efficacy data and instead demonstrates bioequivalence to an approved reference product, the route for true generics. The 505(b)(2) application sits between them: it relies

partly on existing data the applicant did not generate, supplemented by new studies covering the specific change, and is the natural route for a novel delivery reformulation of a known drug.

The **505(b)(1)** application is the full New Drug Application (NDA). The applicant generates the entire body of safety and efficacy evidence, from non-clinical studies through full clinical trials, and submits it. This is the route for a genuinely new molecular entity and is the most demanding, slowest and most expensive.

The **505(j)** application is the Abbreviated New Drug Application (ANDA), the generic route. A generic manufacturer does not repeat the safety and efficacy studies; instead it demonstrates that its product is **bioequivalent** to an already-approved reference product, that it delivers the same drug to the bloodstream at essentially the same rate and extent. Because the reference product has already proven safety and efficacy, bioequivalence is sufficient. This is the cheapest and fastest route, and it is what makes generic medicines affordable.

The **505(b)(2)** application sits between these two and is the one most relevant to novel delivery systems. It is a hybrid NDA that allows the applicant to rely in part on data it did not itself generate, typically the existing safety and efficacy data for the already-approved drug, while supplying new studies only for the specific aspects that are different. It is, in effect, a way to get credit for what is already known about the drug while proving only what is new about the delivery.

17.2 Why 505(b)(2) fits novel delivery

The 505(b)(2) pathway is, in a sense, purpose-built for the entire enterprise of this book. Recall from Chapter 1 that a novel delivery system most often takes a known, approved drug and changes how it is delivered: a new route (a transdermal patch of a previously oral drug), a new dosage form (a depot injection of a daily oral drug), a new combination, or a modified-release version. The drug itself is not new, so its safety and efficacy are already established and documented. Repeating the full clinical programme of a new molecular entity would be wasteful and unethical, exposing patients to trials whose core question, does this drug work and is it safe, has already been answered. But the product is not a simple generic copy either, so the ANDA route does not fit.

The 505(b)(2) pathway resolves this exactly. The applicant relies on the established record of the known drug for its fundamental safety and efficacy, and conducts new studies only to address what the reformulation changes, the pharmacokinetics of the new delivery, its local tolerance, its specific safety considerations. This dramatically reduces the development burden compared with a full NDA while still providing the regulator with the evidence needed for the genuinely new aspects. It is the principal reason that reformulating an established drug into a superior delivery system is commercially attractive, the regulatory-and-commercial driver named in Chapter 1, and it is the route by which many of the depot, transdermal and reformulated products in this book reached the market.

17.3 Complex generics

The flip side of the regulatory picture concerns what happens when someone wants to make a generic copy of an already-marketed novel delivery system. Here a real difficulty arises, captured by the term **complex generic** (Figure 17.2).

Why NDDS copies are "complex generics"

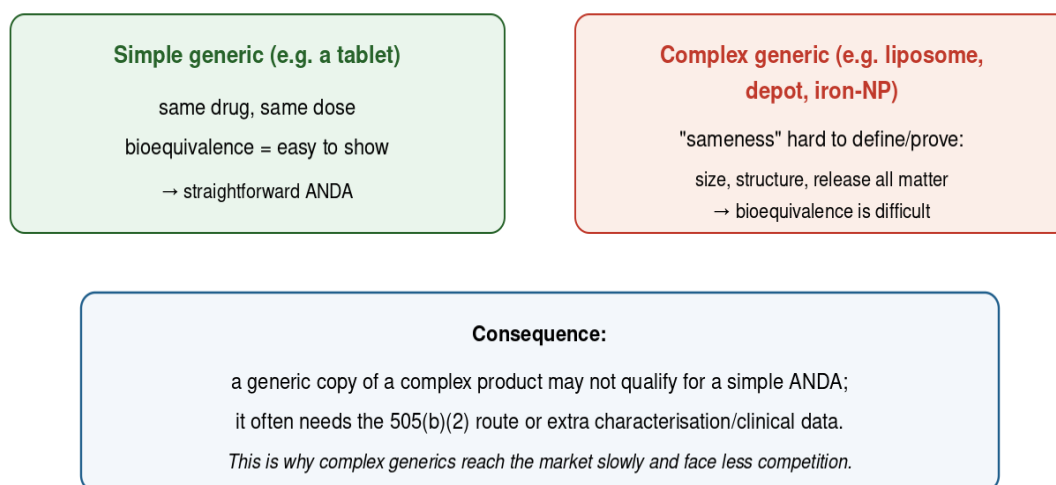


Figure 17.2. *Why copies of novel delivery systems are complex generics. For a simple tablet, bioequivalence is straightforward to demonstrate. For a complex system such as a liposome, depot or iron-nanoparticle injectable, performance depends on physical attributes, size, structure and release, that are difficult to define and match, so conventional bioequivalence may not establish sameness; the copy often needs extra characterisation or the 505(b)(2) route, and consequently reaches the market slowly and with little competition.*

For a simple product such as a conventional tablet, demonstrating bioequivalence is straightforward: the generic and the reference deliver the same drug to the blood at the same rate, which is readily measured. For a complex delivery system, this is far harder, because the product's behaviour depends on physical attributes that are difficult to define and to match. Consider a liposomal product: its performance depends on the liposome size and size distribution, the bilayer composition, the way the drug is loaded and the rate at which it is released, the surface coating, and more. A would-be generic must match the drug and dose and, beyond them, this whole constellation of physical attributes, and proving "sameness" between two complex particulate systems is scientifically demanding. The same applies to iron-carbohydrate nanoparticle injectables, to PLGA depot microspheres, and to other nanomedicines.

The consequence is significant. A complex product often cannot qualify for a simple ANDA, because conventional bioequivalence testing may not adequately demonstrate sameness; the generic applicant may need extensive additional physicochemical characterisation, and sometimes clinical data, often pushing the product toward the 505(b)(2) route instead. This is why complex generics reach the market slowly, why fewer competitors enter, and why prices for these products stay high long after the original patents expire. Regulators, aware of this, have invested in developing product-specific guidance and new bioequivalence methods to make approval of complex generics more feasible, an active area of regulatory science. The practical lesson for the student is that the very complexity that makes a delivery system therapeutically valuable also makes it hard to copy, which has lasting commercial consequences.

17.4 The ICH quality framework

Whatever the marketing pathway, the quality portion of a modern submission is built on the **ICH (International Council for Harmonisation)** quality guidelines, introduced in Chapter 14 and consolidated here.² ICH brings together regulators and industry from major regions to harmonise technical requirements, so that broadly the same scientific quality expectations apply across the United States, Europe, Japan and the many regions, including India, that adopt ICH standards.

The quality guidelines most relevant to delivery systems are: **Q8 (Pharmaceutical Development)**, which sets out Quality by Design, the QTPP, critical quality attributes, design space and control strategy; **Q9 (Quality Risk Management)**, which provides the formal risk-assessment methods; **Q10 (Pharmaceutical Quality System)**, which describes the lifecycle quality system; **Q11**, which extends these principles to the development and manufacture of the drug substance; and **Q12**, which addresses lifecycle management and how changes are handled after approval. A submission for a novel delivery system uses this framework to demonstrate that the product's critical quality attributes are understood and controlled, which is exactly the QbD story of Chapter 14 told to a regulator. Adherence to ICH quality standards is therefore not a separate exercise but the formal expression of the development science the whole unit has described.

17.5 The EMA and the Indian context

Beyond the United States, the **European Medicines Agency (EMA)** is the central regulator for the European Union. Its scientific standards align closely with ICH, and it operates a centralised procedure through which a single marketing authorisation can cover the whole EU, which is the usual route for innovative and biotechnology products. The EMA has been active in developing reflection papers and guidance specifically for nanomedicines and complex products, recognising the same characterisation challenges discussed above.

For the Indian student this material must be set in the home context. The primary Indian regulator is the **Central Drugs Standard Control Organisation (CDSCO)**, which oversees the approval of new drugs and clinical trials in India, with the Indian Pharmacopoeia as the national quality standard. India broadly adopts ICH quality principles, and Indian manufacturers, who are major suppliers of generics and increasingly of complex generics to the world, routinely prepare submissions to US, EU and Indian requirements together. The detailed Indian regulatory provisions for new drugs and clinical trials are set out in current national rules, and a developer working in India aligns the scientific package described in this unit, QbD-based development, full characterisation, non-clinical evaluation, with both CDSCO requirements and the requirements of any export market.

Quick-recall box — regulatory routes US pathways: 505(b)(1) full NDA (new molecule, all data); 505(b)(2) hybrid (relies on others' data + new studies for the change — the NDDS reformulation route); 505(j) ANDA (generic, prove bioequivalence only). Complex generic: a copy of a complex system (liposome, depot, iron-NP) where bioequivalence/sameness is hard to prove, so it needs extra characterisation or 505(b)(2); reaches market slowly, less competition. ICH quality: Q8 development/QbD, Q9 risk, Q10 quality system, Q11 drug substance, Q12 lifecycle. EU = EMA (ICH-aligned, centralised); India = CDSCO + Indian Pharmacopoeia.

Chapter summary

The regulatory route chosen for a delivery system determines what evidence must be generated and how fast and cheaply the product can reach patients. The United States offers three pathways: the 505(b)(1) full New Drug Application, requiring the applicant to generate all safety and efficacy data, for new molecular entities; the 505(j) Abbreviated New Drug Application, requiring only a demonstration of bioequivalence to an approved reference, for true generics; and the 505(b)(2) hybrid application, which lets an applicant rely on existing data for the known drug while conducting new studies only for what the reformulation changes. The 505(b)(2) route is the natural home of novel delivery, because such products take an established drug and change its delivery, so the drug's safety and efficacy are already known and only the delivery is new, which both reduces the development burden and underlies the commercial attractiveness of reformulation. The mirror-image problem is the complex generic: copying an already-marketed complex system such as a liposome or a depot is scientifically hard because performance depends on physical attributes like size, structure and release that are difficult to define and match, so conventional bioequivalence may not suffice, additional characterisation or the 505(b)(2) route is often required, and these products consequently reach the market slowly and retain high prices. Across all pathways the quality package rests on the ICH guidelines Q8 to Q12, which formalise the Quality by Design development science of this unit, while the EMA serves the European Union with ICH-aligned, nanomedicine-aware standards and India's CDSCO, working to the Indian Pharmacopoeia and broadly adopting ICH principles, governs the home market in which Indian manufacturers prepare submissions for domestic and export use together.

Self-assessment

Short-answer (2 marks each)

1. Distinguish 505(b)(1), 505(b)(2) and 505(j) by the data each requires.
2. Why is the 505(b)(2) route suited to a transdermal reformulation of an oral drug?
3. What is a complex generic, and why is its bioequivalence hard to prove?
4. List the ICH quality guidelines Q8–Q12 with their subjects.
5. Name the primary Indian drug regulator and the national quality standard.

Long-answer (5–10 marks each)

6. Compare the three US marketing pathways with a diagram, and explain why 505(b)(2) is the principal route for novel delivery reformulations.
7. Explain the concept of a complex generic, why such products are difficult to copy, and the commercial consequences, with examples.
8. Describe the ICH quality framework and the roles of the EMA and CDSCO, and explain how the development science of this unit maps onto a regulatory submission.

Multiple-choice

9. A depot injection of a previously oral approved drug is most likely submitted via: (a) 505(b)(1) (b) 505(b)(2) (c) 505(j) ANDA (d) no application
10. An ANDA (505(j)) requires the applicant to demonstrate: (a) full new clinical trials (b) bioequivalence to the reference (c) a new design space (d) nothing
11. ICH Q8 covers: (a) quality risk management (b) pharmaceutical development/QbD (c) the quality system (d) lifecycle management

Answer pointers. 1. 505(b)(1) = full NDA, all data; 505(b)(2) = hybrid, relies on existing data + new studies for the change; 505(j) = ANDA, bioequivalence only. 2. The drug's safety/efficacy are already established; only the new delivery needs new data. 3. A generic of a complex system; performance depends on hard-to-match physical attributes, so bioequivalence/sameness is difficult. 4. Q8 development/QbD, Q9 risk, Q10 quality system, Q11 drug substance, Q12 lifecycle. 5. CDSCO; Indian Pharmacopoeia. 6. Per Figure 17.1 and Section 17.2. 7. Per Section 17.3. 8. Per Sections 17.4–17.5. 9. (b). 10. (b). 11. (b).

References

1. US Food and Drug Administration. Applications Covered by Section 505(b)(2): Guidance for Industry. Rockville (MD): FDA; 1999 (draft) and subsequent guidance.
2. ICH. Quality Guidelines Q8(R2), Q9, Q10, Q11, Q12. Geneva: International Council for Harmonisation; 2008–2020.
3. Husaarts L, Mühlebach S, Shah VP, et al. Equivalence of complex drug products: scientific and regulatory challenges. *Ann N Y Acad Sci.* 2017;1407(1):39–49.
4. European Medicines Agency. Reflection papers on nanomedicines and on data requirements for intravenous liposomal products. Amsterdam: EMA; various.
5. Central Drugs Standard Control Organisation. New Drugs and Clinical Trials Rules. New Delhi: CDSCO; current edition.

Chapter 18 - Case Studies in Novel Drug Delivery

Learning outcomes

After working through this chapter the student will be able to:

1. Analyse a marketed delivery system by naming the specific limitation it answers and the strategy it uses, applying the Chapter 1 framework.
2. Explain the design and benefit of liposomal, lipid-nanoparticle, depot, transdermal, ocular, colon-targeted, pulmonary and chronotherapeutic products with verified facts.
3. Distinguish commercial successes from instructive failures and explain why each turned out as it did.
4. Describe the current research status of dendrimer-based cancer delivery without overstating it.
5. Synthesise the recurring principle that the molecule is unchanged and the delivery does the work.

18.1 How to read a case study

Each product below is presented in the same way: the **limitation** of the conventional form, the **strategy** the delivery system uses, and the **outcome**. This is the framework of Chapter 1, and the point of the chapter is to show that every successful delivery product, however sophisticated, is an answer to a precisely named problem (Figure 18.1). Where a product failed, the failure too is informative.

The case studies as limitation → strategy → product

Limitation of the drug	Delivery strategy	Product
nephrotoxic free drug	liposome alters distribution	AmBisome (ampho B)
mRNA degraded, can't enter cell	ionisable lipid LNP	Comirnaty / Spikevax
daily antipsychotic, poor adherence	PLGA microsphere depot	Risperdal Consta
oral hormone, first-pass loss	transdermal patch	estradiol HRT patch
glaucoma drops washed away	rate-controlled ocular insert	Ocusert (pilocarpine)
IBD needs colonic delivery	pH / microbial colon targeting	mesalazine, sulfasalazine
injectable insulin burden	pulmonary dry powder	Afrezza (Exubera failed)
non-specific cancer toxicity	dendrimer / nanocarrier targeting	largely research stage
morning BP surge (rhythm)	chronotherapeutic lag + pulse	bedtime-dosed CR systems

Figure 18.1. The case studies arranged as limitation → strategy → product. In every case the drug molecule is unchanged; the delivery system answers a specific limitation. Reading any marketed delivery product this way, identifying the limitation it solves, is the analytical skill this chapter trains.

18.2 Liposomal amphotericin B (AmBisome)

The limitation is toxicity from non-specific distribution. Amphotericin B is a highly effective antifungal but severely damages the kidney, because the free drug binds cholesterol in human cell membranes as well as ergosterol in fungal membranes. The strategy is to enclose the drug in a liposome, which changes its biodistribution so that less free drug reaches the kidney. The outcome is markedly reduced nephrotoxicity at therapeutic doses, which allows higher and safer dosing in serious systemic fungal infections; AmBisome is widely used in India and internationally.¹ The molecule did not change; the liposome did the work, the founding example of the whole subject.

18.3 mRNA–lipid nanoparticle vaccines (Comirnaty, Spikevax)

The limitation is severe: messenger RNA is a large, fragile, negatively charged molecule that is destroyed within minutes by nucleases and cannot, on its own, cross a cell membrane to reach the ribosomes where it must be translated. Without a delivery system there is no mRNA medicine at all. The strategy is the lipid nanoparticle of Chapter 4, built from four lipids (Figure 18.2): an **ionisable lipid** that is neutral in the blood but becomes positively charged in the acidic endosome to release the RNA into the cytoplasm, a **PEG-lipid** controlling size and circulation, and a **helper phospholipid** and **cholesterol** for structure.² The outcome was the rapid development and global deployment of the mRNA COVID-19 vaccines Comirnaty and Spikevax. This is the strongest modern demonstration that delivery can be the decisive enabling technology: the mRNA sequence is the active, but it is the LNP that turned an undeliverable molecule into one of the most consequential medicines of the era.

mRNA–lipid nanoparticle: four lipids, one job

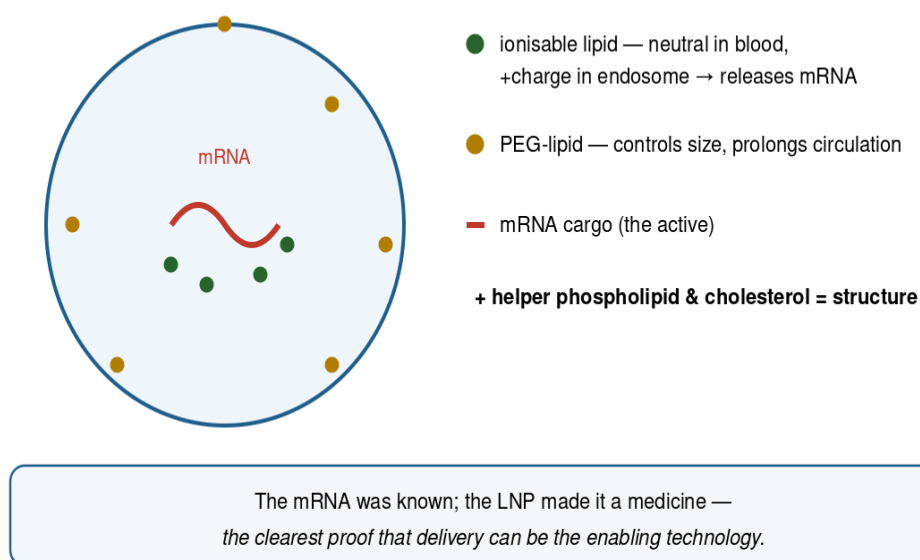


Figure 18.2. The mRNA–lipid nanoparticle. The ionisable lipid, neutral in the bloodstream, becomes positively charged in the acidic endosome and disrupts its membrane to release the mRNA into the cytoplasm; the PEG-lipid controls size and prolongs circulation; helper phospholipid and cholesterol provide structure. The mRNA was the known active, but the LNP made it a usable medicine.

18.4 Depot antipsychotic injection (Risperdal Consta)

The limitation is adherence. Schizophrenia is often treated with daily oral antipsychotics, but daily dosing is precisely what many patients with the illness struggle to maintain, and missed doses cause relapse. The strategy is the PLGA microsphere depot of Chapters 3 and 12: risperidone is dispersed in PLGA microspheres, injected, and released over weeks as the polymer hydrolyses. The outcome is a long-acting injection given every two weeks rather than a tablet taken every day, which removes the daily decision and improves adherence; the polymer clears to endogenous acids so nothing remains.³ The same logic, demonstrated also by leuprolide microspheres (Lupron Depot) for hormone-responsive conditions, makes long-acting injectables one of delivery's clearest clinical wins.

18.5 Transdermal hormone replacement (estradiol patch)

The limitation is first-pass metabolism and the peak–trough profile of oral hormone dosing. Oral estradiol is substantially metabolised by the gut and liver, requiring higher doses and producing fluctuating levels. The strategy is the transdermal patch of Chapter 11: estradiol is a small, lipophilic, potent molecule that satisfies the 500-Dalton criteria, so it crosses the skin and enters the systemic circulation directly, bypassing the liver. The outcome is steady hormone levels from a patch changed once or twice weekly, at a lower total dose, with the convenience and the easy reversibility (remove the patch) that transdermal delivery provides. Estradiol transdermal patches are a long-established example of route change solving a first-pass problem.

18.6 Ocular insert for glaucoma (Ocusert)

The limitation is the rapid loss of eye drops, which are drained and blinked away within minutes so that only a few percent of the dose is absorbed and frequent instillation is needed, a particular burden in a chronic disease such as glaucoma. The strategy is the rate-controlled ocular insert of Chapter 12: Ocusert is a membrane-controlled insert placed in the conjunctival sac that releases pilocarpine at a steady, near zero-order rate over about a week. The outcome is sustained drug delivery from a single weekly placement instead of multiple daily drops, and it stands as the textbook example of zero-order ocular delivery. It also illustrates a limitation of inserts: patient acceptability and the sensation of a foreign body, which constrained its commercial uptake.

18.7 Colon-targeted delivery for inflammatory bowel disease

The limitation is the need to deliver an anti-inflammatory drug to the colonic mucosa while sparing the rest of the body. The strategy is the colon-targeting of Chapter 8: pH-dependent methacrylate coatings, time-dependent and multi-matrix systems, and the microbial azo-bond prodrug approach. The outcome is products that release mesalazine (5-aminosalicylic acid) or budesonide in the colon for ulcerative colitis and Crohn's disease, including the classic sulfasalazine, in which colonic bacteria cleave an azo bond to liberate the active drug exactly where it is needed.⁴ This case shows targeting used not for a nanoparticle but for a conventional oral product, achieving site-specific delivery by reading the chemistry and microbiology of the gut.

18.8 Inhalable insulin: a success and a failure (Afrezza and Exubera)

The limitation is the burden of injected insulin and the wish for a needle-free route, exploiting the lung's vast absorptive surface (Chapter 9). This case is valuable because it contains both a

failure and a success, and the contrast is the lesson. **Exubera**, the first inhaled insulin, was marketed in 2006 but withdrawn in 2007; the science worked, but the product failed commercially because its inhaler device was large and awkward, dosing was confusing and explained in unfamiliar units, and uptake was poor.⁵ **Afrezza**, approved in 2014, uses a small, simple, breath-powered dry-powder inhaler delivering an ultra-rapid-acting insulin, and remains on the market.⁵ The outcome to retain is that a delivery system must succeed not only pharmacologically but as a usable, acceptable product: Exubera's physics were sound and its device was not, and the device killed it. Delivery is engineering, regulation and human factors together, not pharmacokinetics alone.

18.9 Dendrimers for targeted cancer therapy

The limitation is the non-specific toxicity of cancer chemotherapy. The strategy under investigation is the **dendrimer**, a precisely branched, tree-like polymer (developed further in Chapter 19) whose many surface groups can carry drug and targeting ligands, offering a highly defined, monodisperse nanocarrier. The honest outcome is that dendrimer-based targeted oncology remains largely at the research and early-development stage; dendrimers are a vigorous area of laboratory and preclinical study for drug and gene delivery, but they have not yet produced an established mainstream targeted cancer therapy, and concerns about the toxicity of some dendrimer surfaces remain under study. This case is included precisely to teach calibrated judgement: not every promising nanotechnology has yet become a product, and a good pharmaceutical scientist states the research stage accurately rather than overselling it.

18.10 Chronotherapeutic delivery in hypertension

The limitation is a mismatch between constant drug delivery and a disease that varies through the day. Blood pressure and cardiovascular risk rise in the early morning, yet a conventional once-daily tablet may not have its peak effect timed to that surge. The strategy is the chronotherapeutic system of Chapter 2: a dosage form taken at bedtime that holds the drug back through a programmed lag and then releases it (a delayed, pulsatile release) so that the peak drug level coincides with the early-morning blood-pressure rise. The outcome is antihypertensive therapy timed to the body's rhythm rather than delivered flat, aiming to control the morning surge when events are most likely. This case shows the third axis of Chapter 2, timing, applied to a real chronic disease, and rounds out the demonstration that delivery can control rate, place and also time.

18.11 The common thread

Read together, the cases make one point repeatedly. In every instance the active molecule is unchanged: amphotericin B, risperidone, estradiol, pilocarpine, mesalazine, insulin are all old, known drugs, and even the mRNA sequence is simply genetic information. What changed, and what created the therapeutic or commercial value, was the delivery. The successes answered a clearly named limitation with a matched strategy; the failures, like Exubera, usually failed not in their pharmacology but in their practicality. This is the discipline of the whole book reduced to a single habit of mind: name the limitation, match the strategy, and remember that a delivery system must work as a product, not only as a mechanism.

Quick-recall box — the nine cases AmBisome: toxicity → liposome alters distribution → less nephrotoxicity. mRNA-LNP: undeliverable RNA → ionisable-lipid LNP → COVID vaccines. Risperdal Consta: adherence → PLGA depot → 2-

weekly injection. Estradiol patch: first-pass → transdermal → steady levels. Ocusert: drop loss → rate-controlled insert → weekly pilocarpine. IBD: site delivery → colon targeting → mesalazine/sulfasalazine. Inhaled insulin: needle burden → pulmonary DPI → Afrezza succeeded, Exubera failed (bad device). Dendrimers: cancer toxicity → targeted dendrimer → mostly research. Chronotherapy: morning BP surge → bedtime lag+pulse. Thread: molecule unchanged, delivery does the work; products must also be usable.

Chapter summary

The case studies demonstrate, product by product, that a successful delivery system is the answer to a named limitation while the drug molecule itself stays the same. Liposomal amphotericin B reduces nephrotoxicity by changing the drug's distribution; mRNA-lipid nanoparticles made an otherwise undeliverable, fragile molecule into the COVID-19 vaccines through an ionisable-lipid design, the strongest modern proof that delivery can be the enabling technology; the risperidone PLGA microsphere depot converts a daily tablet into a fortnightly injection that solves an adherence problem; the estradiol transdermal patch bypasses first-pass metabolism to give steady hormone levels; the Ocusert insert delivers pilocarpine at a zero-order rate for glaucoma, while illustrating the acceptability limits of inserts; colon-targeted systems deliver mesalazine and the azo-prodrug sulfasalazine to the inflamed colon; and chronotherapeutic antihypertensives time release to the early-morning blood-pressure surge. The inhaled-insulin pair teaches the sharpest practical lesson: Exubera failed and Afrezza succeeded not on pharmacology but on device usability, so a delivery system must work as a usable product and not merely as a sound mechanism. Dendrimers for cancer, by contrast, are presented honestly as still largely investigational, training the judgement to state a technology's true stage. The unifying habit, naming the limitation and matching the strategy, is the analytical core of the discipline.

Self-assessment

Short-answer (2 marks each)

1. What limitation does liposomal amphotericin B overcome, and how?
2. Why is a delivery system essential for an mRNA vaccine?
3. How does a depot antipsychotic improve adherence?
4. Why did Exubera fail while Afrezza succeeded?
5. State the honest current status of dendrimer-based targeted cancer therapy.

Long-answer (5–10 marks each)

6. Analyse any three case studies using the limitation → strategy → outcome framework.
7. Explain the mRNA-LNP vaccine as a delivery achievement, describing the role of each lipid component, with a diagram.
8. “A delivery system must work as a product, not only as a mechanism.” Discuss using the inhaled-insulin case and one other.

Multiple-choice

9. The ionisable lipid in an mRNA-LNP becomes charged: (a) in the bloodstream (b) in the acidic endosome (c) never (d) on the skin
10. Exubera's principal failure was: (a) the insulin did not work (b) device/usability and commercial factors (c) it caused fungal infection (d) it was too cheap

11. Sulfasalazine releases its active in the colon because bacteria: (a) raise pH (b) cleave an azo bond (c) dissolve a coat (d) slow transit

Answer pointers. 1. Nephrotoxicity from non-specific binding; liposome alters distribution, sparing the kidney. 2. mRNA is fragile, large, charged, nuclease-degraded and cannot enter cells unaided; the LNP protects and delivers it. 3. Replaces daily oral dosing with a long-acting injection, removing the daily adherence burden. 4. Pharmacology was sound; Exubera's bulky, confusing device and commercial factors caused failure, Afrezza's simple inhaler succeeded. 5. Largely investigational/research-stage for targeted oncology. 6. Any three per Sections 18.2–18.10. 7. Per Figure 18.2. 8. Per Section 18.8 plus one other. 9. (b). 10. (b). 11. (b).

References

1. Stone NRH, Bicanic T, Salim R, Hope W. Liposomal amphotericin B (AmBisome): a review. *Drugs*. 2016;76(4):485–500.
2. Schoenmaker L, Witzigmann D, Kulkarni JA, et al. mRNA-lipid nanoparticle COVID-19 vaccines: structure and stability. *Int J Pharm*. 2021;601:120586.
3. Park K, Skidmore S, Hadar J, et al. Injectable, long-acting PLGA formulations. *J Control Release*. 2019;304:125–34.
4. Amidon S, Brown JE, Dave VS. Colon-targeted oral drug delivery systems. *AAPS PharmSciTech*. 2015;16(4):731–41.
5. Heinemann L. The failure of Exubera: are we beating a dead horse? *J Diabetes Sci Technol*. 2008;2(3):518–29.

Chapter 19 - Nanosystems and Precision Medicine

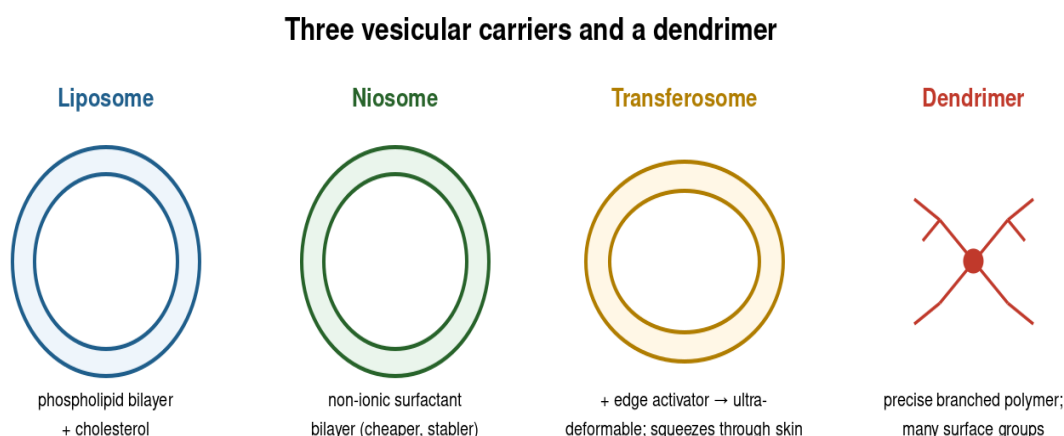
Learning outcomes

After working through this chapter the student will be able to:

1. Compare the principal nanocarrier families, liposomes, niosomes, transferosomes, polymeric nanoparticles, solid lipid nanoparticles and dendrimers, on structure, composition and use.
2. Explain how niosomes and transferosomes were developed to address specific limitations of liposomes.
3. Describe the structure of a dendrimer and the meaning of dendrimer generation.
4. Define precision (personalised) medicine and the shift from one-size-fits-all therapy.
5. Explain how targeted and stimuli-responsive nanosystems serve as the delivery arm of precision medicine.

19.1 The nanocarrier families gathered

This chapter draws together the nanosystems met throughout the book and sets them beside the vision they increasingly serve. The carriers fall into recognisable families (Figure 19.1). The **vesicular carriers**, liposomes, niosomes and transferosomes, are hollow spheres bounded by one or more bilayers, with an aqueous core. The **particulate carriers** include the polymeric nanoparticles of Chapter 3 (PLGA and others), the solid lipid nanoparticles and nanostructured lipid carriers of Chapter 4, and the **dendrimers**, a distinct branched-polymer class. Table 19.1 summarises them; the sections below focus on the carriers not yet treated in detail, the niosome, the transferosome and the dendrimer, and on how each answers a limitation of an earlier carrier.



All offer an aqueous core + bilayer (vesicles) or defined surface (dendrimer) to carry drug of either polarity.

SLN/NLC (Ch 4) add solid-lipid versions; polymeric NPs (Ch 3) add PLGA matrix particles.

Figure 19.1. Three vesicular carriers and a dendrimer. The liposome has a phospholipid bilayer (with cholesterol); the niosome replaces the phospholipid with a non-ionic surfactant for greater stability and lower cost; the transferosome adds an edge activator to make the vesicle ultra-deformable so it can squeeze through the skin. The dendrimer is a precisely

branched polymer with many surface groups. All present an aqueous core plus bilayer (vesicles) or a defined surface (dendrimer) to carry drug, complementing the solid-lipid and polymeric particles of earlier chapters.

19.2 Liposomes recapped, and the niosome

The liposome of Chapter 4 is the reference vesicle: a phospholipid bilayer enclosing an aqueous core, carrying hydrophilic drug in the core and lipophilic drug in the bilayer, stabilised by cholesterol and clinically validated in products such as AmBisome and Doxil. Its drawbacks are practical: natural phospholipids are relatively expensive, vary by source, and are prone to oxidation of their unsaturated chains, which limits shelf life.

The **niosome** was developed to keep the liposome's advantages while removing these drawbacks. A niosome is a vesicle whose bilayer is made not from phospholipid but from **non-ionic surfactants** (often with cholesterol), and structurally it behaves like a liposome, with the same aqueous core and bilayer compartments for drugs of either polarity. The benefits of the substitution are chemical stability (non-ionic surfactants are not prone to the oxidation that affects phospholipids), lower cost, and easier handling and storage. Niosomes are studied for many of the same applications as liposomes, including topical, ocular and systemic delivery, and represent a clear example of changing the building material to fix the limitation of an otherwise excellent carrier.

19.3 The transferosome

The **transferosome** is a vesicle engineered for one specific job: crossing the skin. An ordinary liposome is too large and too rigid to pass through the intact stratum corneum, so it does not deliver drug transdermally to any useful depth. The transferosome adds an **edge activator**, typically a single-chain surfactant such as sodium cholate, to the bilayer. The edge activator destabilises the bilayer just enough to make the vesicle extremely flexible and deformable, so that under the natural hydration gradient across the skin the vesicle can squeeze itself through the narrow intercellular channels of the stratum corneum, deforming to pass through openings far smaller than its resting diameter, and carry its drug across. This **ultra-deformability** is the defining feature. Transferosomes are studied as a way to extend vesicular delivery to the transdermal route and to deliver larger molecules across the skin, complementing the physical methods of Chapter 13. They are an elegant illustration of the book's central habit: a named limitation (vesicles cannot cross the skin) answered by a precise structural modification (make the vesicle deformable).

19.4 Dendrimers

The **dendrimer** is structurally unlike the vesicles. It is a synthetic polymer built by repeatedly branching outward from a central core, so that each successive layer of branches, called a **generation**, doubles the number of arms and the number of surface end-groups. The result is a roughly spherical, highly branched, tree-like macromolecule (the name comes from the Greek for tree) with a precisely defined, reproducible structure and a large number of functional groups on its surface. Polyamidoamine (PAMAM) dendrimers are the most studied family.

Two properties make dendrimers attractive as carriers. First, their structure is **monodisperse and precisely defined**: unlike most polymers, which are mixtures of chain lengths, a dendrimer of a given generation is a single well-defined molecule, which means highly reproducible behaviour. Second, the many surface end-groups provide numerous attachment

points for drug molecules, targeting ligands, solubilising groups and imaging labels, and the interior cavities can host guest molecules; drug can be carried both inside the dendrimer and conjugated to its surface. The dendrimer generation thus tunes size, the number of surface groups and the carrying capacity.

As noted candidly in Chapter 18, dendrimer-based targeted cancer therapy remains predominantly at the research and early-development stage; the chemistry is powerful and well studied, but a caution attends it, namely that some dendrimer surfaces, particularly the amine-terminated ones, can be toxic and must be modified (for example by surface capping or PEGylation) to be tolerated. Dendrimers are best understood at this level as a precisely engineered, high-capacity, but still largely investigational nanocarrier, not yet a mainstream therapeutic platform.

Table 19.1. Comparison of the principal nanocarriers.

Carrier	Building material / structure	Distinctive feature	Stage / examples
Liposome	Phospholipid bilayer + cholesterol	Core + bilayer; either polarity	Clinical (AmBisome, Doxil)
Niosome	Non-ionic surfactant bilayer	More stable, cheaper than liposome	Mostly investigational/topical
Transferosome	Bilayer + edge activator	Ultra-deformable; crosses skin	Investigational (transdermal)
Polymeric NP	PLGA/PLA matrix	Tunable biodegradable matrix	Clinical (depot products)
SLN / NLC	Solid ($\pm$ liquid) lipid core	Solid matrix; controlled release	Mostly topical/investigational
Dendrimer	Branched polymer, generations	Monodisperse; many surface groups	Largely research (oncology)

19.5 Precision medicine and where delivery fits

The final theme of the unit lifts the view from the carrier to the goal it increasingly serves. **Precision medicine** (also called personalised medicine) is the approach of tailoring medical treatment to the individual characteristics of each patient, particularly their genetic make-up, the molecular features of their disease, and relevant biomarkers, rather than treating all patients with a given diagnosis the same way (Figure 19.2). It is a deliberate move away from the **one-size-fits-all** model, in which the same drug at the same dose is given to everyone with a condition and produces benefit in some, no effect in others, and harm in a few, toward an approach that asks which treatment, at which dose, will work for this particular patient.

From one-size-fits-all to precision medicine

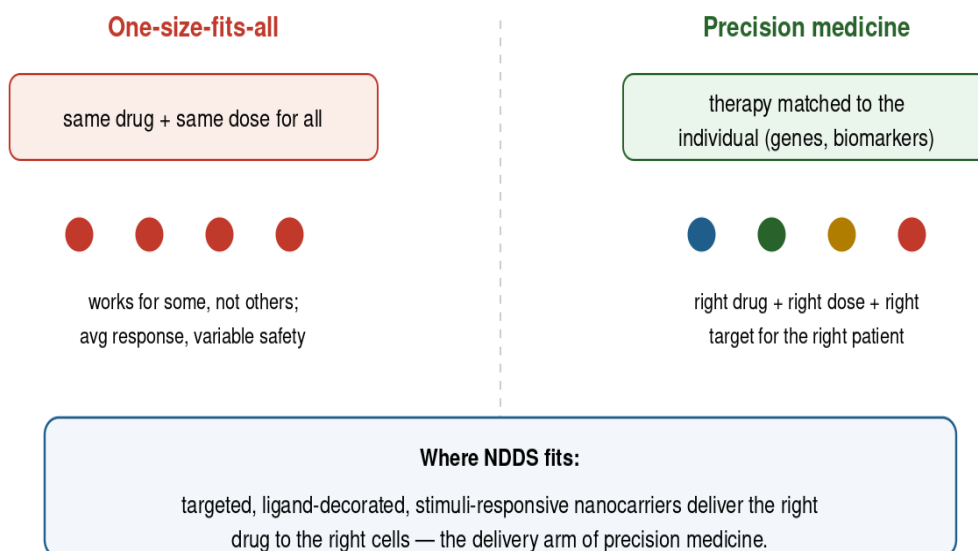


Figure 19.2. *The shift from one-size-fits-all to precision medicine. The conventional model gives the same drug and dose to everyone with a diagnosis, accepting variable response and safety. Precision medicine matches the therapy to the individual's genetic and molecular profile. Targeted, ligand-decorated and stimuli-responsive nanocarriers are the delivery arm of this approach, carrying the right drug to the right cells.*

The connection to this book is direct. Knowing the right drug for a patient is only half the problem; that drug still has to reach the right place in the right amount, which is a delivery problem. The targeted nanosystems of Chapter 2, decorated with ligands that recognise markers over-expressed on a particular patient's tumour cells, and the stimuli-responsive systems that release only in a specific microenvironment, are precisely the tools that let a precisely chosen therapy act precisely where it should. In this sense novel drug delivery is the **delivery arm of precision medicine**: molecular medicine identifies the right target and the right drug for the individual, and delivery science carries that drug to the target while sparing the rest of the body. The vision is not yet fully realised, and honesty requires noting that much of targeted nanomedicine remains a work in progress, with the EPR effect variable in humans (Chapter 2) and many targeted carriers still investigational. But the direction is clear, and it is the direction in which the entire discipline of this book is moving: from delivering a drug to a population, toward delivering the right drug to the right cells of the right patient.

Quick-recall box — nanosystems and precision medicine Vesicles: liposome (phospholipid + cholesterol, clinical); niosome (non-ionic surfactant bilayer → cheaper, more stable); transferosome (+ edge activator → ultra-deformable, crosses skin). Particles: polymeric NP (PLGA), SLN/NLC (solid lipid), dendrimer (branched polymer, generation = layer; monodisperse, many surface groups, mostly research; amine surfaces can be toxic → modify). Precision medicine: therapy matched to the individual's genes/biomarkers, replacing one-size-fits-all. NDDS = the delivery arm: targeted and stimuli-responsive carriers take the right drug to the right cells.

Chapter summary

This chapter assembles the nanocarrier families and connects them to the goal they serve. The vesicular carriers share an aqueous core and bilayer: the liposome is the clinically proven reference, the niosome substitutes a non-ionic surfactant bilayer to gain chemical stability and lower cost, and the transferosome adds an edge activator to become ultra-deformable so it can squeeze through the stratum corneum, each modification answering a specific limitation of the liposome. The particulate carriers comprise the biodegradable polymeric nanoparticles of Chapter 3, the solid lipid nanoparticles and nanostructured lipid carriers of Chapter 4, and the dendrimer, a precisely branched polymer whose generation sets its size and surface-group count, offering a monodisperse, high-capacity but still largely investigational carrier whose amine surfaces may need modification to reduce toxicity. Above the carriers sits the larger movement of precision medicine, which tailors treatment to a patient's genetic and molecular profile in place of the one-size-fits-all model, and within that movement novel drug delivery is the delivery arm: targeted, ligand-decorated and stimuli-responsive nanocarriers carry a precisely chosen therapy to precisely the cells that need it. The realisation is incomplete and parts of targeted nanomedicine remain investigational, but the trajectory of the whole field is unmistakable, from treating a population to treating the individual, with delivery science as the instrument that makes a personalised therapy reach its target.

Self-assessment

Short-answer (2 marks each)

1. How does a niosome differ from a liposome, and what advantage does the difference give?
2. What is an edge activator, and what property does it give a transferosome?
3. What is meant by a dendrimer generation?
4. Define precision medicine and contrast it with one-size-fits-all therapy.
5. In one sentence, how does NDDS serve precision medicine?

Long-answer (5–10 marks each)

6. Compare liposomes, niosomes and transferosomes in structure, building material and distinctive feature, with a diagram, explaining what limitation each later carrier addresses.
7. Describe the structure and properties of dendrimers, the meaning of generation, their attractions as carriers, and the honest current status of dendrimer cancer therapy.
8. Explain precision medicine and the role of targeted and stimuli-responsive nanosystems as its delivery arm.

Multiple-choice

9. A niosome bilayer is made from: (a) phospholipid (b) non-ionic surfactant (c) PLGA (d) gold
10. The ultra-deformability of a transferosome comes from: (a) cholesterol (b) an edge activator (c) PEG (d) a metal core
11. Precision medicine aims to: (a) give everyone the same dose (b) match therapy to the individual's molecular profile (c) avoid all drugs (d) replace delivery systems

Answer pointers. 1. Niosome uses a non-ionic surfactant bilayer instead of phospholipid; gains stability (no oxidation), lower cost. 2. A single-chain surfactant that makes the vesicle

bilayer flexible/deformable, letting it squeeze through skin. 3. Each successive branching layer from the core, doubling arms/surface groups; sets size and capacity. 4. Tailoring treatment to the individual's genes/biomarkers vs the same drug/dose for all. 5. Targeted/stimuli-responsive carriers deliver the right drug to the right cells. 6. Per Figure 19.1 and Sections 19.2–19.3. 7. Per Section 19.4. 8. Per Section 19.5. 9. (b). 10. (b). 11. (b).

References

1. Uchegbu IF, Vyas SP. Non-ionic surfactant based vesicles (niosomes) in drug delivery. *Int J Pharm.* 1998;172(1–2):33–70.
2. Cevc G. Transfersomes, liposomes and other lipid suspensions on the skin: permeation enhancement, vesicle penetration, and transdermal drug delivery. *Crit Rev Ther Drug Carrier Syst.* 1996;13(3–4):257–388.
3. Abbasi E, Aval SF, Akbarzadeh A, et al. Dendrimers: synthesis, applications, and properties. *Nanoscale Res Lett.* 2014;9(1):247.
4. Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med.* 2015;372(9):793–5.
5. Torchilin VP, editor. *Nanoparticulates as Drug Carriers.* London: Imperial College Press; 2006.

Unit V - Sterile Dosage Forms: Parenteral and Ophthalmic

Unit overview. A sterile product is held to a standard no other dosage form must meet: it must contain no living microorganisms and no pyrogens, because it is placed directly into the blood, the tissue or the eye, bypassing every natural defence. That single requirement, sterility, reshapes the water it is made from, the components that go into it, the rooms it is made in, the containers that hold it, and the tests that release it. This unit, the largest in the book, works through that chain. Chapter 20 classifies parenterals and their routes. Chapter 21 covers Water for Injection, the most important raw material. Chapter 22 covers the formulation components that make an injection safe and stable. Chapter 23 covers sterile manufacturing, sterilisation, filtration, lyophilisation and the cleanroom. Chapter 24 covers containers and closures, including Form-Fill-Seal and Blow-Fill-Seal. Chapter 25 turns to ophthalmic products and the special demands of the eye. Chapter 26 closes with the quality-control tests and cGMP compliance that prove a sterile product is fit to use. The discipline throughout is that nothing about a sterile product is casual: every choice traces back to keeping it sterile, pyrogen-free and safe to inject.

Chapter 20 - Parenterals: Introduction and Classification

Learning outcomes

After working through this chapter the student will be able to:

1. Define a parenteral product and state the three non-negotiable quality requirements common to all of them.
2. Distinguish small-volume from large-volume parenterals and give the uses of each.
3. Describe the principal routes of parenteral administration and the volume and vehicle constraints of each.
4. State the advantages and limitations of the parenteral route relative to the oral route.
5. Explain why the intravenous route forbids oils and particulates and why large-volume parenterals omit preservatives.

20.1 What makes a product parenteral

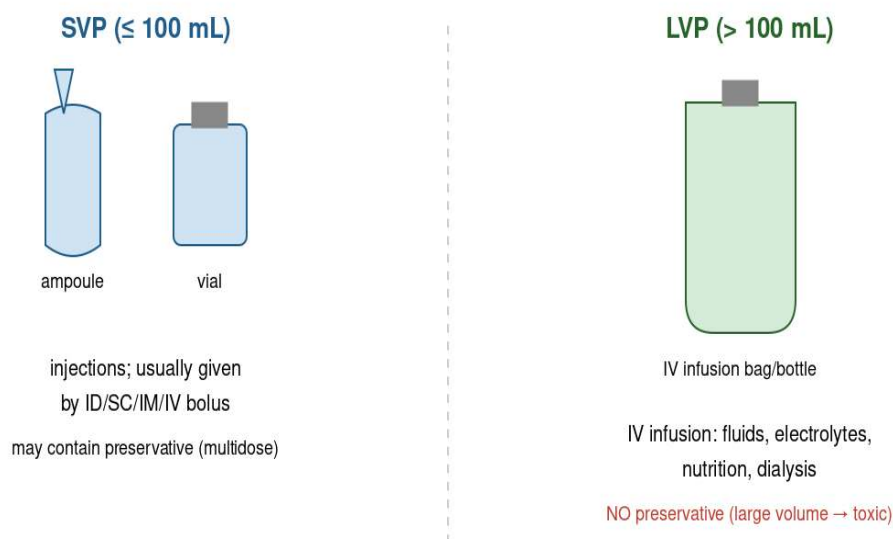
The word **parenteral** comes from the Greek for “beside the intestine”, and it denotes any route of administration that bypasses the gastrointestinal tract, in practice by injection or infusion. Because a parenteral product is introduced directly into the tissues, the bloodstream or another sterile compartment of the body, it bypasses every protective barrier, the skin, the gut wall, the liver, that a swallowed drug must cross. This is the source of the route’s great advantages and of its great demands.

Three quality requirements are absolute for every parenteral product, and they recur through the entire unit. The product must be **sterile**, free of all viable microorganisms, because any organism introduced into the body can cause infection. It must be **free of pyrogens**, fever-producing substances, chiefly bacterial endotoxins, because these provoke a dangerous febrile reaction even when no living organism is present (the distinction between sterility and pyrogen-freedom is examined in Chapters 23 and 26). And, particularly for products given intravenously, it must be **free of particulate matter**, since insoluble particles injected into the blood can lodge in small vessels and cause harm. No oral tablet is held to any of these standards; a parenteral is held to all three at once, and the rest of the unit is, in large part, an account of how each is achieved and proven.

20.2 Small-volume and large-volume parenterals

Parenterals are classified first by volume, because volume drives many design decisions (Figure 20.1).

Small-volume vs large-volume parenterals



Both must be sterile, pyrogen-free and (for IV) particulate-controlled.

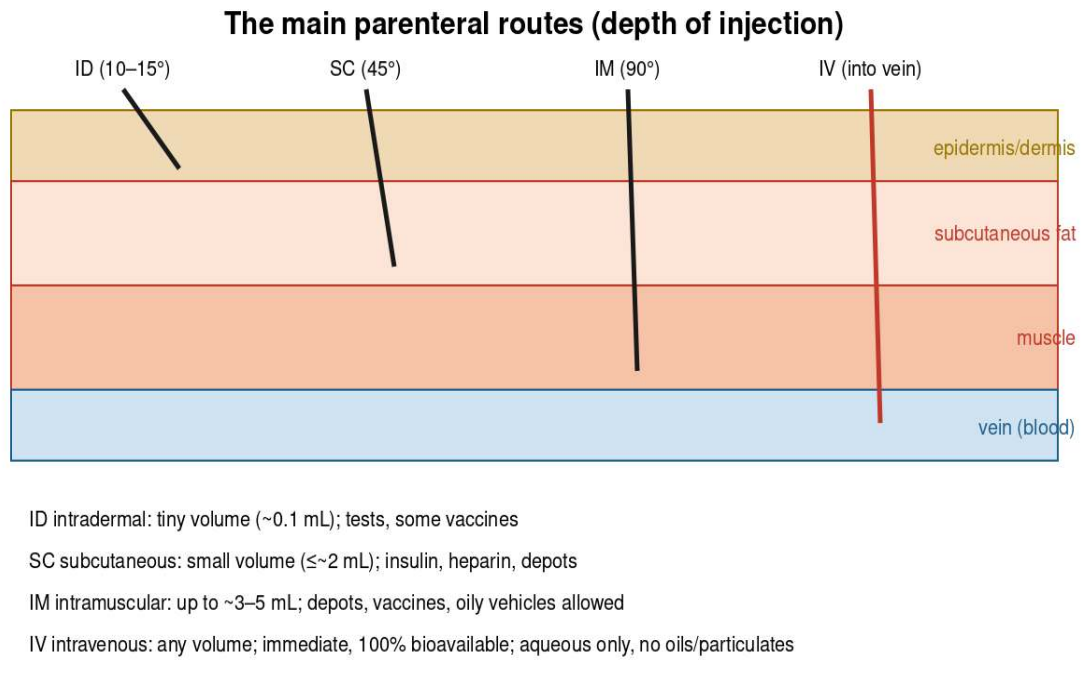
Figure 20.1. Small-volume versus large-volume parenterals. Small-volume parenterals (100 mL or less) are supplied in ampoules and vials and given by various routes; multidose containers may contain a preservative. Large-volume parenterals (more than 100 mL) are intravenous infusions supplied in bags or bottles and must not contain a preservative, because the large volume would deliver a toxic amount of it. Both must be sterile and pyrogen-free, and intravenous products must be particulate-controlled.

A **small-volume parenteral (SVP)** is conventionally defined as a sterile injection of 100 mL or less. SVPs are the familiar injections supplied in glass ampoules (sealed single-use containers) and vials (closed with a rubber stopper, single- or multi-dose), and they are given by any of the injection routes. A multidose SVP, designed to be entered several times, usually contains an antimicrobial preservative to protect the contents against contamination introduced during repeated withdrawals.

A **large-volume parenteral (LVP)** is a sterile solution of more than 100 mL, intended for intravenous infusion. LVPs include the intravenous fluids used for hydration and electrolyte replacement (such as sodium chloride and dextrose solutions), parenteral nutrition, and dialysis and irrigation solutions, supplied in flexible bags or bottles. A defining rule for LVPs is that they **must not contain an antimicrobial preservative**. The reason is straightforward and frequently examined: because the volume administered is so large, the dose of preservative delivered would itself become toxic. An LVP must therefore achieve its safety entirely through sterile manufacture and a single-use, closed container, with no chemical fallback. This single difference, preservative permitted in a small multidose vial but forbidden in a large infusion, captures how volume governs formulation in this field.

20.3 Routes of parenteral administration

Parenterals are classified again by the route and depth of administration, and each route imposes its own limits on the volume that can be given and on the kind of vehicle that is acceptable (Figure 20.2).



Deeper, more vascular routes = faster onset but less reversibility once injected.

Figure 20.2. The principal parenteral routes by depth. The intradermal route places a very small volume into the dermis; the subcutaneous route delivers small volumes into the fat beneath the skin; the intramuscular route reaches the muscle and tolerates larger volumes and oily vehicles; and the intravenous route delivers directly into a vein for immediate, complete availability but accepts only aqueous, particulate-free solutions. Deeper, more vascular routes act faster but are less reversible once injected.

The **intradermal (ID)** route places a very small volume, on the order of 0.1 mL, into the dermis. Its slow absorption and the immunological richness of the skin make it the route for diagnostic tests (such as the tuberculin test) and for certain vaccines.

The **subcutaneous (SC)** route deposits a small volume, generally up to about 2 mL, into the fatty tissue beneath the skin. Absorption is slower than from muscle, which suits self-administered chronic therapies such as insulin and heparin, and the SC space is a common site for depot systems (Chapter 12).

The **intramuscular (IM)** route delivers into skeletal muscle, which is well perfused and tolerates larger volumes, broadly up to around 3 mL (with some limit variation by site). Crucially, the muscle tolerates non-aqueous and oily vehicles and suspensions, so the IM route is used for depot injections, oily formulations and many vaccines.

The **intravenous (IV)** route delivers the product directly into a vein and so straight into the systemic circulation, giving immediate onset and complete (100%) bioavailability, and accommodating volumes from a small bolus to a large infusion. Its strict constraint is that an IV product must be an **aqueous solution free of particulates and, ordinarily, free of oils and undissolved matter**. An oil or an insoluble particle injected into a vein can act as an embolus, blocking a small blood vessel, which is dangerous, so oily vehicles and suspensions, acceptable IM, are not given IV (with rare, specialised exceptions such as lipid emulsions designed for the purpose). This is why an intramuscular oily depot cannot simply be given intravenously, a point that ties the route constraints directly back to patient safety.

Other parenteral routes exist for specialised use, including the intra-arterial, intrathecal (into the cerebrospinal fluid), intra-articular (into a joint), and intracardiac routes, each with its own stringent requirements, but the four above account for the great majority of products.

20.4 Advantages and limitations of the parenteral route

The parenteral route's advantages follow from its directness. It gives a **rapid onset** of action, vital in emergencies, and by intravenous delivery a precise, complete and immediately available dose. It **bypasses the gut and the liver**, so it works for drugs destroyed by stomach acid or gut enzymes, for drugs heavily metabolised on first pass, and for drugs too poorly absorbed orally to be useful, including most proteins and peptides. It allows treatment of patients who **cannot take medicines by mouth**, the unconscious, the vomiting, the uncooperative, or those needing fluids and nutrition directly. And it permits **localised** delivery to a specific site, such as a joint or the spinal fluid, and **sustained** delivery from a depot.

The limitations are equally direct. The route is **invasive**, causing pain and requiring, for most products, a trained person to administer, which raises cost and limits self-use. It carries the **risk of infection** at the injection site and, with poor technique or a non-sterile product, in the bloodstream. Once given, a dose is **difficult or impossible to retrieve**, so an error or an adverse reaction cannot be reversed by stopping absorption as it can with an oral dose; this irreversibility rises with the speed of the route, making intravenous errors the most consequential. And, as the whole unit shows, parenteral products are **expensive and demanding to manufacture**, because of the sterility, pyrogen and particulate requirements. The balance of these advantages and limitations explains why the parenteral route is reserved for situations where its benefits, speed, completeness, gut-and-liver bypass, or the impossibility of oral dosing, justify its costs and risks.

Quick-recall box — parenterals at a glance Three absolutes for every parenteral: sterile, pyrogen-free, (for IV) particulate-free. SVP $\leq$ 100 mL (ampoules/vials; multidose may have preservative); LVP $>$ 100 mL (IV infusion; NO preservative because the large volume would be toxic). Routes by depth: ID ($\sim$ 0.1 mL, tests/vaccines), SC ($\leq$ 2 mL, insulin/heparin/depots), IM (up to $\sim$ 3 mL, tolerates oils/suspensions/depots), IV (any volume, immediate, 100% bioavailable, aqueous and particulate-free only, no oils $\rightarrow$ embolism risk). Trade-off: fast and complete but invasive, infection-prone and irreversible.

Chapter summary

A parenteral product is one given by injection or infusion, bypassing the gastrointestinal tract and with it every natural protective barrier, which is why all parenterals share three absolute requirements, sterility, freedom from pyrogens, and, for intravenous products, freedom from particulate matter, that no oral form must meet. Parenterals are classified by volume into small-volume products of 100 mL or less, supplied in ampoules and vials and permitted a preservative in multidose form, and large-volume products above 100 mL for intravenous infusion, which must omit any preservative because the large volume would deliver a toxic dose of it. They are classified again by route and depth: intradermal for tiny volumes in tests and vaccines, subcutaneous for small volumes such as insulin and depots, intramuscular for larger volumes and for oily vehicles and depots, and intravenous for immediate, complete delivery of aqueous, particulate-free solutions only, since oils and particles would cause embolism. The route's advantages, rapid and complete action, bypass of gut and liver, use in patients who cannot take oral medicines, and localised or sustained delivery, are balanced

against its invasiveness, infection risk, irreversibility and demanding manufacture, which together confine the parenteral route to situations where its directness is genuinely needed and set up the sterility-driven chapters that follow.

Self-assessment

Short-answer (2 marks each)

1. Define a parenteral and state the three absolute quality requirements.
2. Distinguish SVP from LVP by volume and give one use of each.
3. Why must a large-volume parenteral not contain a preservative?
4. Why is an oily vehicle acceptable intramuscularly but not intravenously?
5. Give two advantages and two limitations of the parenteral route.

Long-answer (5–10 marks each)

6. Classify parenteral products by volume and by route, with a diagram, giving the volume and vehicle constraints of each route.
7. Discuss the advantages and limitations of the parenteral route compared with the oral route.
8. Explain why sterility, pyrogen-freedom and particulate control are required for parenterals, relating each to a specific risk.

Multiple-choice

9. A large-volume parenteral is defined as a sterile solution of: (a) ≤ 10 mL (b) ≤ 100 mL (c) > 100 mL (d) exactly 50 mL
10. Which route gives immediate, 100% bioavailability? (a) intradermal (b) subcutaneous (c) intramuscular (d) intravenous
11. Oily depot injections are typically given by the: (a) IV route (b) IM route (c) intra-arterial route (d) intrathecal route

Answer pointers. 1. Product given by injection bypassing the GI tract; sterile, pyrogen-free, particulate-free (IV). 2. SVP ≤ 100 mL (injection); LVP > 100 mL (IV infusion/fluids). 3. The large volume would deliver a toxic dose of preservative. 4. Oils/particles injected IV cause embolism; muscle tolerates oily depots. 5. Advantages: rapid/complete, bypass gut/liver (etc.); limitations: invasive, infection risk, irreversible, costly (any two each). 6. Per Figures 20.1–20.2. 7. Per Section 20.4. 8. Sterility $\rightarrow$ infection, pyrogens $\rightarrow$ fever, particulates $\rightarrow$ embolism. 9. (c). 10. (d). 11. (b).

References

1. Indian Pharmacopoeia Commission. Indian Pharmacopoeia: General monograph on Parenteral Preparations. Ghaziabad: IPC; current edition.
2. Allen LV, Ansel HC. Ansel's Pharmaceutical Dosage Forms and Drug Delivery Systems. 11th ed. Philadelphia: Wolters Kluwer; 2018.
3. Aulton ME, Taylor KMG, editors. Aulton's Pharmaceutics. 6th ed. Edinburgh: Elsevier; 2022.
4. Remington: The Science and Practice of Pharmacy. 23rd ed. London: Academic Press/Pharmaceutical Press; 2020.

Chapter 21 - Water for Injection

Learning outcomes

After working through this chapter the student will be able to:

1. Place Water for Injection within the hierarchy of pharmaceutical waters and state how it differs from Purified Water.
2. State the pharmacopoeial requirements for WFI, including the endotoxin, TOC, conductivity and microbial limits.
3. Describe how WFI is produced, including distillation and the now-permitted validated membrane routes.
4. Explain why WFI is stored hot and circulated continuously, and the role of the distribution loop.
5. Explain why endotoxin, rather than microbial count alone, is the defining concern for WFI.

21.1 Water as the principal ingredient

Water is by far the most-used ingredient in sterile manufacturing: it is the vehicle of most injections, the solvent for compounding, and the medium for cleaning equipment and containers. Because it is present in such quantity and in such intimate contact with the product, the quality of the water sets a ceiling on the quality of everything made from it. A contaminant in the water becomes a contaminant in the product. For sterile products the relevant grade is **Water for Injection (WFI)**, the most highly purified water in routine pharmaceutical use, and understanding its specification and production is foundational to the whole unit.

21.2 The hierarchy of pharmaceutical waters

Pharmacopoeias define several grades of water of ascending purity, and WFI sits at the top of the grades used as ingredients (Figure 21.1). **Potable (drinking) water** is the starting feed; it is safe to drink but not pure enough to be an ingredient, and serves only as the raw water from which purer grades are made. **Purified Water** is produced from potable water by a purification process and meets limits on conductivity (a measure of dissolved ions) and total organic carbon (a measure of dissolved organic matter); it is used for non-sterile products and as feed for WFI, but, crucially, it carries **no endotoxin limit**. **Water for Injection** meets all the chemical limits of Purified Water and adds the requirement that defines it: a limit on **bacterial endotoxins**.

Grades of pharmaceutical water (increasing purity →)

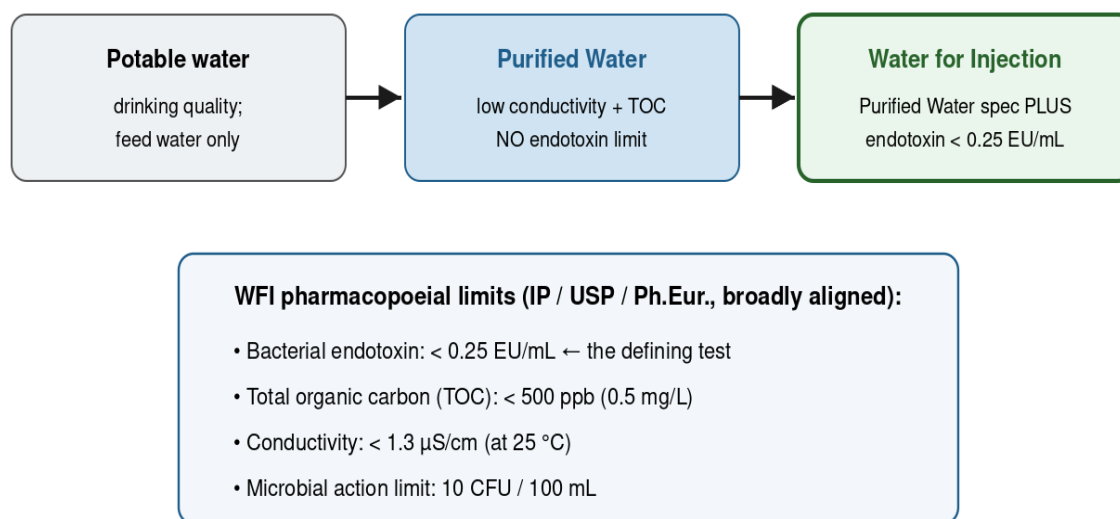


Figure 21.1. The grades of pharmaceutical water. Potable water is feed only; Purified Water meets conductivity and total-organic-carbon limits but has no endotoxin requirement; Water for Injection meets the Purified Water specification and adds the defining bacterial-endotoxin limit. The box lists the WFI pharmacopoeial limits, which are broadly aligned across the Indian, US and European pharmacopoeias.

The single conceptual point to carry away is that **the difference between Purified Water and WFI is the endotoxin requirement**. Both are chemically pure; only WFI is controlled for endotoxin, because only WFI is destined for products that enter the body parenterally, where endotoxin causes fever.

21.3 Pharmacopoeial requirements for WFI

The quality of WFI is fixed by pharmacopoeial limits that are broadly harmonised across the Indian Pharmacopoeia, the United States Pharmacopeia and the European Pharmacopoeia.^{1,2} The principal requirements are:

- **Bacterial endotoxin:** not more than **0.25 EU/mL** (endotoxin units per millilitre). This is the defining limit of WFI.
- **Total organic carbon (TOC):** not more than about **500 ppb** (0.5 mg per litre), limiting dissolved organic matter.
- **Conductivity:** a low value, with a limit of the order of **1.3 µS/cm at 25 °C** (the precise limit is temperature-dependent in the pharmacopoeial staged test), limiting dissolved ionic impurities.
- **Microbial quality:** an action limit commonly set at **10 CFU per 100 mL** (colony-forming units), recognising that WFI as produced and stored is highly pure but not, in the storage loop, sterile in the way a final sterilised product is.

It is worth being precise about that last point, because it is a common source of confusion. WFI in the distribution system is held to a very low microbial limit but is not necessarily “sterile” as a category in the way a terminally sterilised ampoule is; it is kept so pure and so

inhospitable to growth that its bioburden and, above all, its endotoxin stay within limits. When WFI is used to make an injection, the product is subsequently sterilised or aseptically processed (Chapter 23). The reason endotoxin is controlled so strictly is that endotoxin is heat-stable: even if every organism in the water is later killed, the endotoxin those organisms produced survives ordinary sterilisation and would remain in the product. Controlling endotoxin in the water is therefore the only reliable way to keep it out of the product, which is why it defines WFI.

21.4 Production of WFI

Historically, and still as the reference method, WFI is produced by **distillation**. Feed water (itself pre-treated, typically by softening, reverse osmosis and other steps to protect the still) is boiled, and the vapour is condensed to give the product. Distillation is favoured because it separates water from almost everything else: dissolved ions, organic molecules, microorganisms and, importantly, endotoxins are left behind in the residue, because endotoxin is non-volatile and does not carry over with the steam (well-designed stills include features to prevent droplet carry-over that could bring endotoxin across). A multi-effect still or a vapour-compression still produces WFI efficiently at scale.

A significant regulatory development is that the major pharmacopoeias now also permit WFI to be produced by **validated membrane-based methods**, such as reverse osmosis combined with ultrafiltration and other steps, rather than by distillation alone, provided the process is shown to produce water meeting the WFI specification consistently. The European Pharmacopoeia's acceptance of non-distillation routes for WFI, followed by alignment elsewhere, ended distillation's status as the only permitted method.³ The principle to retain is that the method must be **validated to reliably remove endotoxin and impurities to the WFI limits**; distillation does this inherently, and a membrane route must prove that it does so equally reliably. [

21.5 Storage and distribution

Producing WFI to specification is only half the task; keeping it within specification until it is used is the other half, and it is where most failures occur (Figure 21.2). Freshly produced WFI is extremely pure, but pure water left standing at room temperature is an invitation to the few surviving or ingressing bacteria to grow, and as they grow they shed endotoxin, pushing the water out of specification. The storage and distribution system is engineered specifically to prevent this.

Producing and distributing WFI

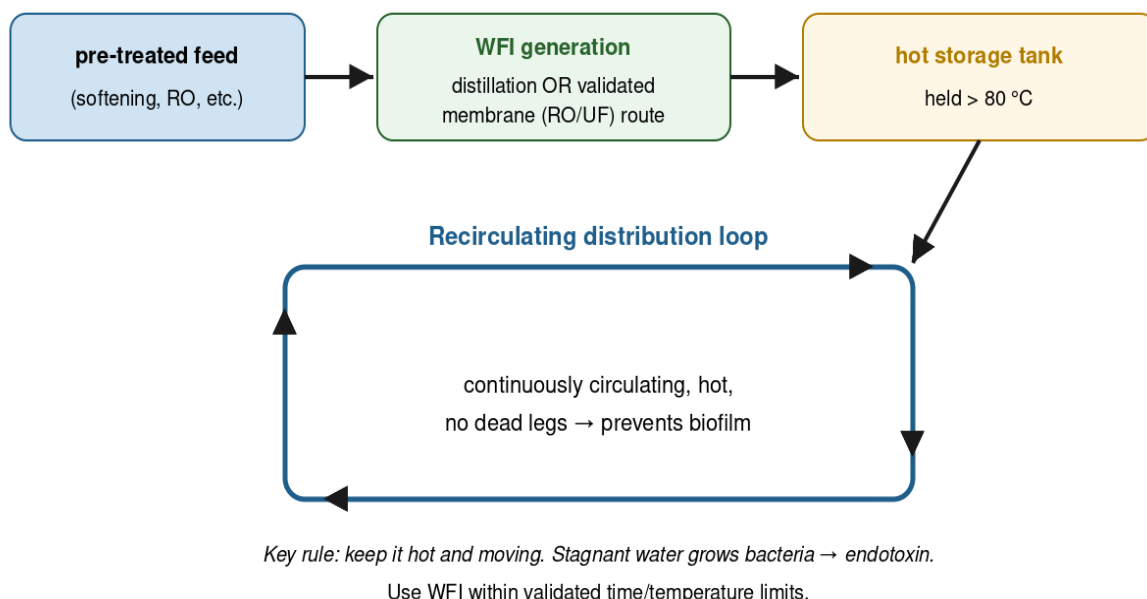


Figure 21.2. Producing and distributing WFI. Pre-treated feed water is converted to WFI by distillation or a validated membrane route, then held in a hot storage tank and circulated through a recirculating distribution loop kept hot and free of dead legs. The governing rule is to keep the water hot and continuously moving, because stagnant water allows bacterial growth and the endotoxin release that follows.

Two principles govern the system. First, **keep it hot**: WFI is commonly stored and circulated at a high temperature, above about 80 °C, because at this temperature bacteria cannot grow, so both bioburden and endotoxin stay low. Second, **keep it moving**: the water is held in a continuously **recirculating distribution loop** rather than in static pipes, because flowing water resists the formation of the **biofilm** (a sticky bacterial layer) that develops on wetted surfaces in stagnant systems. The loop is designed without **dead legs**, sections of pipe where water can stagnate, since these are notorious sites for biofilm and contamination. The system is sanitised periodically (for example by steam or hot water) and monitored, and the WFI is used within validated time and temperature limits. The whole design philosophy reduces to the rule in Figure 21.2: keep it hot and keep it moving, because stagnant water grows bacteria, and bacteria mean endotoxin.

Quick-recall box — Water for Injection WFI = Purified Water spec PLUS an endotoxin limit; that endotoxin limit is the defining difference. Limits: endotoxin < 0.25 EU/mL, TOC < 500 ppb, conductivity < ~1.3 µS/cm (25 °C), microbial action limit 10 CFU/100 mL. Production: distillation (reference; leaves non-volatile endotoxin behind) or validated membrane routes (RO/UF). Storage/distribution: hot (> ~80 °C) and continuously recirculating, no dead legs, to stop biofilm and endotoxin. Why endotoxin matters most: it is heat-stable and survives sterilisation, so it must be excluded at the water stage.

Chapter summary

Water is the principal ingredient of sterile products, so its quality caps the quality of everything made from it, and the grade required is Water for Injection, the purest pharmaceutical water in routine use. Within the hierarchy of waters, potable water is feed only, Purified Water meets conductivity and total-organic-carbon limits, and Water for Injection meets those same chemical limits and adds the bacterial-endotoxin limit that defines it. The pharmacopoeial requirements, broadly harmonised across the Indian, US and European pharmacopoeias, are an endotoxin limit of not more than 0.25 EU/mL, a total-organic-carbon limit around 500 ppb, a conductivity limit of the order of 1.3 $\mu\text{S}/\text{cm}$ at 25 °C, and a microbial action limit commonly of 10 CFU per 100 mL. WFI is produced by distillation, the reference method, which leaves non-volatile endotoxin and other impurities behind, and now also by validated membrane routes such as reverse osmosis with ultrafiltration, provided they reliably meet the specification. Keeping WFI within specification during storage and distribution is achieved by holding it hot, above about 80 °C, and continuously recirculating it through a loop free of dead legs, because stagnant water grows biofilm and the bacteria release endotoxin. The reason endotoxin is the central concern is that it is heat-stable and survives ordinary sterilisation, so it must be excluded at the water stage rather than relied upon to be removed later.

Self-assessment

Short-answer (2 marks each)

1. What single requirement distinguishes WFI from Purified Water?
2. State the endotoxin and TOC limits for WFI.
3. Why does distillation effectively remove endotoxin from water?
4. Why is WFI stored hot and continuously circulated?
5. What is a dead leg and why is it avoided?

Long-answer (5–10 marks each)

6. Describe the hierarchy of pharmaceutical waters and the pharmacopoeial requirements for WFI, with a diagram.
7. Compare distillation and validated membrane methods for producing WFI, explaining the central requirement either method must meet.
8. Explain how WFI is stored and distributed to keep it within specification, and why endotoxin rather than viable count is the defining concern.

Multiple-choice

9. The defining pharmacopoeial limit of WFI is: (a) conductivity (b) TOC (c) bacterial endotoxin ≤ 0.25 EU/mL (d) colour
10. WFI is typically stored at: (a) 4 °C (b) room temperature (c) above ~80 °C (d) frozen
11. Endotoxin must be controlled in the water because it is: (a) volatile (b) heat-stable and survives sterilisation (c) easily filtered (d) harmless

Answer pointers. 1. The bacterial-endotoxin limit. 2. Endotoxin < 0.25 EU/mL; TOC $< \sim 500$ ppb. 3. Endotoxin is non-volatile and stays in the residue while water vaporises (with carry-over prevention). 4. Heat (>80 °C) prevents growth and circulation prevents biofilm, keeping bioburden/endotoxin low. 5. A stagnant pipe section; it harbours

biofilm/contamination. 6. Per Figure 21.1. 7. Per Section 21.4. 8. Per Section 21.5. 9. (c). 10. (c). 11. (b).

References

1. Indian Pharmacopoeia Commission. Indian Pharmacopoeia: monograph on Water for Injection. Ghaziabad: IPC; current edition.
2. United States Pharmacopeia. General chapter <1231> Water for Pharmaceutical Purposes and the Water for Injection monograph. Rockville (MD): USP.
3. European Directorate for the Quality of Medicines. European Pharmacopoeia: Water for Injections monograph (permitting non-distillation production). Strasbourg: EDQM; current edition.
4. Aulton ME, Taylor KMG, editors. Aulton's Pharmaceutics. 6th ed. Edinburgh: Elsevier; 2022.

Chapter 22 - Formulation Components of Parenterals

Learning outcomes

After working through this chapter the student will be able to:

1. Identify the functional categories of components in an injectable formulation and the role of each.
2. Describe aqueous and non-aqueous vehicles and when each is used.
3. Explain the use of buffers, co-solvents, surfactants, antioxidants, chelating agents and preservatives in parenterals, with examples.
4. Explain isotonicity and osmolality, why injections are adjusted to be isotonic, and the consequences of hypotonic and hypertonic products.
5. State which components are restricted or forbidden by route or volume, especially preservatives in LVPs.

22.1 Building a safe injection

An injectable formulation is more than a drug dissolved in water. Around the drug sits a small set of carefully chosen excipients, each present to solve a specific problem of solubility, stability, comfort or safety (Figure 22.1). Because the product is injected, every excipient must itself be safe by that route and at that amount, which sharply limits the choices, recalling the regulatory constraints of Chapter 6. This chapter works through the functional categories in turn and then treats isotonicity, the property that most distinguishes parenteral formulation from oral.

Functional components of an injectable formulation

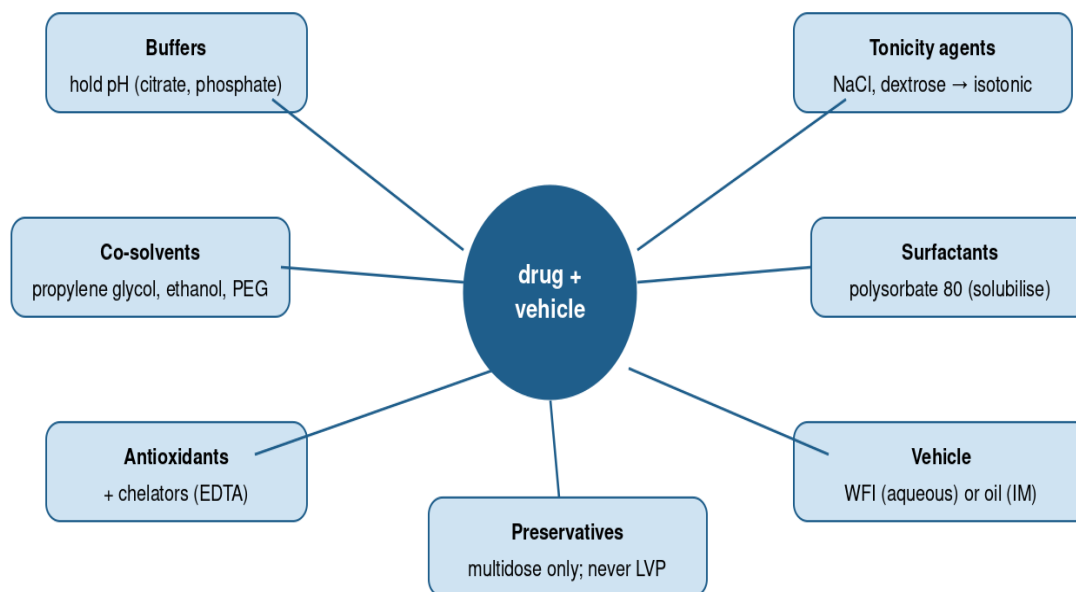


Figure 22.1. The functional components of an injectable formulation around the drug and vehicle: buffers to hold pH, tonicity agents to match plasma, co-solvents and surfactants to dissolve poorly soluble drugs, antioxidants and chelating agents to prevent oxidation, and, in multidose products only, antimicrobial preservatives. Each component must be safe by the intended route and is restricted accordingly.

22.2 Vehicles

The **vehicle** is the medium in which the drug is dissolved or dispersed, and it makes up the bulk of the product. The default and overwhelmingly preferred vehicle is **water**, specifically Water for Injection (Chapter 21), because water is physiologically compatible and most drugs and the body itself are water-based. The great majority of injections are aqueous.

When a drug is too poorly soluble in water, or is unstable in water, a **non-aqueous vehicle** may be used. **Fixed oils** of vegetable origin (such as sesame, castor or arachis oil, meeting pharmacopoeial purity) are used for oil-soluble drugs and for depot injections that release slowly from an oily reservoir. Oily vehicles are restricted to the intramuscular and subcutaneous routes and must never be given intravenously, because an oil injected into a vein causes embolism (Chapter 20). **Water-miscible co-solvents** (Section 22.4) provide a middle path for drugs that are sparingly soluble in water but do not need a fully oily vehicle.

22.3 Buffers

Many drugs are stable, soluble or both only within a narrow pH range, and the pH of a solution can drift during storage as the drug or excipients react. A **buffer** is a combination of a weak acid and its conjugate base that resists changes in pH, holding the formulation at the value where the drug is most stable and best tolerated. Common parenteral buffers include **citrate, phosphate and acetate** systems. The formulator must balance two demands: the buffer should hold the drug at its optimum pH, but the product should also be reasonably close to blood pH (about 7.4) and have a low enough buffer capacity that, once injected, it is

quickly neutralised by the blood's own buffering, so it does not cause pain or irritation at the injection site. A strongly buffered product far from physiological pH is poorly tolerated, so buffer choice and concentration are a compromise between stability and comfort.

22.4 Co-solvents and surfactants

Two strategies help dissolve a poorly water-soluble drug in an injection. **Co-solvents** are water-miscible organic solvents added to water to increase the solubility of a lipophilic drug; the common parenteral co-solvents are **propylene glycol, ethanol, glycerol and the polyethylene glycols (PEG 300, PEG 400)**. They are effective but must be used within safe limits, because in excess they can cause pain, haemolysis or systemic toxicity, and a drug solubilised in a co-solvent system may precipitate when the product is diluted into blood or an infusion, which must be checked. **Surfactants**, chiefly the non-ionic **polysorbate 80 (Tween 80)** introduced in Chapter 4, solubilise lipophilic drugs by incorporating them into micelles and also stabilise proteins and emulsions; polysorbate 80 is widely used in parenteral and vaccine formulations. Surfactants too are used at the lowest effective concentration, as they can cause irritation and, in some cases, hypersensitivity reactions.

22.5 Antioxidants and chelating agents

Some drugs are degraded by oxidation, a reaction driven by dissolved oxygen and often catalysed by trace metal ions. Two complementary excipients protect against this. **Antioxidants** are substances more readily oxidised than the drug, so they are consumed preferentially and spare the drug; examples used in parenterals include **sodium metabisulfite, sodium bisulfite, ascorbic acid and the tocopherols**. **Chelating agents** bind and sequester the trace metal ions (iron, copper) that catalyse oxidation, removing the catalyst; the standard parenteral chelator is **disodium edetate (EDTA)**. The two are often used together, the chelator removing the catalyst and the antioxidant mopping up oxidation that still occurs, and they are reinforced by physical measures such as purging the solution and the container headspace with an inert gas (nitrogen) to exclude oxygen.

22.6 Antimicrobial preservatives

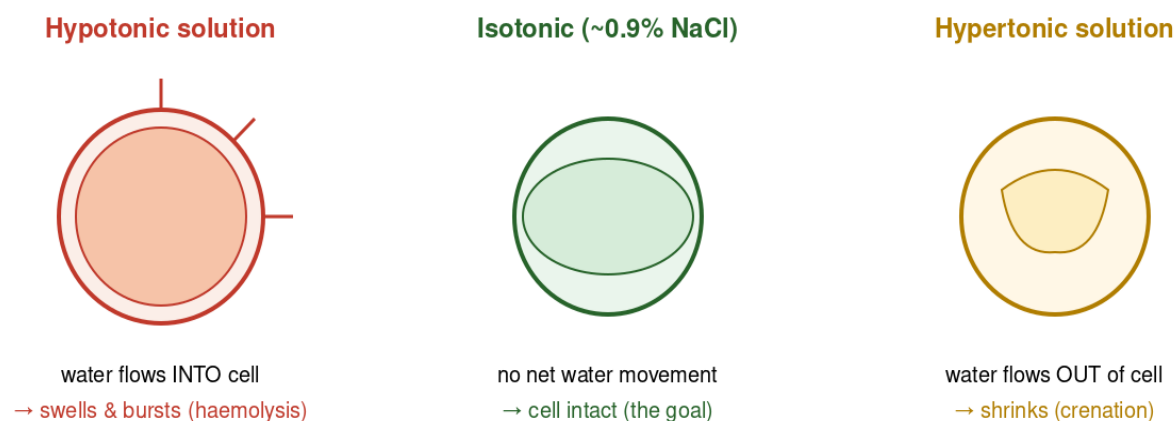
A **preservative** is an antimicrobial agent added to prevent the growth of microorganisms that may be introduced into the product. Its role in parenterals is specific and bounded. Preservatives are added to **multidose containers**, which are entered repeatedly and so risk contamination at each withdrawal; the preservative protects the remaining contents between uses. Common parenteral preservatives include **benzyl alcohol, the parabens (methyl- and propylparaben), phenol and phenylmercuric compounds**, used at low, defined concentrations.

Two firm rules apply and are frequently examined. First, preservatives are **not used in single-dose containers**, which are used once and discarded, so no protection between uses is needed. Second, and most importantly, preservatives are **forbidden in large-volume parenterals**: because such a large volume is administered, the dose of preservative delivered would itself be toxic (Chapter 20). Preservatives are also avoided or restricted in products for routes where they are especially harmful, such as intrathecal and intraocular injection, and in products for neonates, in whom benzyl alcohol in particular has caused serious toxicity. The general principle is that a preservative is a calculated compromise, accepting a small dose of a mildly toxic antimicrobial to protect a multidose product, and where that compromise cannot be safely made, the product must instead be single-use and preservative-free.

22.7 Isotonicity and osmolality

The property that most sharply distinguishes parenteral formulation is **tonicity**, because an injection is placed directly among the body's cells, and the osmotic relationship between the solution and those cells matters immediately (Figure 22.2).

Why injections are made isotonic (effect on red blood cells)



Osmolarity matched to plasma (~285–310 mOsm/L) protects cells.

Adjusted with NaCl/dextrose; tonicity calculated (e.g. NaCl equivalents / freezing-point depression).

Figure 22.2. Why injections are made isotonic. In a hypotonic solution, water flows into the red blood cells, which swell and burst (haemolysis). In a hypertonic solution, water flows out and the cells shrink (crenation). In an isotonic solution, matched to the osmolality of plasma, there is no net water movement and the cells stay intact, which is the goal. Tonicity is adjusted with sodium chloride or dextrose and calculated by methods such as sodium chloride equivalents or freezing-point depression.

Osmolality (or osmolarity) measures the concentration of dissolved particles in a solution, and **tonicity** describes the effect of that concentration on cells across a membrane. Blood plasma has an osmolality of roughly 285 to 310 mOsm/L. A solution with the same effective osmotic concentration as plasma is **isotonic**, and it causes no net movement of water across the red-cell membrane, so the cells are unharmed. A **hypotonic** solution is more dilute than plasma, so water flows into the cells, which swell and may rupture, an effect called **haemolysis** that is dangerous when it happens to red blood cells in the bloodstream. A **hypertonic** solution is more concentrated, so water flows out of the cells, which shrink, an effect called **crenation**. The clear goal is to formulate injections to be isotonic, so that they neither lyse nor shrink the cells they contact; the classic reference is **0.9% w/v sodium chloride**, “normal saline”, which is isotonic with blood.

When a drug solution is not naturally isotonic, a **tonicity-adjusting agent**, usually **sodium chloride or dextrose**, is added to bring it to isotonicity. The amount needed is calculated by established methods, principally the **sodium chloride equivalent method** (each drug is assigned the weight of NaCl osmotically equivalent to 1 g of the drug) and the **freezing-point depression method** (a solution is made isotonic when its freezing-point depression matches that of plasma, about 0.52 °C). These calculations are a standard part of parenteral and

ophthalmic formulation. A practical qualification: strict isotonicity is the ideal, but slightly hypertonic products are tolerated when given slowly into a large vein, where they are rapidly diluted by the flowing blood, which is why some concentrated injections and infusions are acceptable despite not being exactly isotonic.

Quick-recall box — injectable components Vehicle: WFI (default) or fixed oil (IM/SC only, never IV). Buffer: holds pH for stability (citrate/phosphate/acetate), but low capacity so blood neutralises it (comfort). Co-solvents: propylene glycol, ethanol, PEG (watch for precipitation on dilution). Surfactant: polysorbate 80 (solubilise/stabilise). Oxidation: antioxidants (metabisulfite, ascorbic acid) + chelator (EDTA) + nitrogen purge. Preservatives: multidose only; NEVER in LVP (toxic dose); avoid intrathecal/neonatal. Tonicity: make isotonic (~0.9% NaCl, ~285–310 mOsm/L); hypotonic → haemolysis, hypertonic → crenation; adjust with NaCl/dextrose via NaCl-equivalent or freezing-point methods.

Chapter summary

An injectable formulation surrounds the drug with a small set of excipients, each chosen to solve a specific problem and each constrained to be safe by the route of administration. The vehicle is normally Water for Injection, with fixed oils reserved for oil-soluble or depot products given intramuscularly or subcutaneously but never intravenously. Buffers such as citrate, phosphate and acetate hold the formulation at the pH where the drug is stable, balanced against the need to stay tolerable on injection. Co-solvents like propylene glycol, ethanol and the polyethylene glycols, and surfactants such as polysorbate 80, dissolve poorly water-soluble drugs, used at the lowest effective level and checked for precipitation on dilution. Oxidation-prone drugs are protected by antioxidants such as sodium metabisulfite together with the chelating agent disodium edetate and by nitrogen purging. Antimicrobial preservatives protect multidose containers but are excluded from single-dose products and absolutely forbidden in large-volume parenterals, where the volume would make the preservative dose toxic, and avoided in intrathecal and neonatal products. The property that most defines parenteral formulation is tonicity: injections are made isotonic with plasma so that they neither swell and burst cells (haemolysis, from hypotonic solutions) nor shrink them (crenation, from hypertonic solutions), using sodium chloride or dextrose calculated by the sodium chloride equivalent or freezing-point depression methods, with the reference being 0.9% normal saline.

Self-assessment

Short-answer (2 marks each)

1. Name the default vehicle for injections and one situation requiring a non-aqueous vehicle.
2. Why must a parenteral buffer have a relatively low buffer capacity?
3. How do antioxidants and chelating agents complement each other?
4. State the two firm rules limiting preservative use in parenterals.
5. Define isotonic, hypotonic and hypertonic and the effect of each on red cells.

Long-answer (5–10 marks each)

6. Describe the functional components of an injectable formulation, giving the role and an example of each, with a diagram.

7. Explain isotonicity and osmolality, the consequences of hypotonic and hypertonic injections, and how tonicity is adjusted and calculated.
8. Discuss the use and restriction of antimicrobial preservatives in parenterals, including why they are forbidden in LVPs and restricted in certain routes and populations.

Multiple-choice

9. A solution that causes red cells to swell and burst is: (a) isotonic (b) hypotonic (c) hypertonic (d) buffered
10. The standard parenteral chelating agent is: (a) polysorbate 80 (b) disodium edetate (EDTA) (c) benzyl alcohol (d) dextrose
11. Antimicrobial preservatives are forbidden in: (a) single-dose ampoules (b) multidose vials (c) large-volume parenterals (d) all injections

Answer pointers. 1. Water for Injection; oil-soluble drug or oily depot (non-aqueous). 2. So blood quickly neutralises it after injection, avoiding pain/irritation. 3. Chelator removes metal catalysts; antioxidant is preferentially oxidised, sparing the drug. 4. Not in single-dose containers; never in LVPs (toxic dose). 5. Isotonic = no net water movement; hypotonic = water in → haemolysis; hypertonic = water out → crenation. 6. Per Figure 22.1. 7. Per Section 22.7. 8. Per Section 22.6. 9. (b). 10. (b). 11. (c).

References

1. Indian Pharmacopoeia Commission. Indian Pharmacopoeia: Parenteral Preparations and excipient monographs. Ghaziabad: IPC; current edition.
2. Akers MJ. Sterile Drug Products: Formulation, Packaging, Manufacturing and Quality. New York: CRC Press; 2010.
3. Allen LV, Ansel HC. Ansel's Pharmaceutical Dosage Forms and Drug Delivery Systems. 11th ed. Philadelphia: Wolters Kluwer; 2018.
4. Aulton ME, Taylor KMG, editors. Aulton's Pharmaceutics. 6th ed. Edinburgh: Elsevier; 2022.

Chapter 23 - Sterile Manufacturing: Sterilisation, Filtration, Lyophilisation and Environmental Control

Learning outcomes

After working through this chapter the student will be able to:

1. Distinguish terminal sterilisation from aseptic processing and explain why terminal sterilisation is preferred where possible.
2. Describe the principal sterilisation methods and the meaning of F_0 and the sterility assurance level.
3. Explain sterile filtration, the role of the $0.22\ \mu\text{m}$ filter, and its place in aseptic processing.
4. Describe the principle and the three phases of lyophilisation and why it is used.
5. Describe the cleanroom classification system and the principle of environmental control in sterile manufacture.

23.1 Two philosophies of achieving sterility

Every sterile product reaches sterility by one of two routes, and the choice between them is the most consequential decision in sterile manufacturing (Figure 23.1). The principle was introduced in Chapter 13 and is developed fully here.

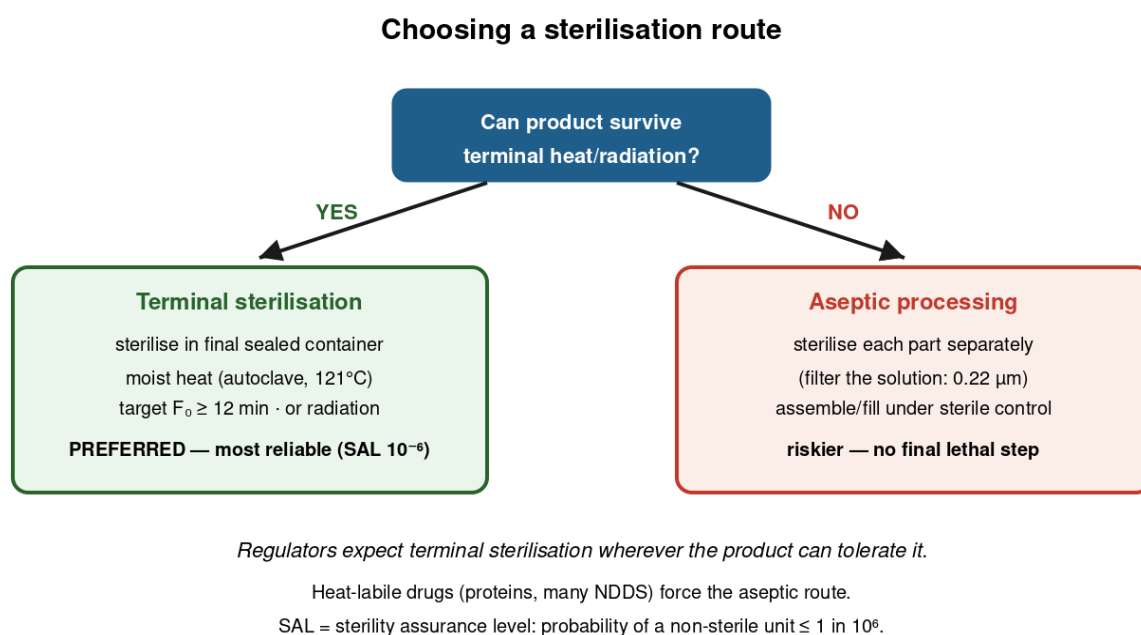


Figure 23.1. Choosing a sterilisation route. If the product can withstand a terminal process, it is sterilised in its final sealed container by moist heat or radiation, the preferred and most reliable route. If the product is heat- or radiation-labile, it must be made by aseptic processing: each component is sterilised separately, the solution is sterile-filtered through a $0.22\ \mu\text{m}$ filter, and the product is assembled and filled under sterile conditions, a route that is inherently riskier because there is no final lethal step.

In **terminal sterilisation**, the product is filled and sealed in its final container and then sterilised as a whole, most often by moist heat in an autoclave. This is the preferred route, and regulators expect it to be used whenever the product can survive it, for one decisive reason: the sterilising step is applied to the finished, sealed article, so it kills any organism introduced at any earlier stage. Sterility is assured by a validated lethal process acting on the final product.

In **aseptic processing**, the product cannot survive a terminal process, so sterility must be built up step by step. Each component, the solution, the container, the closure, is sterilised separately by a method it can tolerate (the solution typically by filtration), and these sterile components are then brought together and filled under conditions designed to exclude all contamination. The weakness, as Chapter 13 noted, is that there is no final lethal step: sterility depends entirely on maintaining control throughout assembly, so any lapse contaminates the product irrecoverably. Aseptic processing is therefore inherently riskier and is reserved for products, many proteins, vaccines and heat-labile novel systems, that cannot be terminally sterilised. The governing rule is simple: terminally sterilise if you can; process aseptically only if you must.

23.2 Methods of sterilisation and the measures of success

The principal sterilisation methods are chosen by what the product and components can tolerate. **Moist heat (steam under pressure, in an autoclave)** is the workhorse, typically operating at 121 °C; it is highly effective because moist heat denatures microbial proteins efficiently, and it is the default for aqueous products and equipment that can withstand it. **Dry heat (hot-air oven)**, at higher temperatures, is used for glassware, oils and powders that steam cannot penetrate or would damage, and at high temperature it also destroys (depyrogenates) endotoxin. **Radiation sterilisation** (gamma or electron beam) is used for heat-sensitive single-use plastics and devices. **Gaseous sterilisation** with ethylene oxide is used for heat- and moisture-sensitive devices that cannot be irradiated. **Filtration** (Section 23.3) sterilises heat-labile liquids by removing rather than killing organisms.

Two quantitative concepts express how well sterilisation works. The **sterility assurance level (SAL)** is the probability that a given unit remains non-sterile after the process; the standard target for a terminally sterilised product is an SAL of 10^{-6} , meaning no more than one non-sterile unit in a million. The **F₀ value** is used for moist-heat sterilisation: it expresses the lethality of a heat cycle as the equivalent number of minutes at 121 °C, allowing cycles run at different temperatures or with heat-up and cool-down phases to be compared on a common scale. A widely used reference is an F₀ of at least **12 minutes** for an overkill moist-heat cycle, though the required F₀ depends on the bioburden and heat resistance of the specific product. These measures turn “sterile” from a vague claim into a validated, quantified outcome.

23.3 Sterile filtration

For a heat-labile liquid that cannot be autoclaved, sterility is achieved by **sterile filtration**: the solution is passed through a membrane filter whose pores are small enough to retain microorganisms, and the filtered liquid emerges sterile. The standard sterilising-grade filter has a nominal pore size of **0.22 micrometres (0.2 µm)**, which is small enough to retain bacteria. Filtration removes organisms rather than killing them, so the filtrate must then be filled under aseptic conditions, which is why sterile filtration is the central solution-sterilising step of aseptic processing (Figure 23.1).

Two cautions are important. First, an ordinary 0.22 μm filter retains bacteria but does not reliably retain viruses or, by itself, remove dissolved endotoxin, so it assures freedom from bacterial contamination, not from every possible contaminant. Second, the integrity of the filter must be tested (for example by a bubble-point test) before and after use, because a flawed or damaged filter would let organisms through and there is no later step to catch the failure. Sterile filtration is powerful precisely because it works on heat-labile products, but it carries the full burden of aseptic risk.

23.4 Lyophilisation (freeze-drying)

Many of the most valuable parenteral drugs, proteins, peptides, some antibiotics and vaccines, are unstable in aqueous solution and would degrade long before their shelf life if stored as a liquid. **Lyophilisation (freeze-drying)** solves this by removing the water to leave a dry, stable solid that is reconstituted with Water for Injection just before use. The process removes water by **sublimation**, the direct conversion of ice to vapour under low pressure, which avoids exposing the fragile drug to the heat that ordinary evaporative drying would require. It proceeds in three phases (Figure 23.2).

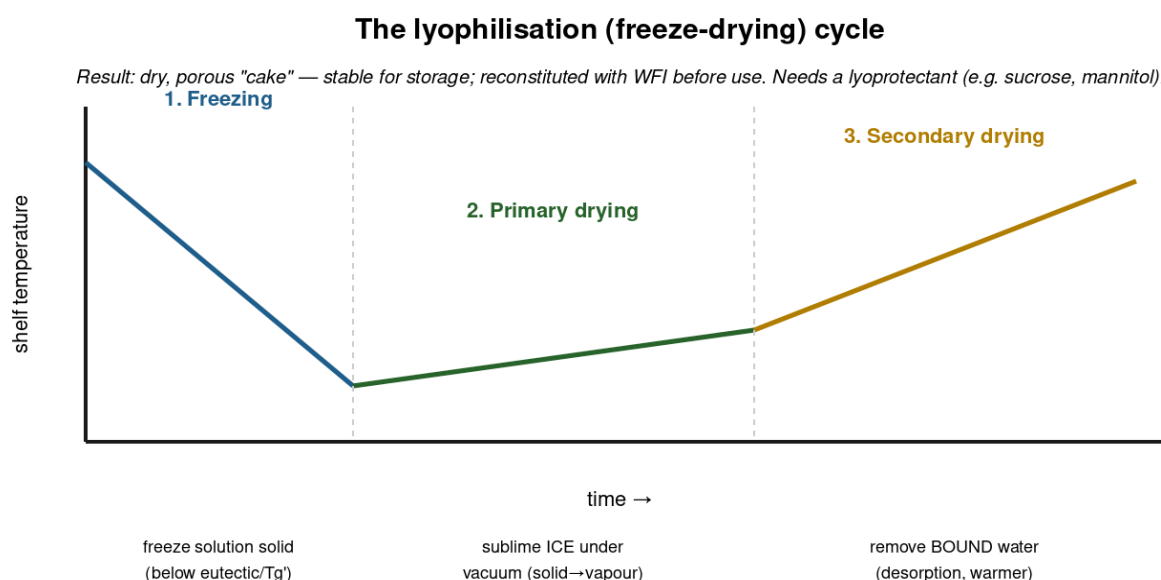


Figure 23.2. The three phases of lyophilisation. In freezing, the solution is cooled until it is fully solid. In primary drying, the pressure is lowered and the ice is removed by sublimation directly from solid to vapour at low temperature. In secondary drying, the temperature is raised to desorb the remaining bound water. The result is a dry, porous cake that is stable on storage and reconstituted with Water for Injection before use; a lyoprotectant such as sucrose or mannitol protects the drug and forms the cake.

In the **freezing** phase, the solution is cooled until it is completely solid, with the drug and excipients locked in a frozen matrix. In **primary drying**, the chamber pressure is reduced to a deep vacuum and gentle heat is applied so that the frozen water sublimates directly to vapour and is removed, while the temperature is kept low enough that the product stays solid; this is the longest phase and removes the bulk of the water. In **secondary drying**, the temperature is raised further to drive off the small amount of water that remains bound to the solid by desorption, bringing the residual moisture down to the low level needed for stability. The product of the cycle is a dry, porous **cake** that dissolves rapidly when WFI is added. A

lyoprotectant (a sugar such as sucrose, trehalose or mannitol) is almost always included; it protects the drug, especially a protein, from the stresses of freezing and drying and gives the cake its structure. Lyophilisation is expensive and slow, but for an otherwise unstable biologic it is often the only way to achieve an acceptable shelf life.

23.5 Environmental control and cleanrooms

Aseptic processing in particular, and sterile manufacturing in general, can only succeed in an environment where airborne and surface contamination is rigorously controlled. This is the function of the **cleanroom**: a manufacturing space in which the air is filtered (through high-efficiency particulate air, HEPA, filters), the airflow, temperature, humidity and pressure are controlled, and the number of airborne particles and microorganisms is held within defined limits. Cleanrooms are held at positive pressure relative to surrounding areas so that air flows outward and contamination cannot drift in, and the most critical operations are protected by **unidirectional (laminar) airflow** that sweeps particles away from the product.

Cleanrooms are classified by how clean their air must be, and two systems are in common use. The **GMP grades A, B, C and D** classify zones by permitted particle and microbial limits, with **Grade A** being the cleanest, reserved for the most critical operations such as the actual filling of an aseptic product and typically provided by a laminar-airflow workstation or isolator, and Grades B, C and D being progressively less stringent surrounding zones. In parallel, the **ISO classification (ISO 14644)** numbers cleanliness classes, with **ISO 5** corresponding broadly to the critical Grade A environment and higher ISO numbers denoting less clean areas. The essential idea, examined more fully in Chapter 26, is that the cleanliness of the environment is matched to the criticality of the operation: the moment the sterile product is exposed, at the point of fill, demands the cleanest air, while support areas can be less stringent. Personnel, who are the largest single source of contamination, are controlled through gowning, training and disciplined movement, and the whole environment is continuously monitored.

Quick-recall box — sterile manufacturing Two routes: terminal sterilisation (sterilise final sealed product; preferred; SAL 10^{-6}) vs aseptic processing (sterilise parts separately, filter solution, fill under control; riskier, no final lethal step). Methods: moist heat (autoclave 121 °C, measured by F_0 , target $\geq \sim 12$ min), dry heat (glass/oils/powders, depyrogenates), radiation (plastics), ethylene oxide (devices), filtration (heat-labile liquids). Sterile filter = 0.22 μm (retains bacteria, not viruses/endotoxin; integrity-tested). Lyophilisation: freezing $\rightarrow$ primary drying (sublimation of ice) $\rightarrow$ secondary drying (desorb bound water); gives stable cake; uses a lyoprotectant. Cleanrooms: HEPA-filtered, positive pressure, laminar flow at the critical point; Grade A (cleanest, = ISO 5) for filling, then B/C/D; people are the main contamination source.

Chapter summary

Sterility is achieved by one of two philosophies. Terminal sterilisation treats the finished, sealed product with a validated lethal process, usually moist heat, and is preferred wherever the product can survive it because the sterilising step acts on the final article. Aseptic processing, used only for heat- or radiation-labile products, sterilises each component separately, filters the solution and assembles everything under controlled conditions, and is inherently riskier because no final lethal step exists. The main sterilisation methods, moist heat, dry heat, radiation, ethylene oxide and filtration, are chosen by what the product tolerates, and their success is quantified by the sterility assurance level, with a target of one

non-sterile unit in a million, and by the F_0 value, which expresses moist-heat lethality as equivalent minutes at 121 °C. Sterile filtration through a 0.22 µm membrane removes bacteria from heat-labile liquids but not viruses or endotoxin, requires integrity testing, and feeds aseptic filling. Lyophilisation stabilises water-sensitive drugs by removing water through sublimation in three phases, freezing, primary drying and secondary drying, leaving a reconstitutable porous cake protected by a lyoprotectant. All of this depends on a controlled environment: HEPA-filtered, positive-pressure cleanrooms with unidirectional airflow at the critical points, classified by GMP Grades A to D or by ISO 14644 with Grade A (ISO 5) reserved for the exposed product at filling, and with personnel managed as the principal source of contamination.

Self-assessment

Short-answer (2 marks each)

1. Distinguish terminal sterilisation from aseptic processing and state which is preferred and why.
2. Define the sterility assurance level and give the standard target.
3. What is the F_0 value and what reference value is commonly used?
4. What pore size defines a sterilising-grade filter, and what does it fail to remove?
5. Name the three phases of lyophilisation.

Long-answer (5–10 marks each)

6. Compare terminal sterilisation and aseptic processing with a decision diagram, and explain why each is used.
7. Describe the lyophilisation cycle and its three phases with a diagram, explaining why the process is used and the role of a lyoprotectant.
8. Explain the principle of cleanroom environmental control, the GMP grade and ISO classification systems, and how cleanliness is matched to the operation.

Multiple-choice

9. The preferred sterilisation approach where the product can tolerate it is: (a) aseptic processing (b) terminal sterilisation (c) filtration only (d) no sterilisation
10. In lyophilisation, ice is removed during primary drying by: (a) melting (b) evaporation (c) sublimation (d) filtration
11. The cleanest GMP cleanroom grade, used for aseptic filling, is: (a) Grade D (b) Grade C (c) Grade B (d) Grade A

Answer pointers. 1. Terminal = sterilise final sealed product (preferred, acts on finished article); aseptic = sterilise parts and assemble under control (riskier). 2. Probability a unit is non-sterile; target 10^{-6} . 3. Equivalent minutes at 121 °C of a moist-heat cycle; commonly ≥ 12 min. 4. 0.22 µm; does not reliably remove viruses or dissolved endotoxin. 5. Freezing, primary drying, secondary drying. 6. Per Figure 23.1. 7. Per Figure 23.2. 8. Per Section 23.5. 9. (b). 10. (c). 11. (d).

References

1. Indian Pharmacopoeia Commission. Indian Pharmacopoeia: Methods of Sterilisation. Ghaziabad: IPC; current edition.

2. ISO 14644-1: Cleanrooms and associated controlled environments — Classification of air cleanliness by particle concentration. Geneva: ISO; 2015.
3. Tang XC, Pikal MJ. Design of freeze-drying processes for pharmaceuticals: practical advice. *Pharm Res.* 2004;21(2):191–200.
4. Akers MJ. *Sterile Drug Products: Formulation, Packaging, Manufacturing and Quality.* New York: CRC Press; 2010.
5. Aulton ME, Taylor KMG, editors. *Aulton's Pharmaceutics.* 6th ed. Edinburgh: Elsevier; 2022.

Chapter 24 - Containers, Closures, and Form-Fill-Seal/Blow-Fill-Seal Technology

Learning outcomes

After working through this chapter the student will be able to:

1. Explain why the container–closure system is part of the product and the requirements it must meet.
2. Describe the pharmacopoeial types of glass and select the appropriate type for a product.
3. Describe elastomeric closures and the properties required of them.
4. Explain extractables and leachables, sorption, and container–closure integrity, and how each is evaluated.
5. Describe Form-Fill-Seal and Blow-Fill-Seal technology and explain why BFS suits preservative-free products.

24.1 The container is part of the product

For a sterile product the container and its closure are not mere packaging; they are functional components of the product itself (Figure 24.1). The container–closure system must keep the product sterile for its entire shelf life, protect it from light, moisture and oxygen, not react with or contaminate the contents, allow the product to be withdrawn and administered, and maintain a hermetic seal. A failure of any of these, a glass that leaches alkali into the solution, a stopper that sheds particles, a seal that admits microorganisms, can ruin a product that was perfectly made. The selection and evaluation of containers and closures is therefore a core part of sterile-product development.

The container–closure system

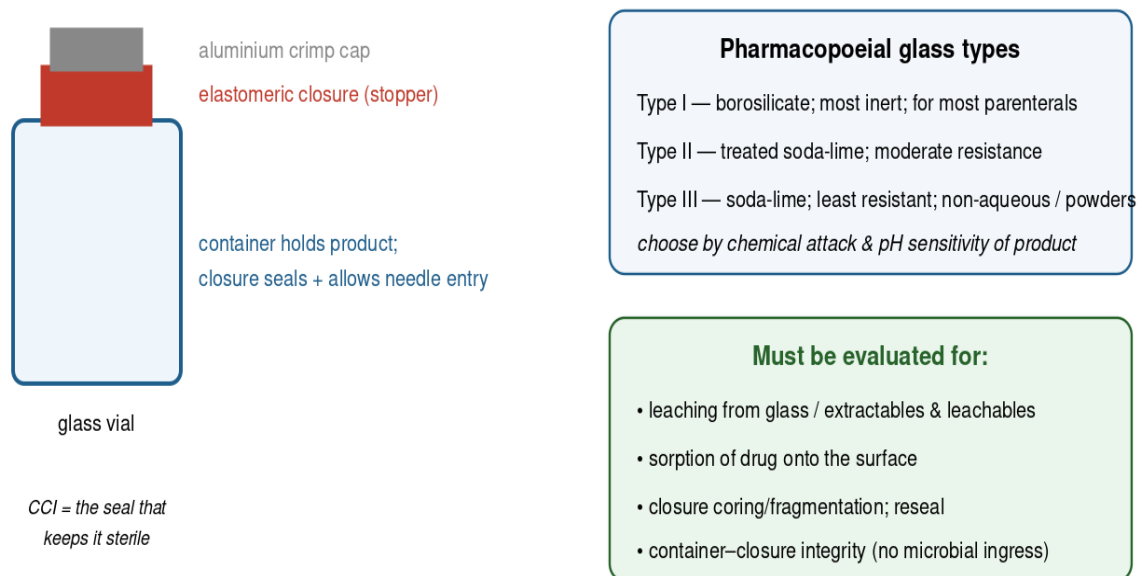


Figure 24.1. The container–closure system, shown for a vial: the glass container holds the product, the elastomeric closure seals it while allowing needle entry, and the aluminium crimp cap secures the closure. The pharmacopoeial glass types (I, II, III) differ in chemical resistance, and the system as a whole must be evaluated for leaching, sorption, closure coring and reseal, and above all for container–closure integrity, the seal that maintains sterility.

24.2 Glass containers and the pharmacopoeial glass types

Glass is the traditional and still dominant material for parenteral containers because it is inert, impermeable, transparent (allowing inspection) and able to withstand heat sterilisation. Its drawbacks are that it is breakable, heavy, and not entirely inert: glass can release small amounts of alkali into an aqueous product, raising the pH, and certain solutions can attack the glass surface. To manage this, pharmacopoeias classify glass into types by chemical resistance.¹

Type I is borosilicate glass, the most chemically resistant and most inert, with low alkali release; it is suitable for almost all parenteral products, including sensitive and alkaline solutions, and is the default choice for injections. **Type II** is a soda-lime glass that has been surface-treated (dealkalised) to improve its resistance, suitable for many aqueous products of moderate sensitivity, particularly acidic and neutral ones. **Type III** is ordinary soda-lime glass with the least chemical resistance, used for non-aqueous products and for powders, where attack on the glass is not a concern. The formulator selects the glass type by how chemically aggressive and pH-sensitive the product is: a borosilicate Type I for a sensitive injectable, a cheaper Type III where the contents will not interact with the glass.

24.3 Elastomeric closures

The closure of a vial is an **elastomeric (rubber) stopper**, and it has demanding and partly conflicting jobs. It must seal the container hermetically to maintain sterility and prevent

leakage; it must be soft and elastic enough to be pierced by a needle and then to **reseal** when the needle is withdrawn (especially for multidose vials); it must withstand sterilisation; and it must be compatible with the product, neither leaching harmful substances into it nor adsorbing the drug from it. Elastomeric closures are made from formulated rubber compounds containing many additives, and two specific problems must be controlled. **Coring** or **fragmentation** is the shedding of small pieces of the stopper into the product when it is pierced by a needle, which would introduce particulates; closures are designed and tested to minimise it. And the rubber's additives are a source of leachables (Section 24.4). Many modern closures are coated (for example with a fluoropolymer film) to reduce both leaching and interaction with the product.

24.4 Extractables, leachables, sorption and integrity

Four interactions between the product and its container–closure system must be evaluated, and they are a standard subject of examination.

Extractables and leachables are chemical species that migrate from the container or closure into the product. **Extractables** are compounds that can be forced out of the packaging material under exaggerated laboratory conditions (aggressive solvents, heat), representing what the material could potentially release; **leachables** are the compounds that actually migrate into the specific product under real storage conditions, and are a subset of the extractables. Both matter because a substance leached from glass, rubber or plastic into an injectable is then administered to the patient, so the packaging must be shown not to release harmful amounts of anything.

Sorption is the reverse problem: the loss of drug (or preservative) from the product by adsorption onto, or absorption into, the container or closure surface. For a low-dose or protein drug this can remove a meaningful fraction of the dose, reducing potency, so it too must be assessed.

Container–closure integrity (CCI) is the most important property of all, because it is the property that maintains sterility. CCI is the ability of the sealed system to prevent the ingress of microorganisms (and the loss of contents or, where relevant, of headspace gas) over the product's shelf life. A product can be perfectly sterile when sealed and become contaminated if the seal leaks, so CCI must be demonstrated, by tests such as dye-ingress, vacuum or pressure decay, helium leak, or high-voltage methods. The container–closure–integrity test is, in effect, the proof that the package will keep the product sterile until use, and it is a required part of the evaluation of any sterile container system.

24.5 Form-Fill-Seal and Blow-Fill-Seal technology

A powerful way to reduce contamination risk is to minimise the human handling and environmental exposure of the container, and this is the idea behind the integrated “fill-seal” technologies.

In **Form-Fill-Seal (FFS)**, the container is formed, filled with product and sealed in one continuous, automated operation on a single machine, rather than using pre-made containers handled separately. Reducing the steps and the exposure reduces the opportunity for contamination. FFS is widely used for sachets, blisters and some liquid containers.

Blow-Fill-Seal (BFS) is the most important form of this approach for sterile liquids and is examined in detail (Figure 24.2). In BFS, a single machine performs three steps in immediate succession inside a controlled zone. In the **blow** step, a tube of molten plastic (typically

polyethylene or polypropylene) is extruded and blown into a mould to **form** the container; because the plastic emerges hot from the extruder, the freshly formed container is essentially sterile inside. In the **fill** step, the sterile product is filled into the still-warm container immediately, through the same machine. In the **seal** step, the container is hermetically sealed, all before it leaves the machine.

Blow-Fill-Seal (BFS): form, fill and seal in one machine

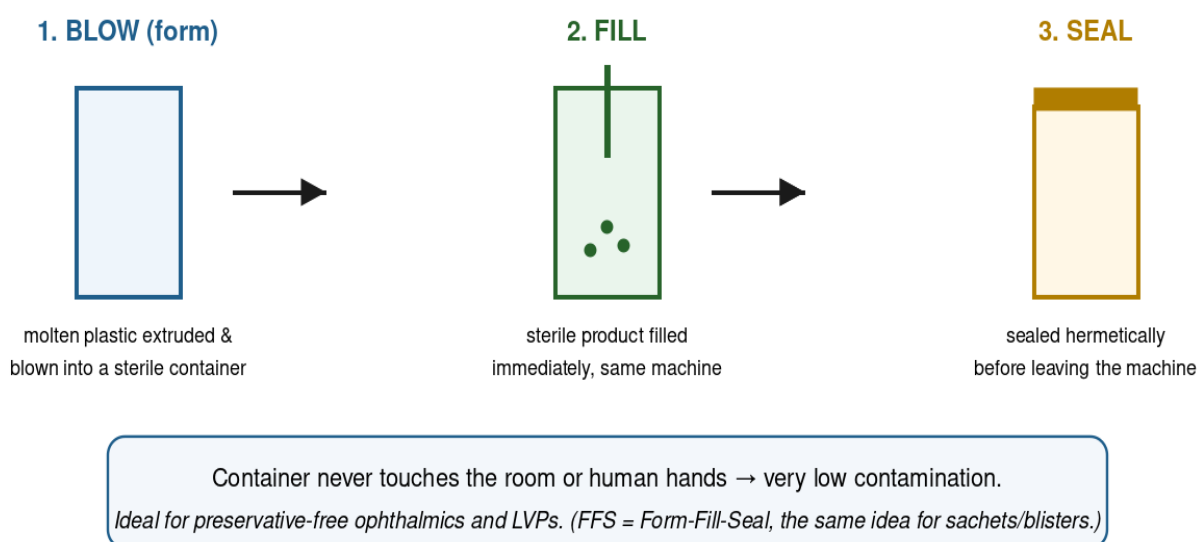


Figure 24.2. The Blow-Fill-Seal process. A single machine blows a tube of molten plastic into a sterile container (form), fills it with sterile product immediately (fill), and seals it hermetically (seal), all in one continuous operation inside a controlled zone. Because the container is formed, filled and sealed without contact with the room or human hands, contamination risk is very low, which makes BFS ideal for preservative-free ophthalmics and for large-volume parenterals. FFS (Form-Fill-Seal) applies the same integrated principle to sachets and other formats.

The decisive advantage of BFS is that the container is **formed, filled and sealed without ever being touched by human hands or exposed to the room environment**, which gives an extremely low contamination risk, an example of the aseptic-by-design philosophy. This makes BFS especially valuable for two product types met elsewhere in this unit. It suits **large-volume parenterals**, which must be preservative-free and made to a high assurance of sterility. And it is the enabling technology for **preservative-free ophthalmic products**: because each BFS unit is made sterile and used once, the eye drop needs no preservative, which avoids the ocular-surface toxicity that preservatives can cause on chronic use, as Chapter 25 discusses. BFS is thus a technology that directly serves the goal of safer, preservative-free sterile products.

Quick-recall box — containers and closures Container–closure system is part of the product: keeps it sterile, protects, does not react. Glass types: I borosilicate (most inert, default for injectables), II treated soda-lime (moderate), III soda-lime (least resistant, non-aqueous/powders). Elastomeric closure: seals, allows needle entry, must reseal, low coring/fragmentation; often coated. Four evaluations: extractables (potential) vs leachables (actual migration in), sorption (drug lost to the

surface), and CCI (container–closure integrity = the seal that maintains sterility). FFS/BFS: form-fill-seal in one machine; BFS blows-fills-seals plastic with no human contact → very low contamination → ideal for LVPs and preservative-free ophthalmics.

Chapter summary

For a sterile product the container and closure are functional parts of the product, required to maintain sterility for the whole shelf life, protect the contents, avoid reacting with them, and permit withdrawal. Glass dominates because it is inert, impermeable and heat-resistant, and pharmacopoeias grade it by chemical resistance into Type I borosilicate (most inert, the default for injectables), Type II treated soda-lime (moderate resistance), and Type III soda-lime (least resistant, for non-aqueous products and powders). The elastomeric closure must seal hermetically, be pierceable and reseal for multidose use, withstand sterilisation, and resist shedding fragments (coring), with coated rubbers now common. Four product–package interactions must be evaluated: extractables, the compounds a material can release under forcing conditions, and the leachables that actually migrate into the product; sorption, the loss of drug onto the package surface; and above all container–closure integrity, the seal’s ability to keep microorganisms out, which is what preserves sterility and is proven by leak tests. Finally, integrated Form-Fill-Seal and especially Blow-Fill-Seal technologies form, fill and seal a container in one automated operation with no human contact or room exposure, giving very low contamination risk and making BFS the technology of choice for preservative-free large-volume parenterals and ophthalmics.

Self-assessment

Short-answer (2 marks each)

1. Why is the container–closure system considered part of the product?
2. Which glass type is most chemically resistant, and for what is it used?
3. What is coring, and why does it matter for an elastomeric closure?
4. Distinguish extractables from leachables.
5. What is container–closure integrity and why is it the most important package property?

Long-answer (5–10 marks each)

6. Describe the pharmacopoeial glass types and the properties required of elastomeric closures, with a diagram of the container–closure system.
7. Explain extractables and leachables, sorption, and container–closure integrity, and how each is evaluated.
8. Describe Blow-Fill-Seal technology with a diagram and explain why it is ideal for preservative-free ophthalmic and large-volume products.

Multiple-choice

9. The most inert pharmacopoeial glass, used for most parenterals, is: (a) Type I borosilicate (b) Type II (c) Type III (d) soda-lime
10. The package property that maintains sterility over shelf life is: (a) sorption (b) extractables (c) container–closure integrity (d) colour
11. BFS is especially valued because the container is: (a) made of glass (b) formed, filled and sealed without human contact (c) reusable (d) preserved chemically

Answer pointers. 1. It maintains sterility and can react with/contaminate the contents, affecting the product. 2. Type I borosilicate; most parenterals including sensitive/alkaline ones. 3. Shedding of stopper fragments into the product on needle entry; introduces particulates. 4. Extractables = potential release under forcing conditions; leachables = actual migration into the product in use. 5. The seal's ability to prevent microbial ingress; it preserves sterility. 6. Per Figure 24.1. 7. Per Section 24.4. 8. Per Figure 24.2 and Section 24.5. 9. (a). 10. (c). 11. (b).

References

1. Indian Pharmacopoeia Commission. Indian Pharmacopoeia: Glass containers and Rubber closures for parenteral preparations. Ghaziabad: IPC; current edition.
2. United States Pharmacopeia. General chapters on packaging, glass containers and elastomeric closures. Rockville (MD): USP.
3. Akers MJ. Sterile Drug Products: Formulation, Packaging, Manufacturing and Quality. New York: CRC Press; 2010.
4. Aulton ME, Taylor KMG, editors. Aulton's Pharmaceutics. 6th ed. Edinburgh: Elsevier; 2022.

Chapter 25 - Ophthalmic Products

Learning outcomes

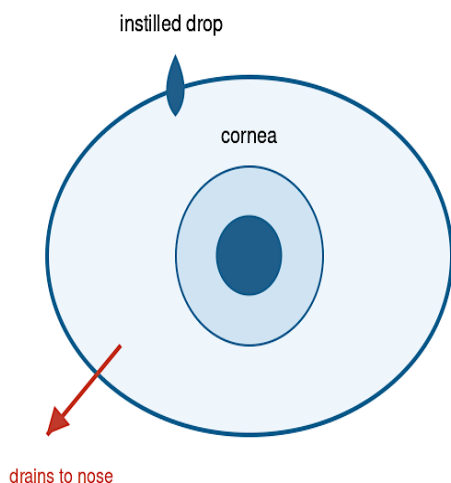
After working through this chapter the student will be able to:

1. Explain the physiological factors that make ophthalmic delivery inefficient and the resulting low ocular bioavailability.
2. State the essential quality requirements of an ophthalmic product: sterility, tonicity, pH and comfort.
3. Describe the main ophthalmic dosage forms, solutions, suspensions, gels, ointments and inserts, and their relative residence times.
4. Explain residence-time enhancers, including viscosity agents, in-situ gels and mucoadhesives.
5. Explain why preservative-free systems are preferred for chronic use and how BFS enables them.

25.1 The eye as a delivery target

The eye is protected against foreign substances by an efficient set of clearance mechanisms, and those same mechanisms make it a difficult target for drug delivery. An instilled eye drop faces the problem first raised in Chapter 1 and developed in Chapter 12: it is cleared from the eye surface within minutes, so that only a very small fraction of the dose is absorbed (Figure 25.1).

Why eye drops are inefficient (and how to fix it)



Losses (within minutes):

- blinking & reflex tearing wash drop away
- nasolacrimal drainage carries it off
- small eye capacity ($\sim 30 \mu\text{L}$) vs drop ($\sim 50 \mu\text{L}$)

Result: typically < 5% of dose absorbed

Fixes (raise residence time):

- $\uparrow$ viscosity (cellulose, hyaluronic acid)
- in-situ gels (Ch 12) · mucoadhesives
- ointments, gels, inserts (longer contact)

Same problem first seen in Ch 1 & Ch 12 — the eye clears drops fast; delivery design slows the clearance.

Figure 25.1. Why eye drops are inefficient. A drop is washed away within minutes by blinking and reflex tearing and is drained through the nasolacrimal duct to the nose; the

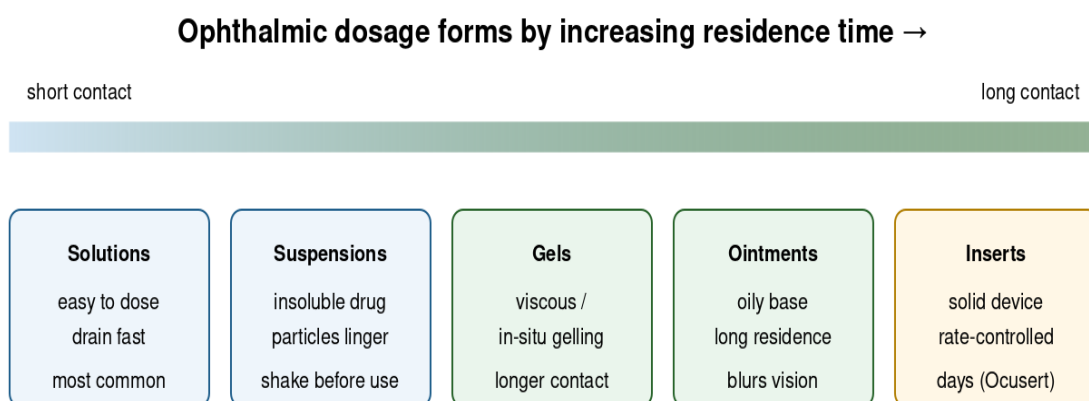
eye's small fluid capacity (around 30 μL) cannot even hold a standard drop (around 50 μL), so the excess overflows. The result is that typically less than 5% of the instilled dose is absorbed. Delivery design raises residence time by increasing viscosity, gelling in place, using mucoadhesives, or using ointments, gels and inserts.

Several factors combine. **Blinking and reflex tearing** sweep the drop off the cornea almost at once. **Nasolacrimal drainage** carries the fluid down the tear duct to the nose, where some is even absorbed systemically (a route of unwanted systemic side effects from eye drops). And the eye's **fluid capacity is tiny**, roughly 30 microlitres, while a standard drop is around 50 microlitres, so much of the drop simply overflows and is lost. The combined effect is that the **ocular bioavailability** of a conventional eye drop is typically very low, often quoted as under 5% of the instilled dose. The whole of ophthalmic formulation is, in large part, a campaign to overcome this rapid loss.

25.2 Essential requirements of an ophthalmic product

Whatever its form, an ophthalmic product must meet a set of requirements driven by the sensitivity of the eye (Figure 25.2).

It must be **sterile**, because the eye, especially a damaged or post-surgical eye, is vulnerable to infection that can threaten sight; an ophthalmic product is held to the same sterility standard as a parenteral. It must be **isotonic** with the tear fluid, for the same osmotic reasons as an injection (Chapter 22), since a strongly hypotonic or hypertonic drop is irritating; the reference, as for parenterals, is around 0.9% sodium chloride equivalent. It must be **buffered to a comfortable pH**, ideally near the pH of tears (about 7.4), balanced against the pH at which the drug is stable and soluble; the eye tolerates a modest range, but a strongly acidic or alkaline drop causes stinging and reflex tearing that washes the drug away faster. It should have appropriate **viscosity**, both for comfort and, as Section 25.3 shows, to prolong contact. And it must be **comfortable and non-irritant**, because an irritating product triggers exactly the tearing and blinking that clear it from the eye, so comfort is not a luxury but a determinant of efficacy.



All ophthalmics must be: sterile · isotonic · buffered near tear pH (~7.4) · comfortable

Preservative-free (single-dose / BFS) preferred for chronic use to avoid ocular-surface toxicity.

Figure 25.2. *Ophthalmic dosage forms arranged by increasing residence time, from easily dosed but fast-draining solutions, through suspensions and viscous or in-situ gels, to long-residence ointments and rate-controlled inserts. Every ophthalmic form, whatever its*

residence time, must be sterile, isotonic, buffered near tear pH and comfortable, and preservative-free systems are preferred for chronic use.

25.3 Dosage forms and residence-time enhancers

Ophthalmic dosage forms can be arranged along a spectrum of increasing contact time with the eye (Figure 25.2), and moving along that spectrum is the principal way to raise bioavailability.

Solutions are the simplest and most common form: easy to instil and accurately dosed, but cleared fastest. **Suspensions** are used when the drug is insoluble; the suspended particles can linger slightly longer in the eye, and they provide a reservoir of drug, but they must be properly resuspended before use. **Gels**, whether simply viscous or in-situ gelling, increase contact time substantially. **Ointments**, based on an oily vehicle such as a petrolatum–mineral oil base, give the longest contact of the conventional forms and are often used at night, with the trade-off that they blur vision and feel greasy. **Inserts** are solid devices placed in the eye that release drug at a controlled rate over hours to days; the rate-controlled Ocusert (pilocarpine) of Chapter 12 is the classic example, giving the longest and most controlled delivery of all.

The strategies for prolonging residence time, the **residence-time enhancers**, are the practical core of the chapter and tie directly back to earlier units. **Viscosity-increasing agents**, such as the cellulose derivatives (hydroxypropyl methylcellulose) and hyaluronic acid, thicken the formulation so it resists drainage and clings to the eye longer. **In-situ gelling systems** (Chapter 12) are instilled as easily dosed liquid drops that gel on contact with the eye, triggered by the temperature, pH or ion content of the tear film, combining the convenience of a drop with the residence of a gel; this is one of the most elegant solutions to the ocular-delivery problem. **Mucoadhesive polymers** (Chapter 3), such as chitosan and carbomers, bind to the mucin of the tear film and hold the formulation against the eye surface. In each case the goal is the same: keep the drug on the eye longer so that more of it is absorbed before it is cleared.

25.4 Preservatives and the case for preservative-free systems

Multidose eye drops, like multidose injections, have traditionally contained an **antimicrobial preservative** to protect the contents against contamination introduced during repeated use, the most common being **benzalkonium chloride**. Benzalkonium chloride is an effective preservative and also has a mild permeation-enhancing effect, but it carries a real cost on the ocular surface: with chronic, frequent use, as in lifelong glaucoma therapy, it can damage the corneal and conjunctival surface, causing irritation, dryness and inflammation. For a patient using preserved drops several times a day for years, this cumulative surface toxicity is a genuine clinical problem.

This is why **preservative-free ophthalmic systems** are increasingly preferred for chronic use. The challenge is to protect a product against contamination without a preservative, and the solution is to make it **single-use** or to package it so that the contents cannot be contaminated. **Blow-Fill-Seal (BFS) technology** (Chapter 24) is the key enabler: BFS produces small, single-dose plastic units that are formed, filled and sealed sterile without human contact, so each unit is sterile and used once and needs no preservative. The connection across the unit is direct: the manufacturing technology of Chapter 24 makes possible the preservative-free ophthalmic product that the sensitivity of the chronically treated eye demands. Preservative-free unit-dose drops, and preservative-free multidose

bottles with special non-return valve systems, are now widely used for dry-eye and glaucoma therapy for exactly this reason.

Quick-recall box — ophthalmics Problem: blinking, tearing, nasolacrimal drainage and tiny eye capacity ($\sim 30\ \mu\text{L}$ vs $\sim 50\ \mu\text{L}$ drop) clear a drop in minutes $\rightarrow$ typically $< 5\%$ absorbed. Every ophthalmic must be sterile, isotonic ($\sim 0.9\%$ NaCl equivalent), buffered near tear pH (~ 7.4), and comfortable (discomfort $\rightarrow$ tearing $\rightarrow$ faster loss). Forms by residence time: solution $<$ suspension $<$ gel $<$ ointment $<$ insert. Residence-time enhancers: viscosity agents (HPMC, hyaluronic acid), in-situ gels (temp/pH/ion), mucoadhesives (chitosan, carbomer). Preservatives: benzalkonium chloride common but toxic to the surface on chronic use $\rightarrow$ preservative-free single-dose / BFS preferred for glaucoma and dry eye.

Chapter summary

The eye clears an instilled drop within minutes through blinking, reflex tearing and nasolacrimal drainage, and its fluid capacity is far smaller than a standard drop, so conventional eye drops deliver typically less than 5% of the dose, which makes prolonging contact the central aim of ophthalmic formulation. Every ophthalmic product, regardless of form, must be sterile because the eye is vulnerable to sight-threatening infection, isotonic with tears to avoid irritation, buffered near the tear pH of about 7.4 balanced against drug stability, suitably viscous, and above all comfortable, since discomfort provokes the very tearing and blinking that clear the drug away. The dosage forms range along a spectrum of increasing residence time from solutions, the simplest and fastest-cleared, through suspensions and gels to ointments and finally rate-controlled inserts such as Ocusert, and moving along this spectrum raises bioavailability. The practical residence-time enhancers are viscosity-increasing agents like cellulose derivatives and hyaluronic acid, in-situ gelling systems that are dosed as liquids and gel on the eye, and mucoadhesive polymers that bind the tear-film mucin. Finally, while multidose drops have relied on preservatives such as benzalkonium chloride, that agent damages the ocular surface on the chronic frequent use typical of glaucoma, so preservative-free single-dose systems are now preferred, made possible by Blow-Fill-Seal technology that produces sterile, contact-free, single-use units, tying the manufacturing advances of the previous chapter directly to safer ophthalmic care.

Self-assessment

Short-answer (2 marks each)

1. Give three reasons why eye drops have low ocular bioavailability.
2. State four essential requirements of an ophthalmic product.
3. Why is comfort a determinant of ophthalmic efficacy rather than only a matter of patient preference?
4. Name three residence-time enhancers and how each works.
5. Why are preservative-free systems preferred for chronic ophthalmic use?

Long-answer (5–10 marks each)

6. Explain the physiological factors limiting ocular drug delivery and the requirements an ophthalmic product must meet, with a diagram.
7. Describe the ophthalmic dosage forms in order of residence time and the residence-time enhancers used to improve bioavailability.

8. Discuss the role of preservatives in ophthalmic products, the problem with benzalkonium chloride in chronic use, and how BFS enables preservative-free delivery.

Multiple-choice

9. The approximate ocular bioavailability of a conventional eye drop is: (a) ~50% (b) ~25% (c) < 5% (d) ~90%
10. An in-situ gelling eye drop is advantageous because it is: (a) a solid (b) dosed as a liquid but gels on the eye (c) preserved (d) hypertonic
11. The preservative associated with chronic ocular-surface toxicity is: (a) sucrose (b) benzalkonium chloride (c) EDTA (d) sodium chloride

Answer pointers. 1. Blinking/tearing, nasolacrimal drainage, small eye capacity vs drop volume. 2. Sterile, isotonic, buffered near tear pH, comfortable (also appropriate viscosity). 3. Discomfort triggers tearing/blinking that clears the drug, lowering absorption. 4. Viscosity agents (HPMC/hyaluronic acid), in-situ gels (gel on the eye), mucoadhesives (bind mucin). 5. Preservatives like benzalkonium chloride cause cumulative ocular-surface toxicity on frequent chronic use. 6. Per Figure 25.1. 7. Per Figure 25.2 and Section 25.3. 8. Per Section 25.4. 9. (c). 10. (b). 11. (b).

References

1. Gaudana R, Ananthula HK, Parenky A, Mitra AK. Ocular drug delivery. AAPS J. 2010;12(3):348–60.
2. Baranowski P, Karolewicz B, Gajda M, Pluta J. Ophthalmic drug dosage forms: characterisation and research methods. ScientificWorldJournal. 2014;2014:861904.
3. Indian Pharmacopoeia Commission. Indian Pharmacopoeia: Ophthalmic Preparations. Ghaziabad: IPC; current edition.
4. Aulton ME, Taylor KMG, editors. Aulton's Pharmaceutics. 6th ed. Edinburgh: Elsevier; 2022.

Chapter 26 - Quality Control and Compliance

Learning outcomes

After working through this chapter the student will be able to:

1. Distinguish sterility from freedom from pyrogens and explain why both must be tested.
2. Describe the sterility test and the bacterial endotoxin (LAL) test, including the three LAL methods.
3. State the particulate-matter limits for small- and large-volume parenterals and describe how they are measured.
4. Describe the clarity and extractable-volume tests and the key labelling requirements.
5. Explain cGMP, cleanroom grades and environmental monitoring as the system that assures sterile-product quality.

26.1 Proving a sterile product is fit to use

The preceding chapters built a sterile product; this chapter proves it. Quality control of parenterals rests on a set of pharmacopoeial tests, each targeting one of the absolute requirements of Chapter 20, sterility, freedom from pyrogens, and freedom from particulates, together with tests of identity, strength, clarity and deliverable volume (Figure 26.1). Underpinning all of them is the compliance framework of current Good Manufacturing Practice, which ensures the product was made under conditions that make a passing test trustworthy. The chapter takes the tests first, then the compliance framework.

Quality-control tests for parenteral products

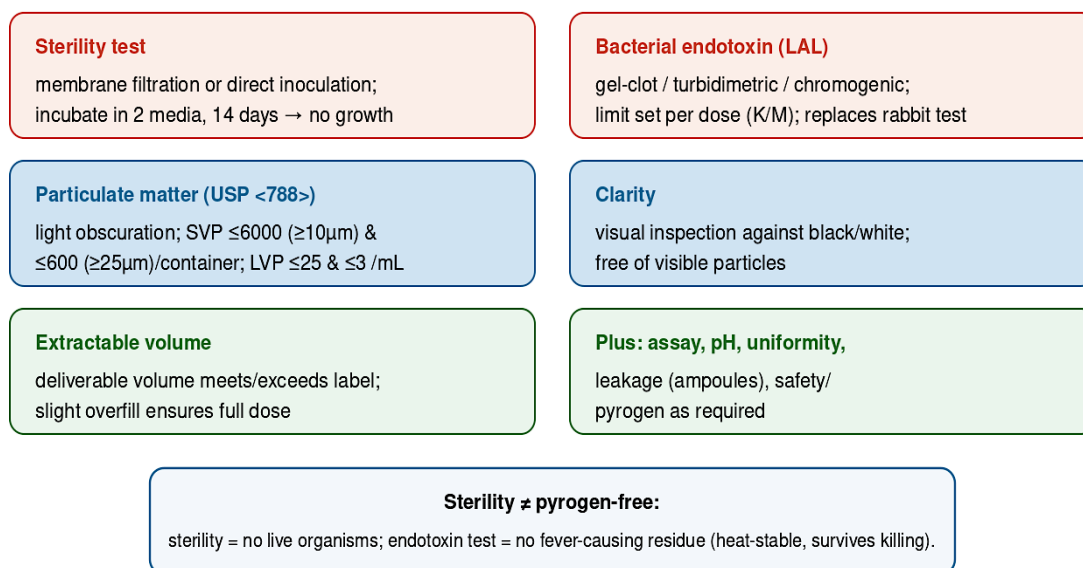


Figure 26.1. The quality-control tests for parenteral products: the sterility test, the bacterial endotoxin (LAL) test, the particulate-matter test (USP <788>), the clarity test, the extractable-volume test, and supporting tests of assay, pH and uniformity. The footnote captures the essential distinction that sterility and pyrogen-freedom are separate properties

requiring separate tests, because endotoxin is heat-stable and survives the killing of the organisms that produced it.

26.2 Sterility versus pyrogens: two distinct properties

A point that must be fixed before the tests make sense is that **sterility and freedom from pyrogens are different properties**. **Sterility** means the absence of viable (living) microorganisms. **Pyrogens** are fever-producing substances, and the most important are **bacterial endotoxins**, fragments of the cell walls of Gram-negative bacteria (lipopolysaccharide). The crucial fact is that endotoxin is **heat-stable**: it survives the heat that kills the bacteria. A product can therefore be perfectly sterile, every organism dead, and still contain enough endotoxin from those dead organisms to cause a dangerous fever when injected. Sterility and pyrogen-freedom must therefore be assured and tested separately, which is why there are two distinct tests below, and why endotoxin is controlled all the way back at the water stage (Chapter 21).

26.3 The sterility test

The **sterility test** demonstrates the absence of viable microorganisms in the product. Pharmacopoeial methods use two approaches.¹ The preferred is **membrane filtration**, in which the product is filtered through a sterilising membrane that retains any organisms, and the membrane is then incubated in growth media; this suits filterable solutions and is the standard choice. The alternative is **direct inoculation**, in which the product is added directly to the growth media, used for products that cannot be filtered. In both cases the inoculated media are incubated, typically for **14 days**, in two media chosen to grow a broad range of aerobic and anaerobic bacteria and fungi, and the test passes if no growth appears.

The sterility test has an inherent and important limitation: it is a test on a **sample** of the batch, not on every unit, so a passing test gives statistical assurance, not absolute proof, that the whole batch is sterile. A small number of contaminated units could be missed by sampling. This is precisely why sterility cannot be relied upon to be “tested into” a product and must instead be **built in** by a validated sterilisation process and GMP (the same philosophy as QbD in Chapter 14); the sterility test confirms the process worked, it does not by itself create confidence in an uncontrolled process.

26.4 The bacterial endotoxin test (LAL)

Pyrogen testing is now done chiefly by the **bacterial endotoxin test**, which uses **Limulus Amebocyte Lysate (LAL)**, a reagent derived from the blood of the horseshoe crab that clots or reacts in the presence of bacterial endotoxin.² The LAL test has largely replaced the older **rabbit pyrogen test** (in which a rise in the body temperature of injected rabbits signalled pyrogens), because LAL is more sensitive, quantitative, reproducible and avoids animal use. Three LAL methods are used: the **gel-clot** method, in which endotoxin causes the lysate to form a firm gel (a simple positive/negative or semi-quantitative endpoint); the **turbidimetric** method, which measures the increasing turbidity as the reaction proceeds; and the **chromogenic** method, which measures colour released from a substrate cleaved during the reaction. The endotoxin limit for a given product is calculated from the maximum dose the patient may receive, so that the permitted endotoxin per dose stays below the threshold that causes a febrile reaction, and the product passes if its measured endotoxin is below that limit. This is the same kind of limit applied to Water for Injection (0.25 EU/mL) in Chapter 21.

26.5 Particulate matter, clarity and extractable volume

Particulate matter in injections, undissolved particles that could lodge in blood vessels, is controlled by **USP <788>** and the corresponding IP test.³ The primary method is **light obscuration**, in which the solution is passed through a sensor and particles are counted and sized by the light they block; a microscopic method serves as a second-stage backup. The limits distinguish small- from large-volume products. For **small-volume parenterals** the light-obscuration limits are not more than **6000 particles of 10 µm or larger per container** and not more than **600 particles of 25 µm or larger per container**. For **large-volume parenterals** the limits are expressed per millilitre: not more than **25 particles of 10 µm or larger per mL** and not more than **3 particles of 25 µm or larger per mL**. These numbers are worth memorising for examinations.

Clarity is assessed by **visual inspection** of each unit against both black and white backgrounds under defined illumination, to detect visible particles; injections must be essentially free of visible particulate matter, a check distinct from the sub-visible particulate count above.

Extractable (deliverable) volume confirms that the container actually delivers the labelled dose. Because some liquid is always lost in wetting the container and the syringe, containers are filled with a slight **overfill** above the nominal volume, and the extractable-volume test verifies that the volume that can be withdrawn meets or exceeds the labelled amount, so the patient receives a full dose.

Supporting tests, of assay (drug content), pH, uniformity of dosage units, and, for ampoules, leakage, complete the release testing.

26.6 Labelling and documentation

The label of a parenteral product must carry the information needed for safe use, including the name and strength of the drug, the volume, the route(s) of administration, the names and amounts of added substances such as preservatives, the storage conditions, the expiry date and batch number, and any warnings. For products where the route matters critically, the label must make the route unmistakable; an intrathecal product, for instance, must be clearly distinguished to prevent a fatal wrong-route error. **Documentation** is itself a GMP requirement: every step of manufacture and testing is recorded in batch records, so that the history of any batch can be reconstructed and so that a passing result is traceable and trustworthy, the principle that “if it was not documented, it was not done” (developed in quality-assurance courses).

26.7 cGMP, cleanrooms and environmental monitoring

The tests above are only meaningful if the product was made under conditions that prevent contamination in the first place, which is the role of **current Good Manufacturing Practice (cGMP)**. cGMP is the body of regulations and principles governing how medicines are manufactured to assure their quality, covering premises, equipment, personnel, materials, processes, documentation and quality systems. For sterile products, the most distinctive cGMP requirements concern the manufacturing environment.

Sterile manufacturing takes place in classified **cleanrooms**, introduced in Chapter 23 and detailed here (Figure 26.2). The **GMP grades A, B, C and D** classify zones by their permitted airborne particle and microbial limits. **Grade A** is the cleanest and is the local zone for the most critical operations, the actual exposure and filling of the sterile product, with

essentially no tolerated viable contamination; it corresponds broadly to **ISO 5** and is provided by unidirectional (laminar) airflow, a laminar-flow workstation or an isolator. **Grade B** is the clean background environment surrounding a Grade A zone in aseptic processing. **Grades C and D** are progressively less stringent zones for earlier and less critical preparation steps. The cleaner zones are nested inside the less clean ones, so contamination decreases as one approaches the exposed product.

Cleanroom grades and environmental monitoring

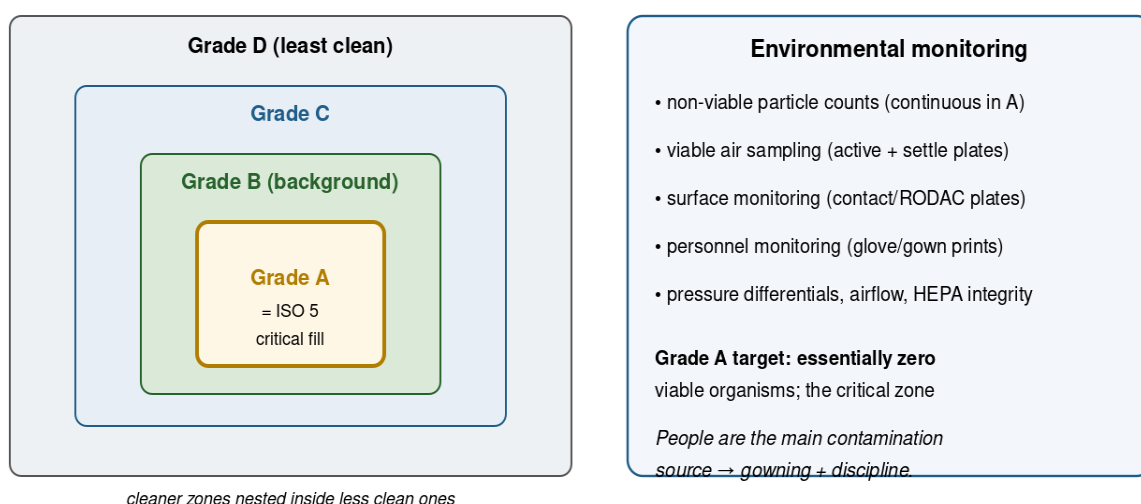


Figure 26.2. Cleanroom grades and environmental monitoring. The GMP grades A to D classify zones by cleanliness, with the cleanest Grade A (broadly ISO 5) reserved for the critical filling operation and nested inside progressively less clean Grade B, C and D zones. Environmental monitoring continuously checks non-viable particle counts, viable air and surface contamination, personnel, and the airflow, pressure and HEPA-filter parameters that maintain the classification. Because people are the principal contamination source, gowning and discipline are central.

These conditions are continuously verified by **environmental monitoring**, the routine measurement of contamination in the manufacturing environment. It includes counting **non-viable particles** in the air (continuously in the critical Grade A zone), sampling for **viable microorganisms** in the air (by active air samplers and settle plates) and on **surfaces** (by contact plates), monitoring **personnel** (glove and gown sampling), and checking the engineering parameters, room pressure differentials, airflow patterns and HEPA-filter integrity, that maintain the classification. Because **personnel are the single largest source of contamination** in a cleanroom, shedding particles and organisms constantly, their control through gowning, training and disciplined behaviour is central, and personnel monitoring is a key part of the programme. Environmental monitoring closes the loop: it provides ongoing evidence that the controlled environment which makes a sterile product possible is actually being maintained, so that the release tests of this chapter can be trusted.

Quick-recall box — QC and compliance Sterility $\neq$ pyrogen-free (endotoxin is heat-stable, survives killing $\rightarrow$ test both). Sterility test: membrane filtration (preferred) or direct inoculation; 2 media, 14 days; sample-based, so build sterility in, don't test it in. Endotoxin: LAL (gel-clot / turbidimetric / chromogenic), replaced rabbit test; limit per dose. Particulates (USP <788>, light obscuration):

SVP ≤ 6000 ($\geq 10\mu\text{m}$) & ≤ 600 ($\geq 25\mu\text{m}$) per container; LVP ≤ 25 & ≤ 3 per mL. Clarity: visual, no visible particles. Extractable volume: overfill ensures full dose. Labelling: drug, strength, volume, route, added substances, storage, expiry, batch. cGMP cleanrooms: Grade A (=ISO 5, critical fill) nested in B/C/D; environmental monitoring of particles, viables, surfaces, people; personnel are the main contamination source.

Chapter summary

Quality control proves that a sterile product meets the absolute requirements set out at the start of the unit, and it begins from the distinction that sterility, the absence of living organisms, and freedom from pyrogens are separate properties, because endotoxin is heat-stable and survives the killing of the bacteria that made it, so both must be tested separately. The sterility test, by membrane filtration or direct inoculation with a fourteen-day incubation in two media, confirms the absence of viable organisms but, being sample-based, gives statistical assurance only, which is why sterility must be built in by a validated process rather than tested in. The bacterial endotoxin test using Limulus Amebocyte Lysate, in its gel-clot, turbidimetric or chromogenic forms, has largely replaced the rabbit pyrogen test and measures endotoxin against a dose-based limit. Particulate matter is controlled by light obscuration under USP <788> to limits of 6000 and 600 particles per container at 10 and 25 μm for small-volume products and 25 and 3 particles per millilitre for large-volume products, while clarity is checked visually and extractable volume confirms a full deliverable dose through deliberate overfill. Labelling must convey drug, strength, volume, route, added substances, storage and batch details unambiguously, and complete documentation makes every batch traceable. All of this rests on current Good Manufacturing Practice, above all the classified cleanroom environment, with the cleanest Grade A (ISO 5) zone reserved for the critical filling step and nested within less clean Grade B, C and D areas, and on continuous environmental monitoring of particles, viable organisms, surfaces and personnel, recognising that people are the principal source of contamination, so that the release tests are made trustworthy by the controlled conditions under which the product was made.

Self-assessment

Short-answer (2 marks each)

1. Distinguish sterility from freedom from pyrogens, and explain why both are needed.
2. Describe the two methods of the sterility test and its incubation period.
3. Name the three LAL methods and state what LAL has replaced.
4. State the USP <788> particulate limits for small- and large-volume parenterals.
5. Why are containers deliberately overfilled?

Long-answer (5–10 marks each)

6. Describe the principal quality-control tests for parenterals with a diagram, explaining the limitation of the sterility test.
7. Explain the difference between sterility and pyrogen testing and describe the bacterial endotoxin test and its methods.
8. Describe cGMP cleanroom grades and environmental monitoring, and explain how they make the release tests trustworthy.

Multiple-choice

9. The bacterial endotoxin test uses a reagent derived from the: (a) rabbit (b) horseshoe crab (c) sheep (d) mouse
10. The cleanest cleanroom grade, used for aseptic filling, corresponds broadly to: (a) ISO 9 (b) ISO 8 (c) ISO 5 (d) ISO 14
11. For a small-volume parenteral, the USP <788> light-obscuration limit at $\geq 10 \mu\text{m}$ is: (a) 25 per mL (b) 6000 per container (c) 3 per mL (d) 600 per container

Answer pointers. 1. Sterility = no live organisms; pyrogen-free = no fever-causing endotoxin, which is heat-stable and survives killing, so both must be tested. 2. Membrane filtration (preferred) and direct inoculation; ~14 days in two media. 3. Gel-clot, turbidimetric, chromogenic; replaced the rabbit pyrogen test. 4. SVP ≤ 6000 ($\geq 10 \mu\text{m}$) and ≤ 600 ($\geq 25 \mu\text{m}$) per container; LVP ≤ 25 and ≤ 3 per mL. 5. To ensure the full labelled dose can be withdrawn despite wetting losses. 6. Per Figure 26.1. 7. Per Sections 26.2 and 26.4. 8. Per Section 26.7. 9. (b). 10. (c). 11. (b).

References

1. Indian Pharmacopoeia Commission. Indian Pharmacopoeia: Sterility test and Bacterial Endotoxin test. Ghaziabad: IPC; current edition.
2. United States Pharmacopeia. General chapters <71> Sterility Tests and <85> Bacterial Endotoxins Test. Rockville (MD): USP.
3. United States Pharmacopeia. General chapter <788> Particulate Matter in Injections. Rockville (MD): USP.
4. World Health Organization / national GMP guidance for sterile products; EU GMP Annex 1 (Manufacture of Sterile Medicinal Products). Current editions.
5. Aulton ME, Taylor KMG, editors. Aulton's Pharmaceutics. 6th ed. Edinburgh: Elsevier; 2022.

Glossary of Key Terms

Absorption window — the region of the gastrointestinal tract, usually the upper small intestine, where a drug is efficiently absorbed; beyond it, absorption is poor.

Aseptic processing — manufacture of a sterile product by sterilising each component separately and assembling them under contamination-excluding conditions, used when the product cannot survive terminal sterilisation.

Biodegradable — able to be broken down in the body into smaller, ideally endogenous or readily cleared, fragments.

Biocompatible — not provoking an unacceptable biological response (toxicity, inflammation, immune reaction).

Bioequivalence — demonstration that a product delivers its drug to the systemic circulation at essentially the same rate and extent as a reference product.

Brick-and-mortar model — the description of the stratum corneum as corneocyte “bricks” embedded in a lipid “mortar”.

Container–closure integrity (CCI) — the ability of a sealed package to prevent microbial ingress, thereby maintaining sterility.

Controlled release — release of drug at a predetermined, ideally constant (zero-order) rate.

Critical quality attribute (CQA) — a measurable product property that must be controlled within limits to ensure quality.

Dendrimer — a precisely branched, monodisperse polymer whose successive generations multiply its surface end-groups.

Design space — the demonstrated combination of material attributes and process parameters that assures product quality; its boundary is the edge of failure.

Edge activator — a single-chain surfactant added to a vesicle bilayer to make it ultra-deformable, as in a transferosome.

Encapsulation efficiency — the fraction of the drug used in manufacture that ends up entrapped in the carrier.

Endotoxin — a heat-stable, fever-producing component (lipopolysaccharide) of Gram-negative bacterial cell walls; the principal pyrogen.

Enhanced permeability and retention (EPR) effect — the passive accumulation of long-circulating nanoparticles in tumours through leaky vasculature and poor lymphatic drainage.

F₀ value — the lethality of a moist-heat sterilisation cycle expressed as equivalent minutes at 121 °C.

First-pass metabolism — pre-systemic metabolism of an orally absorbed drug in the gut wall and liver before it reaches the systemic circulation.

Gastro-retentive system — a dosage form designed to remain in the stomach to prolong release above an absorption window or for local action.

Hydrophilic–lipophilic balance (HLB) — a number expressing the relative weight of a surfactant's hydrophilic and lipophilic parts, predicting its function.

In-situ gelling system — a formulation that is liquid on administration and forms a gel at the site, triggered by temperature, pH or ions.

In-vitro–in-vivo correlation (IVIVC) — a predictive relationship between in-vitro dissolution and in-vivo absorption; Level A is point-to-point.

Isotonic — having the same effective osmotic concentration as blood plasma, so that cells neither swell nor shrink.

Lyophilisation — freeze-drying; removal of water by sublimation to yield a stable, reconstitutable solid.

Lipid nanoparticle (LNP) — a multi-lipid nanoparticle, including an ionisable lipid, used to deliver nucleic acids such as mRNA.

Mononuclear phagocyte system (MPS) — the network of macrophages, concentrated in liver and spleen, that clears opsonised particles from the blood.

Mucoadhesion — binding of a formulation to the mucus layer to prolong residence at a mucosal surface.

Niosome — a vesicle whose bilayer is built from non-ionic surfactants rather than phospholipids.

Opsonisation — adsorption of plasma proteins onto a particle, marking it for clearance by the mononuclear phagocyte system.

Parenteral — administered by a route that bypasses the gastrointestinal tract, in practice by injection or infusion.

Polydispersity index (PDI) — a measure of the width of a particle-size distribution; low values indicate uniformity.

Precision medicine — tailoring treatment to an individual's genetic and molecular profile rather than treating all patients with a diagnosis identically.

Process analytical technology (PAT) — a framework for measuring and controlling critical attributes during manufacture so that quality is built in.

Pyrogen — a fever-producing substance; bacterial endotoxin is the most important.

Quality by Design (QbD) — development approach in which quality is built into a product through understanding of product and process, beginning from the QTPP.

Quality target product profile (QTPP) — a prospective definition of the quality characteristics a product must have to deliver its intended benefit.

Stratum corneum — the thin, lipid-rich outermost layer of the skin that is the rate-limiting barrier to transdermal permeation.

Sterility — the absence of all viable microorganisms; distinct from freedom from pyrogens.

Sterility assurance level (SAL) — the probability that a unit remains non-sterile after a sterilisation process; the target is 10^{-6} .

Surface plasmon resonance (SPR) — the collective oscillation of a metal nanoparticle's surface electrons with incident light, exploited in gold nanoparticles.

Terminal sterilisation — sterilisation of the product in its final sealed container; the preferred route where the product can survive it.

Tonicity — the effect of a solution's osmotic concentration on cells across a membrane (isotonic, hypotonic, hypertonic).

Transfersome — an ultra-deformable vesicle, made flexible by an edge activator, able to squeeze across the skin.

Water for Injection (WFI) — the highest-purity pharmaceutical water, distinguished from Purified Water by a bacterial-endotoxin limit.

Zeta potential — the effective surface charge of a particle at the slipping plane; a large magnitude indicates colloidal stability.

Subject Index

Locators are chapter numbers, given the per-chapter structure of the book.

A — Absorption window, 7; Aseptic processing, 13, 23; Antioxidants, 22; Amphotericin B (AmBisome), 1, 4, 18.

B — Benzalkonium chloride, 25; Bioadhesion (gastric), 7; Biodegradable polymers, 3; Biodistribution, 16; Biorelevant media (FaSSIF/FeSSIF), 10; Biowaiver, 10, 17; Blow-Fill-Seal (BFS), 24, 25; Buccal delivery, 9; Buffers (parenteral), 22; Bulk vs surface erosion, 3.

C — Caco-2 assay, 15; Case studies, 18; Chelating agents (EDTA), 22; Chitosan, 3; Chronotherapeutics, 2, 18; Cleanrooms (Grades A–D, ISO), 23, 26; Colon targeting (pH/microbial/time), 8; Complex generics, 17; Container–closure integrity, 24; Controlled release, 2; Co-solvents, 22; Critical quality attributes, 14.

D — Dendrimers, 18, 19; Design space, 14; Dissolving microneedles, 13; Doxil (liposomal doxorubicin), 2, 4; Dynamic light scattering, 15.

E — Edge of failure, 14; Endotoxin (LAL, limit), 21, 26; Enhanced permeability and retention (EPR), 2, 16; Eudragit (L100/S100), 8; Excipients (GRAS, IID), 6; Extractables and leachables, 24; Extractable volume, 26.

F — First-pass metabolism, 1, 9, 11; Floating systems (effervescent, HBS), 7; Form-Fill-Seal, 24; F_0 value, 23.

G — Gastro-retentive systems, 7; Gelatin (Type A/B), 3; Glass types (I/II/III), 24; Gold nanoparticles, 5; GRAS, 6.

H — HLB and Span/Tween, 4; Hydrodynamically balanced system, 7; Hydrogel-forming microneedles, 13.

I — Implants (bio/non-degradable), 12; Inactive Ingredient Database (IID/IIG), 6; In-situ gels, 12, 25; Inhaled insulin (Afrezza/Exubera), 9, 18; Iontophoresis, 13; Iron oxide (ferumoxytol), 5; Isotonicity, 22, 25; IVIVC (Levels A/B/C), 10.

L — Lecithin and phospholipids, 4; Lipid nanoparticles (mRNA), 4, 18; Liposomes, 4, 19; Lyophilisation, 23.

M — Mesoporous silica, 5; Metal–organic frameworks, 5; Microneedles (five types), 13; Microspheres (PLGA depot), 3, 12; Mucoadhesion, 3, 9; Mononuclear phagocyte system, 16.

N — Nasal delivery (nose-to-brain), 9; Niosomes, 19; Non-clinical evaluation, 16; Novel excipient pathway, 6.

O — Ocular inserts (Ocuser, Ozurdex, Vitrasert), 12, 18; Ophthalmics, 25; Osmolality, 22.

P — PAMPA, 15; Parenteral routes (ID/SC/IM/IV), 20; Particulate matter (USP <788>), 26; PCL, 3; PEGylation (stealth), 2, 16; Permeation enhancers (chemical), 11; PLA/PGA/PLGA, 3; Polydispersity index, 15; Precision medicine, 19; Preservatives, 22, 25; Process analytical technology, 13, 14; Pulmonary delivery (DPI, respirable size), 9; Pyrogens, 20, 26.

Q — Quality by Design, 14; QTPP/CQA/CPP, 14; Quick-recall (use of), How to Use This Book.

R — Regulatory pathways (505(b)(1)/(b)(2)/(j)), 17; Release kinetics (zero/first-order, Higuchi), 2; Reservoir vs matrix patch, 11; Risperdal Consta, 3, 18.

S — SLN and NLC, 4; Sonophoresis, 13; Sterility test, 26; Sterilisation (terminal/aseptic, methods), 23; Stratum corneum, 11; Surface plasmon resonance, 5; Surfactants, 4, 22; SVP vs LVP, 20.

T — Targeting (passive/active), 2; Tonicity (haemolysis/crenation), 22; Transdermal delivery, 11; Transferosomes, 19; 500-Dalton rule, 11.

V — Vehicles (aqueous/oily), 22; Vesicular carriers, 4, 19.

W — Water for Injection (specs, production, distribution), 21.

Z — Zero-order release, 2; Zeta potential, 15.