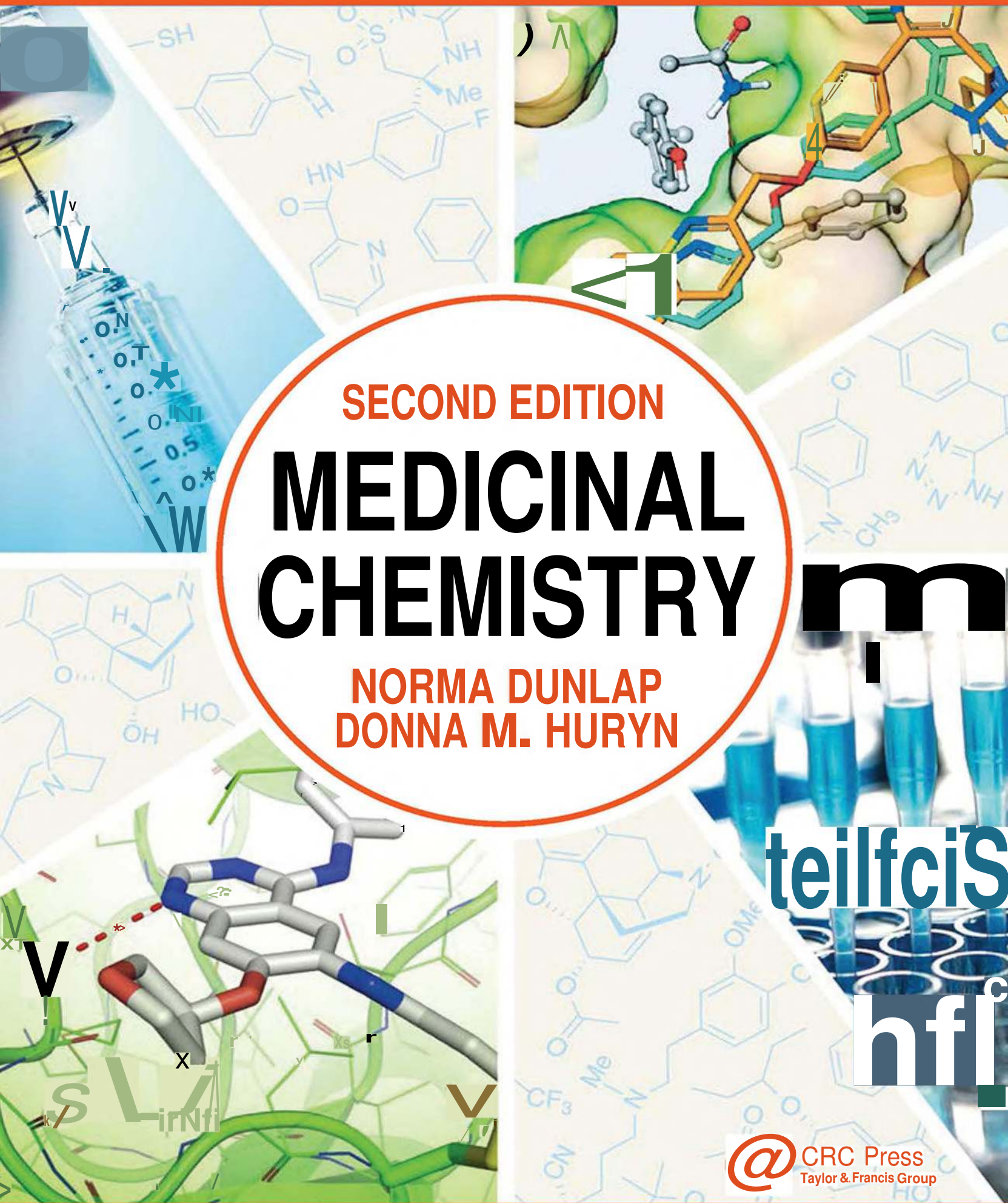


SECOND EDITION

MEDICINAL CHEMISTRY

NORMA DUNLAP
DONNA M. HURYN

CRC Press
Taylor & Francis Group

MEDICINAL CHEMISTRY

This book provides a concise introduction to the medicinal chemistry of therapeutically active compounds. Written in an accessible style, it offers a brief overview of the drug development process, from discovery to approved drug. Emphasizing medicinal chemistry strategies, the content is organized by target as well as therapeutic area drug classes. It covers existing drugs and the discovery process that led to their development. Aimed at students and professionals interested in drug discovery, *Medicinal Chemistry* requires little prior knowledge, making it ideal for classroom and training seminar use.

- Written by experts in the field, this outstanding new edition explains medicinal chemistry on chemical mechanism terms, allowing for a more interesting and user-friendly text.
- Includes valuable insights toward the various pathways taken by the pharmaceutical industry in drug discoveries.
- Enhanced text includes questions raised and suggestions made from students in the authors' medicinal chemistry class.
- Content provided to enhance the text includes a solutions manual, PowerPoint slides for lecture outlines, and especially for the earlier chapters, a selection of videos. An accompanying website will be updated to include additional material: (1) Journal Club exercises—current journal articles with guiding questions for student analysis (two per chapter); (2) additional case studies; (3) organic synthesis analysis—selected syntheses, both discovery and process, will be discussed; and (4) updates on current advances in the field of medicinal chemistry.
- The book benefits a diverse audience, including upper-level undergraduates and graduates who are studying in the fields of medicinal chemistry and drug discovery, as well as scientists working in the pharmaceutical industry.

MEDICINAL CHEMISTRY

SECOND EDITION

NORMA DUNLAP
AND DONNA M. HURYN


CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

Second edition published 2026
by CRC Press
2385 NW Executive Center Drive, Suite 320, Boca Raton FL 33431

and by CRC Press
4 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

CRC Press is an imprint of Taylor & Francis Group, LLC

© 2026 Taylor & Francis Group, LLC

Reasonable efforts have been made to publish reliable data and information, but the authors and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, access www.copyright.com or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. For works that are not available on CCC please contact mpkbookspermissions@tandf.co.uk

Trademark notice: Product or corporate names may be trademarks or registered trademarks and are used only for identification and explanation without intent to infringe.

Library of Congress Cataloging-in-Publication Data
[Insert LoC Data here when available]

ISBN: 978-1-032-73716-4 (hbk)
ISBN: 978-1-032-73715-7 (pbk)
ISBN: 978-1-003-46560-7 (ebk)

DOI: 10.1201/9781003465607

Typeset in Utopia Std
by Apex CoVantage, LLC

CONTENTS

Preface xi

Acknowledgments xiii

About the Authors xv

PART I DRUG DISCOVERY AND DEVELOPMENT 1

Chapter 1 Historical Perspective and Overview of Drug Discovery 3

Learning Objectives 3

1.1 Introduction 3

1.2 Historical Perspective of the Practice of Medicinal Chemistry 3

1.2.1 Many Medicines Were Originally Isolated from Plants 4

1.2.2 Semisynthetic Drugs Are Prepared from Natural Products 6

1.3 Overview of the Practice of Medicinal Chemistry in the Modern Era 8

1.3.1 Identification of Complex Structures Began in the Early 1900s 8

1.3.2 Advances in Total Synthesis Allow Access to Natural Products and Analogs 10

1.3.3 Medicinal Chemistry Currently Is Based on the Molecular Causes of Disease 13

1.4 Summary 15

Case Study 1 Discovery of Artemisinin 16

Case Study 2 Discovery of Taxol 16

1.5 Review Questions 18

1.6 Application Questions 18

Further Reading 19

Chapter 2 Drug Discovery: Hit and Lead Identification 21

Learning Objectives 21

2.1 Introduction 21

2.2 Overview of Drug Discovery Process 21

2.3 Identification of Hits and Leads 23

2.3.1 Various Types of Assays Are Used during the Hit and Lead Identification Process 23

2.4 Sources of Hits and Leads in Drug Discovery: Compounds with Known Bioactivity 24

2.4.1 Plant-Based Sources Are the Starting Points for Many Drugs 24

2.4.2 Microbes and Higher Animals May Be Sources for Hits, Leads, and Drugs 26

2.4.3 Hits and Leads Are Often Based on Natural Ligands 27

2.4.4 Compounds with Reported Bioactivity Are Important Sources of Hits and Leads 28

2.4.5 Clinical Observations May Uncover Unexpected Activities of Known Bioactive Compounds 29

2.5 Sources of Hits and Leads in Drug Discovery: Synthetic Libraries 30

2.5.1 Legacy Libraries Contain Compounds from Prior Discovery Campaigns 30

2.5.2 Diversity Libraries Are Large Sets of Compounds That Sample Broad Chemical Space 31

2.5.3 Combinatorial Libraries Are Sets of Closely Related Analogs That Are Prepared by Parallel or Split-Pool Synthesis 32

2.6 Methods for Screening Compound Libraries 35

2.6.1 High-Throughput Screening Assays Large Numbers of Compounds by a Variety of Different Techniques 35

2.6.2 Fragment-Based Screening Detects Direct Binding to a Protein Target 37

2.6.3 Virtual Libraries Can Be Screened by Computational Algorithms, Including Artificial Intelligence (AI) 40

2.7 Hit Triage Provides High-Quality Hits from Screening 42

2.7.1 False Positives and Nuisance Compounds Need to Be Eliminated from Primary Hit Lists 43

2.7.2 Hits Identified in Screening Campaigns Need to Be Prioritized 44

2.8 Summary 45

Case Study 1 Example of Fragment-Based Screening 46

Case Study 2 Hit Identification Using a Focused Library 47

2.9 Review Questions 49

2.10 Application Questions 49

Further Reading 51

Chapter 3
Lead Optimization: Drug–Target Interactions and the Pharmacophore 53

Learning Objectives 53
3.1 Introduction 53
3.2 Optimization of Drug–Target Interactions 55
3.2.1 Optimization of Binding Depends on Specific Molecular Interactions 55
3.2.2 Specific Interactions Contribute to the Overall Binding of Drugs to Their Targets 56
3.3 Tools Used to Develop Structure–Activity Relationships 59
3.3.1 Historical Methods for Quantifying Structure–Activity Relationships Were Based on Physicochemical Parameters and Presence or Absence of Defined Groups 60
3.3.2 Modern Computational Methods for QSAR 63
3.3.3 Ligand- and Structure-Based Drug Design Rely on Knowledge of the Pharmacophore 64
3.3.4 Pharmacophore Models Are Widely Used in Ligand-Based Drug Design 65
3.3.5 Structure-Based Drug Design Requires Knowledge of the Structure of the Drug Target 69
3.4 Summary 71
Case Study 1 Example of the Use of a Pharmacophore Model In Virtual Screening 72
Case Study 2 Example of the Use of Structure-Based Drug Design 74
3.5 Review Questions 77
3.6 Application Questions 77
Further Reading 80

Chapter 4
Lead Optimization: Properties Optimized and Medicinal Chemistry Strategies 81

Learning Objectives 81
4.1 Introduction 81
4.2 Properties Evaluated during Lead Optimization 81
4.2.1 Biological Activity Is Optimized as a Component of the Lead Optimization Process 82
4.2.2 Physical and Pharmaceutical Properties Are Optimized in Parallel with Biological Properties 83
4.3 Optimization of Lead Properties 86
4.3.1 Rules and Metrics Have Been Developed to Aid the Lead Optimization Process 86
4.4 Specific Strategies to Address Deficiencies in Lead Structures 88
4.4.1 Lipophilicity Is Optimized so That a Drug Can Reach Its Target 88
4.4.2 Leads Need to Be Optimized for Solubility in Aqueous Media 89
4.4.3 Optimization of Metabolic Stability Is a Necessary Component of Lead Optimization 89
4.4.4 Targeting Strategies and the Use of Prodrugs to Improve Drug Properties 90
4.5 Medicinal Chemistry Strategies Applied during Lead Optimization 93
4.5.1 Increasing Steric Bulk Affects Binding, Lipophilicity, and Metabolic Stability 93
4.5.2 Bioisosteres Are Used as Substitutions for Specific Functional Groups 94
4.5.2.1 Nonclassical Bioisosteres: Cyclic/Acyclic 97
4.5.2.2 Nonclassical Bioisosteres: Functional Group Replacements 97
4.5.3 Scaffold Hopping Is a Subset of Bioisosterism 100
4.5.4 Transition State Mimetics Are Common Features of Enzyme Inhibitors 101
4.5.5 Conformational Constraints Can Be Used to Improve Binding to the Drug Target 103
4.5.6 Privileged Structures Are Structural Templates Found in Multiple Drugs 105
4.6 Summary 106
Case Study 1 Lead Optimization of Losartan 107
Case Study 2 Remdesivir, First Drug Approved to Treat COVID-19 109
4.7 Review Questions 112
4.8 Application Questions 113
Further Reading 116

Chapter 5
The Process of Developing a Drug from an Optimized Lead 117

Learning Objectives 117
5.1 Introduction 117
5.2 Patenting and Drug Discovery and Development 119
5.2.1 Three Versions of Patents Are Commonly Filed during the Drug Discovery Process 119
5.3 Process Chemistry Research 120
5.4 Characterization of Pharmacokinetic Properties of Leads 122
5.4.1 Absorption Is One Determinant of How Much Drug Reaches Its Biological Target 123
5.4.2 Distribution Determines in Which Tissues Drugs Accumulate 126
5.4.3 Metabolism and Elimination Are the Body's Method for Removing Drugs 127
5.5 Drug Toxicity Determination 133
5.5.1 Indications of Toxicity Are Identified Early in the Drug Discovery Stage by *In Vitro* Assays 133
5.5.2 Toxicity Testing Is Done with *In Vivo* Assays in Multiple Animal Species 135
5.5.3 Formulation of the Final Drug Product 137
5.6 Clinical Trials in Humans 137
5.7 Summary 138
Case Study 1 Discovery and Development of Ubrogepant (Ubrelvy) 139
Case Study 2 Development of Rivaroxaban (Xarelto) 142
5.8 Review Questions 146
5.9 Application Questions 146
Further Reading 148

PART II CLASSES OF DRUG TARGETS 149

Chapter 6 Receptors, Ion Channels, and Transporters as Drug Targets 151

Learning Objectives 151

6.1 Introduction 151

6.2 G Protein–Coupled Receptors as Drug Targets 152

6.2.1 GPCRs Are Defined by Structural Features 153

6.2.2 Ligands Bind to GPCRs and May Activate or Inactivate a Receptor 158

6.2.3 Drug Activity Is Classified by Examining Biological Response as Compared to Dose 160

6.2.4 Theories Have Been Developed to Explain How Drugs Binding at the Same Receptor Site May Have Different Types of Activity 164

6.2.5 Drugs Acting at GPCRs Include Drugs to Treat Allergy, Cardiovascular Disease, Asthma, and Ulcers 165

6.3 Ion Channels (Ionotropic Receptors) as Drug Targets 168

6.3.1 Ligand-Gated Ion Channels Open or Close in Response to Binding of a Drug or Natural Ligand 170

6.3.2 Neuromuscular Blockers, Sedatives, and Anti-Nausea Agents Are Examples of Drugs Acting at LGICs 171

6.3.3 Voltage-Gated Ion Channels Open or Close in Response to Changes in Membrane Potential 173

6.3.4 Blockers of Sodium, Calcium, and Potassium VGICs Include Local Anesthetics, Analgesics, Anticonvulsants, and Anti-Arrhythmia Drugs 175

6.4 Nuclear Receptors as Drug Targets 177

6.4.1 Drugs That Target Nuclear Receptors Are Used to Treat Cancer, Inflammation, and Diabetes and to Regulate the Physiological Effects of Sex Hormones 180

6.5 Transporters as Drug Targets 183

6.5.1 Drugs Acting at Transporters Treat a Wide Array of Neuropsychiatric Disorders, Hypertension, and Diabetes 184

6.6 Summary 187

Case Study 1 Discovery of Ramelteon 188

Case Study 2 Discovery of Rosiglitazone 189

6.7 Review Questions 191

6.8 Application Questions 191

Further Reading 193

Chapter 7 Enzymes as Drug Targets 195

Learning Objectives 195

7.1 Introduction 195

7.2 Effect of Small-Molecule Modulators on Enzymes 198

7.2.1 Modulators Are Classified by How They Interact with an Enzyme 198

7.2.2 Effects on Enzyme Kinetics Are Used to Characterize the Mechanism of Small-Molecule Modulators 200

7.2.3 Enzymes Are Frequent Targets of Drug Action 202

7.3 Kinases and Kinase Inhibitor Drugs 203

7.3.1 Kinase Inhibitors Work through Multiple Binding Modes and Are Effective Anti-Cancer Drugs and Anti-Inflammatory Agents 206

7.4 Proteases and Protease Inhibitors 212

7.4.1 Protease Inhibitors Are Designed on the Basis of the Enzyme's Mechanism of Action and Are Useful for Treating HIV and Bacterial Infections, as Well as Cardiovascular Disease 214

7.5 Polymerases and Polymerase Inhibitors 216

7.5.1 Inhibitors of Polymerases Can Mimic the Substrate or Bind at an Allosteric Site 217

7.6 Esterases, Phosphodiesterases, and Their Inhibitors 219

7.6.1 Inhibitors of Phosphodiesterases Treat Cardiovascular Disease, Respiratory Disease, and Erectile Dysfunction 220

7.7 Oxidoreductases, Cyclooxygenases, and Inhibitors 224

7.7.1 Inhibitors of HMG-CoA Reductase Are Used to Lower Cholesterol Levels 224

7.7.2 Inhibitors of Cyclooxygenase Are Effective Anti-Inflammatory Drugs 227

7.8 Summary 229

Case Study 1 Discovery of the Protein Tyrosine Kinase Inhibitor Imatinib 230

Case Study 2 Discovery of Grazoprevir, an HCV NS3/4A Protease Inhibitor 231

7.9 Review Questions 236

7.10 Application Questions 236

Further Reading 238

Chapter 8 Protein–Protein and Lipid Structure Interactions as Drug Targets 239

Learning Objectives 239

8.1 Introduction 239

8.2 Protein–Protein Interactions as Drug Targets 240

8.3 Modulating Protein–Protein Interactions 241

8.3.1 One Strategy to Prevent Binding between Two Proteins Is to Mimic One of the Protein Partners 242

8.3.2 Protein–Protein Interactions Can Be Modulated by Chemically Induced Proximity 248

8.3.2.1 Bifunctional Degraders 249

8.3.2.2 Molecular Glues 252

8.3.3 Small Molecules That Bind to a Protein Can Stabilize or Destabilize It, Leading to Therapeutic Effects 253

8.4 Lipids of Cell Membranes as Drug Targets 258

8.4.1 Drugs Acting at Lipids in Membranes Act as Anti-Infective Agents 258

8.5 Summary 260

Case Study 1 Discovery of Vepdegestrant 261

Case Study 2 Discovery of Maraviroc 264

8.6 Review Questions 267
8.7 Application Questions 267
Further Reading 269

Chapter 9
DNA and RNA as Drug Targets 271
Learning Objectives 271
9.1 Introduction 271
9.2 DNA as a Drug Target 273
9.2.1 Intercalation Is One of the Main Mechanisms by Which Small Molecules Interact with DNA 274
9.2.2 Drugs May Bind in the Minor Groove of DNA through Reversible Noncovalent Interactions 276
9.2.3 Drugs That Interact with DNA by Irreversible Mechanisms Are Anti-Cancer Agents 278
9.2.4 Compounds That Cleave DNA after Binding and Generating Free Radicals Are Powerful Drugs Used to Treat Cancer 281
9.3 RNA as a Drug Target 284
9.3.1 Compounds Targeting Bacterial Ribosomal RNA Inhibit Protein Synthesis and Are Useful Anti-Infective Agents 284
9.3.2 Antisense Therapies Targeting RNA Are Effective in Treating Genetic Diseases and Infections 285
9.3.3 Small Molecules Can Bind to RNA and Modify Splicing 288
9.4 Summary 289
Case Study 1 Doxorubicin and Analogs 290
Case Study 2 Discovery of Mipomersen 292
9.5 Review Questions 294
9.6 Application Questions 294
Further Reading 296

PART III SELECTED THERAPEUTIC AREAS 297

Chapter 10
Anti-Cancer Drugs 299
Learning Objectives 299
10.1 Introduction 299
10.2 Drugs Targeting DNA Replication and Mitosis 301
10.2.1 Drugs May Target DNA and DNA Processing Enzymes, Such as Polymerase and Topoisomerase 301
10.2.2 Drugs May Target Biosynthesis of DNA Building Blocks 306
10.2.3 Drugs May Target Structural Proteins Involved in Cell Division 311
10.2.4 Effects of Anti-Tumor Compounds May Be Optimized by Targeting Delivery Systems 311
10.3 Drugs Targeting Oncogenes and Signaling Pathways 313
10.3.1 Kinase Inhibitors Target Mutant or Overexpressed Kinases Found in Cancer Cells 313
10.3.2 RAS Inhibitors Target a Previously "Undruggable" Protein 319
10.4 Inhibition of Angiogenesis 320
10.5 Drugs Targeting the Ubiquitin–Proteasome Pathway 321
10.6 Drugs Targeting Epigenetic Processes 324
10.6.1 Nucleosides with Modified Bases Can Act as DNA Methyltransferase Inhibitors 325
10.6.2 Histone Deacetylase Inhibitors Bind to Zinc in the Enzyme's Active Site 326
10.7 Drugs Targeting Hormone-Dependent Tumors 327
10.8 Summary 329
Case Study 1 Discovery of Vorinostat 330
Case Study 2 Selective PARP1 Inhibitor Saruparib 331
10.9 Review Questions 335
10.10 Application Questions 335
Further Reading 336

Chapter 11
Antiviral and Antifungal Agents 337
Learning Objectives 337
11.1 Introduction 337
11.2 Antiviral Agents: Common Viral Structures, Replication Process, and Impact 338
11.3 Antiviral Drug Discovery 340
11.4 Antiviral Drugs Targeting Viral Life Cycle 342
11.4.1 Drugs May Target Viral Attachment or Entry 342
11.4.2 Drugs May Target Uncoating of Viral Particles 343
11.4.3 Many Antiviral Drugs Inhibit Transcription by Targeting the Active Site of Polymerases 345
11.4.4 Prodrugs Improve the Properties and Effectiveness of Nucleosides 348
11.4.5 Some HIV Drugs That Target Reverse Transcriptase Bind at Allosteric Sites 350
11.4.6 Viral Protease Inhibitors Have Been Developed by Using Structure-Based Drug Design 352
11.4.7 Inhibitors of the Hepatitis C Virus NS5A Replication Complex Exhibit Antiviral Activity 357
11.4.8 Drugs to Treat Retroviruses May Target the Viral Integrase 357
11.4.9 Influenza Drugs' Targets Include Viral Budding and an Endonuclease 359
11.5 Antifungal Agents: Common Fungal Structures, Replication Process, and Impact 362
11.6 Antifungal Drug Discovery 364

11.7 Drugs Targeting Fungi 364
11.7.1 Polyene Drugs Target the Fungal Cell Membrane 364
11.7.2 Azoles and Allylamines Inhibit Ergosterol Biosynthesis 365
11.7.3 Some Antifungal Drugs Inhibit Cell Replication 367
11.7.4 Echinocandins Inhibit Fungal Cell Wall Synthesis 367
11.8 Summary 370
Case Study 1 Discovery of Daclatasvir 371
Case Study 2 Discovery of Rilpivirine 372
11.9 Review Questions 375
11.10 Application Questions 375
Further Reading 377

Chapter 12
Antibacterial and Antiparasitic Drugs 379
Learning Objectives 379
12.1 Introduction 379
12.2 Antibacterial Agents: Bacterial Structure, Points of Drug Interaction, and Impact 380
12.3 Antibacterial Drug Discovery 381
12.4 Bacterial Cell Wall Synthesis Inhibitors 384
12.4.1 β-Lactam Antibacterial Compounds Are the Largest Class of Cell Wall Synthesis Inhibitors 385
12.4.2 Glycopeptide Antibacterial Agents Inhibit Cell Wall Synthesis by a Different Mechanism Than β-Lactams 392
12.5 Bacterial Protein Synthesis Inhibitors 394
12.5.1 Aminoglycosides, Tetracyclines, and Erythromycins Are Protein Synthesis Inhibitors Based on Natural Products 394
12.5.2 Oxazolidinones Are Inhibitors of Protein Synthesis 399
12.6 Drugs Targeting Bacterial DNA Replication 400
12.6.1 Quinolone Antibacterial Agents Inhibit DNA Gyrase 400
12.6.2 Sulfonamide Antibacterial Compounds Inhibit Folic Acid Biosynthesis 402
12.7 Drugs Affecting Bacterial Cell Membrane Permeability 404
12.8 Antiparasitic Drugs: Common Parasite Features, Infectious Process, and Impact 405
12.9 Antiparasitic Drug Discovery 405
12.9.1 Current Drugs Used to Treat Malaria Are Derivatives of Natural Products 406
12.9.2 There Are Few Effective Treatments for Trypanosomal Diseases 409
12.9.3 Drugs That Treat Helminthic Diseases Are Used in Both Veterinary and Human Medicine 411
12.10 Summary 412
Case Study 1 Discovery of Linezolid 413
Case Study 2 SQ109 for Multi-Drug-Resistant Tuberculosis 415
12.11 Review Questions 417
12.12 Application Questions 417
Further Reading 418

Chapter 13
Drugs Acting on the Central Nervous System 419
Learning Objectives 419
13.1 Introduction 419
13.2 Drugs Targeting Anxiety and Depression 422
13.2.1 Positive Allosteric Modulators at the $GABA_A$ Receptor Are Sedatives and Anticonvulsants 423
13.2.2 Drugs That Increase Monoamine Neurotransmitter Levels or Act as Partial Agonists at Neurotransmitter Receptors Are Antidepressants 426
13.3 Drugs Targeting Psychosis 432
13.3.1 Typical Antipsychotics Act as Dopamine Receptor Antagonists 433
13.3.2 The Mechanism of Atypical Antipsychotics Involves Dopamine Receptor Antagonism Combined with Serotonin Receptor Antagonism 435
13.3.3 Drugs Targeting Glutamate, Phosphodiesterase 10A, and the Muscarinic Acetylcholine Receptor Also Act as Antipsychotics 436
13.4 Drugs Targeting Pain 439
13.4.1 Drugs Used to Treat Pain Include Opiates and Nonsteroidal Anti-Inflammatory Drugs as Well as Antidepressants 439
13.4.2 Novel Approaches to Pain Treatment Include Selective Ion Channel Blockers 440
13.4.3 Selective Serotonin Agonists Are Used to Treat Migraine 441
13.5 Drugs Targeting Neurodegeneration 443
13.5.1 Approved Small-Molecule Drugs to Treat Alzheimer's Disease Affect Acetylcholine or Glutamate Levels 444
13.5.2 Recent Approaches to Modifying the Course of Alzheimer's Disease with Small Molecules Have Focused on Decreasing Production of Amyloid Plaques 446
13.5.3 Drugs Developed to Treat Parkinson's Disease Are Based on Increasing Dopamine Levels 447
13.6 Drugs of Abuse 450
13.6.1 Stimulants Include Both Legal and Illegal Drugs 450
13.6.2 Hallucinogens Include Serotonin Receptor Agonists and *N*-methyl-D-Aspartate Antagonists 452
13.6.3 Cannabinoids Have Potential Medical Uses 453
13.7 Summary 454
Case Study 1 Development of Vilazodone 455
Case Study 2 Discovery of Donepezil 456
13.8 Review Questions 459
13.9 Application Questions 459
Further Reading 461

Index 463

PREFACE

This textbook is an overview of the subject of medicinal chemistry within the context of drug design and discovery. Both of us were trained as synthetic organic chemists and had our "first" careers in medicinal chemistry departments at major pharmaceutical companies. Upon moving into academia, we both took on teaching a one-semester class in medicinal chemistry. Anyone who teaches this type of course quickly realizes the difficulty in bringing together the many aspects that comprise this subject, such as organic chemistry, pharmacology, and biochemistry, as well as introducing therapeutic areas in medicinal chemistry in just one semester. After teaching this course for many years using various review articles as text, as well as material written ourselves, we decided to write a textbook. We intended to write a text that includes the historical background of drug discovery but emphasizes modern practices in design, discovery, and development. We focus on drug targets and how current knowledge of these targets drives medicinal chemistry and drug discovery. While it is not possible to cover all therapeutic areas in a one-semester text, selected areas and classes of drugs are also included. Overall, we intend for this text to meet the need for a modern, one-semester textbook that ties together the many different aspects of medicinal chemistry as it is currently practiced.

The target audience is primarily upper-level undergraduate students and beginning graduate students in organic chemistry and medicinal chemistry. Other students who might find the text useful include pharmacy students, biochemistry students, and anyone planning to pursue a career in the drug discovery field. In addition to science majors, we have had students from fields as diverse as finance, engineering, and biomedical ethics take our courses, all with a common interest in learning about drug discovery. The book should also be useful for professional scientists, such as chemists, biologists, and patent professionals, who are just entering the drug discovery field and may not have had direct experience with medicinal chemistry. It is expected that those using this textbook have a basic knowledge of organic chemistry as well as basic biology.

To present this broad subject in a way that is logical and covers essential topics in a one-semester format, our approach is to divide the text into three sections. In Part I, "Drug Discovery and Development," the first five chapters cover historical background in drug discovery, as well as the modern drug discovery and development process, with an emphasis on medicinal chemistry strategies used in various phases. These chapters give the essential background of hit and lead discovery, modifications driven by knowledge of both target and pharmacophore structure, common medicinal chemistry strategies, and the central importance of pharmacokinetics. They provide the background that subsequent chapters build upon. In Part II, "Classes of Drug Targets," Chapters 6–9 focus on the major drug targets, namely, their structure and function, and how medicinal chemists approach each. Examples of specific drugs developed for each type of target are included. Targets include receptors and ion channels, enzymes, protein–protein and lipid interactions, and nucleic acids. Each of those is broken down into more specific classes of important drug targets. These chapters can be covered in any order, but reference topics in Chapters 1–5. In Part III, "Selected Therapeutic Areas," the final chapters, Chapters 10–13, are devoted to selected important therapeutic areas. While there are many more chapters that could have been written on additional areas of interest, these topics were selected as examples that tie together components introduced in the earlier chapters and represent both historically important agents and future opportunities. The topics included are anti-cancer, antibacterial, and antiviral drugs and drugs acting on the central nervous system. A subset of Chapter 13 is devoted to drugs of abuse, informed by student

interest on this topic. Any of these last chapters on specific therapeutic areas can be considered optional or used independently, but they also rely on the material covered in Chapters 1–9.

In writing this book, we have tried to emphasize modern practices while also retaining some historical background as a way to show progress in the field and to help explain complex concepts. Full color figures are used throughout to identify key structural features and highlight specific points of significance. Key terms are highlighted in boldface type and are further explained in an expanded glossary. Each chapter ends with problems, divided into review questions and application questions. Review questions are based on understanding material presented in the chapter, whereas application problems require critical thinking skills to apply previously covered concepts to new problems. Some application problems ask the student to research a topic, intending to provide relevant experience to students. Answers to review questions are printed in the back of the book. Each chapter also includes two case studies on individual drugs that highlight relevant subjects in those chapters. While some are historical examples, most case studies focus on compounds and recently developed drugs.

In this second edition, each chapter has been updated to include advances in the field. Changes include the introduction of new topics, such as AI/machine learning techniques in drug discovery, biased GPCR drug design, an explanation of sustainable practices in drug discovery, and the new modalities (e.g., PROTACS and molecular glues) that are advancing our ability to develop novel therapeutics. In addition, innovative new drugs/drug candidates, including those targeting previously undruggable targets, such as RAS, and new diseases, such as COVID-19, have been added to specific therapeutic area chapters. These changes reflect the exciting rapid advancements in medicinal chemistry.

Norma Dunlap

Donna M. Huryn

ACKNOWLEDGMENTS

We gratefully acknowledge the contributions of the following scientists and instructors, for their advice and critique in the development of this book:

Galia Blum (The Hebrew University), Justin M. Chalker (Flinders University), Amber F. Charlebois (SUNY Geneseo), Emily Dykhuizen (Purdue University), Steven M. Firestine (Wayne State University), Lekh Nath Sharma Gautam (St. Jude Children's Research Hospital), George Greco (Goucher College), Wayne C. Guida (University of Southern Florida), Geneive Henry (Susquehanna University), David Hunt (The College of New Jersey), William Maio (New Mexico State University), Paul L. Ornstein (Roosevelt University), Timothy S. Snowden (The University of Alabama), Erika A. Taylor (Wesleyan University), Nanette Wachter (Hofstra University), and Simon Ward (University of Sussex).

Special thanks go to Jay Kostman, for reviewing the text, particularly for providing expert feedback on antiviral and antibacterial chapters; to Stephen Kostman, Matthew Kostman for text preparation, and Miles Rizzo for preparation of several figures; and to Chelsea Harmon and Melissa Grenier, for the review of the text and preparation of answer keys. We are grateful for the input of many students over the past years. Finally, our thanks also go to the people at Taylor & Francis, particularly Hilary LaFoe and Varalika Kathuria, for their enthusiasm, support, and guidance on this project.

Norma Dunlap
Donna M. Huryn
May 2025

RESOURCES FOR INSTRUCTORS AND STUDENTS

The teaching and learning resources for instructors and students are available online. The instructor resources on the Taylor & Francis website are password-protected and available only to adopting instructors. The student resources on the Taylor & Francis website are available to everyone. We hope these resources will enhance student learning and make it easier for instructors to prepare dynamic lectures and activities for the classroom.

Instructor Resources

Instructor resources are available on the Routledge website, located at www.routledge.com/9781032737164. The website provides access not only to the teaching resources for this book but also to all other Taylor & Francis textbooks. Adopting instructors can obtain access to the site from their sales representative.

Art of Medicinal Chemistry

The images from the book are available in two convenient formats: PowerPoint® and JPEG. They have been optimized for display on a computer. Figures are searchable by figure number, by figure name, or by keywords used in the figure legend from the book.

Answer Keys to Application Questions

Answers to application questions are available to qualified instructors.

Student Resources

The resources for students are available on the Medicinal Chemistry Student website, located at www.routledge.com/9781032737164.

Journal Club

The Journal Club recommends journal articles that complement topics in the textbook to improve students' critical analysis of research and to promote a better understanding of the research process. These are examples taken from the recent literature and include early-stage experimental compounds as well as compounds that are at the clinical trial stage. Each Journal Club document provides background information on the chosen paper as well as questions and discussion points to stimulate in-class discussion. We have found that primary literature is very helpful in exposing upper-level students to research articles, giving students practice in comprehension of real-life research questions, and in teaching them how all aspects of medicinal chemistry come together in drug development. Answers will be provided to instructors only.

ABOUT THE AUTHORS



Norma Dunlap is Professor Emerita of Chemistry at Middle Tennessee State University, where she specialized in synthetic organic chemistry and medicinal chemistry. Her research, done primarily with undergraduates and MS students, was focused on the design and synthesis of bioactive peptidomimetics. She holds a PhD in organic chemistry from the University of Wyoming and was an NIH postdoctoral fellow at the University of Pennsylvania. She began her career as a senior research scientist at Hoffmann–La Roche in drug discovery. After moving to academia, she was a lecturer at Vanderbilt University before moving to Middle Tennessee State University.



Donna M. Huryn is Professor of Practice in the Chemistry Department at the University of Pennsylvania, where her research focuses on medicinal chemistry projects on novel drugs to treat kidney disease, cancer, and rare genetic diseases. She holds a PhD from the University of Pennsylvania. She started her career in industry at Hoffmann–La Roche, and then Wyeth Research, before joining academia; she was previously a faculty member at the University of Pittsburgh School of Pharmacy. She is a fellow of the American Association for the Advancement of Science and the American Chemical Society and was inducted into the ACS Division of Medicinal Chemistry Hall of Fame. She is a recipient of the ACS Philadelphia Local Section Award, the Philip S. Portoghese Lectureship, and the Garvin-Olin Medal. From 2014 to 2023, she was Associate Editor of *ACS Medicinal Chemistry Letters*. In 2019, she co-founded Empowering Women in Organic Chemistry (EWOC) and continues to serve on the organizing committee.

PART I

DRUG DISCOVERY AND DEVELOPMENT

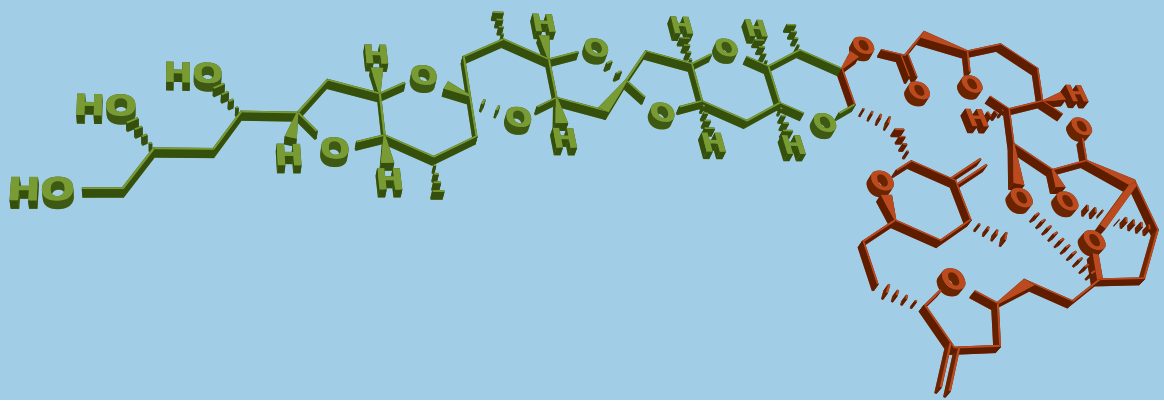
CHAPTER 1
Historical Perspective and Overview of Drug Discovery

CHAPTER 2
Drug Discovery: Hit and Lead Identification

CHAPTER 3
Lead Optimization: Drug–Target Interactions and the Pharmacophore

CHAPTER 4
Lead Optimization: Properties Optimized and Medicinal Chemistry Strategies

CHAPTER 5
The Process of Developing a Drug from an Optimized Lead



Historical Perspective and Overview of Drug Discovery

1

LEARNING OBJECTIVES

- Understand how *medicinal chemistry* is defined.
- Become familiar with the historical progression of therapeutics from plant preparations to single pure substances.
- Learn how organic chemistry contributed to the identification and synthesis of drugs.
- Describe how scientists' understanding of drug action in the body has evolved over time.

1.1 INTRODUCTION

Even before we are born, we are under assault by foreign organisms and by defects in our bodies that may ultimately lead to an overwhelming variety of diseases and conditions. We fight back by means of our immune system and various repair mechanisms, and when these prove insufficient, we rely on external substances: medicines. **Medicinal chemistry** is the branch of chemistry that focuses on the design, synthesis, and development of new medicines or drugs. Hidden in this simple definition are several areas of science, including organic chemistry, biochemistry, and pharmacology. The vast majority of drugs are small organic molecules, as opposed to inorganic molecules or macromolecules, such as proteins. Hence, in order to study these drugs, an understanding of the structure and nature of functional groups present in organic compounds is required. To understand how drugs exert their action in a living system, it is also necessary to have a basic understanding of **pharmacology**, the study of drug action. The process of how new drugs are designed involves examining drug–macromolecule interactions at a molecular level, which includes various aspects of biochemistry. Overall, the process of drug discovery and development in medicinal chemistry includes the iterative process of understanding drug–target interactions, designing improved analogs, synthesizing new molecules, and evaluating compounds.

1.2 HISTORICAL PERSPECTIVE OF THE PRACTICE OF MEDICINAL CHEMISTRY

Medicinal chemistry has been practiced, in some form or another, for thousands of years, progressing from the use of certain plants or plant parts through preparation of semi-pure plant extracts to isolation of pure active molecules and, finally, to the

DOI: 10.1201/ 9781003465607-2

invention of molecules specifically designed to treat a certain disease or condition. Until fairly recently, our source of medicines was limited to substances that occurred in nature.

There are numerous recorded instances from approximately 3,000 BC and AD 1,000 of plants being used to treat disease. These include descriptions of the poppy plant (called Gil or Joy) by ancient Sumerians (3,000 BC), in the Ebers Papyrus (Egypt, 1550 BC), and in the writings of Galen (AD 129–199). During this period, special priests, shamans, and medicine men and women administered medicine. There was an aura of mystery and religion involved in their knowledge of the magical plants that could heal, prolong life, and treat pain and other symptoms of disease. Many of these substances are still used today, either in their crude plant form or as the pure active component of the plant.

During the early part of the second millennium, scientists began to realize that a specific ingredient in the plant was responsible for its beneficial activity; therefore, there was a focus on obtaining semi-pure and standardized plant preparations to be used as medicines. In this time period appeared publications of **herbals** and **pharmacopeias**, which are written collections of descriptions of plants useful for medicinal purposes. These publications listed plants that had been used for thousands of years and described treatments using those plants for specific conditions. Some important examples include an herbal published in Switzerland in 1470 by Bartholomeus Anglicus; the *Nuovo Receptario Composito*, published in Venice in 1498; the *First London Pharmacopoeia*, published in 1618; and *The Divine Husbandman's Materia Medica Classic*, published in 1596 in China. These publications provided specific information on the use and preparation of various plants that had been used for thousands of years, a number of which are still used today. During this time period, the Swiss physician Paracelsus (1493–1541), in particular, attempted to find the healing element in both organic and inorganic pharmaceutical preparations. Some specific examples of medicinal plants and their uses are summarized in **Table 1.1**.

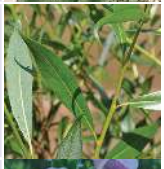
As can be seen from the examples in Table 1.1, there is a long history of the use of medicinal plants, going back to the earliest recorded history. While most were used to treat specific medical conditions, others, such as *Erythroxylum coca*, were used as stimulants or euphorics. Some others originated from the arrow poisons (known as woorara), used as a paralytic to hunt animals for centuries in South America. In the 1800s, animal studies using woorara, carried out by Sir Benjamin Brodie and others, showed that the heart kept beating while the animal became paralyzed, eventually leading to the use of the active ingredient, tubocurarine, in anesthesia.

The science of chemistry moved into a more modern understanding of the composition of materials between the 1700s and the 1900s. With this expansion of the field of chemistry, there was a focus on the isolation of pure compounds and attributing biological activity to specific substances even though the structure was often unknown. By the late 1800s, the knowledge of organic compounds had advanced to a basic understanding of functional group arrangements, along with the correct assignment of many structures. This period also marked the introduction of the first semi-synthetic drugs: those prepared by chemical derivatization of natural products.

1.2.1 Many Medicines Were Originally Isolated from Plants

Many early medicines fall into the category of nitrogen-containing plant products, called **alkaloids**. The first alkaloid obtained in pure form was morphine, isolated by Sertürner from a type of poppy plant (*Papaver somniferum*) in 1805 and named for Morpheus, the Greek god of dreams. It became widely used both to treat pain and as a cough suppressant. Many other medicinally active plant ingredients were isolated during this period. For example, Pierre Pelletier and Jean Caventou isolated colchicine from the autumn crocus in 1820, Friedrich Gaedcke isolated cocaine in 1855, and Nagayoshi Nagai isolated ephedrine from ma huang in 1887. Another example is the isolation of salicylic acid from willow bark. Salicin, originally isolated from the bark in 1826 by Johann Buchner, is an aryl glycoside. Crude isolation in this case involved the preparation of a simple tea or aqueous extract of willow bark and leaves, followed by further extraction and crystallization.

Table 1.1 Examples of Medicinal Plants, the Historical Documentation of Their Use, and Their Active Ingredient

Plant (Common Name)	Historical Mention	Use	Active Agent	
Papaver somniferum (opium poppy)	3000–300 BC (Sumeria, Egypt, Greece, China) 1500s (Paracelsus in Switzerland): laudanum (~12% morphine)	Analgesic	Morphine	
Dichroa febrifuga (Chang shan or Chinese quince)	2735 BC (China) one of the "50 fundamental herbs" in traditional Chinese medicine	Malaria, fever	Febrifugine	
Ephedra sinica (ma huang)	2735 BC (China)	Cough, asthma	Ephedrine	
Cinchona (fever tree)	1600s (South Americas); brought to Europe by Jesuit missionaries	Malaria, fever	Quinine	
Colchicum autumnale (autumn crocus)	AD 300 (Egypt) AD 560 (Arabia); source of hermodactyls (medicinal roots)	Joint pain, gout, rheumatism	Colchicine	
Erythroxylum coca ("the divine plant of the Incas")	AD 600 (South America), used by the Moche and Inca civilization possibly as early as 3000 BC 1596 (Spanish explorers mention its effects)	Stimulant, euphoric	Cocaine	
Chondrodendron tomentosum (ourari/woorari = "killer of birds")	1516 (South America), brought to England by Walter Raleigh, after visiting Guyana	paralytic in anesthesia (originally used for hunting)	Tubocurarine	
Salix alba (willow)	500 BC (Greece: Hippocrates; Sumeria, Assyria) 1700s (Native Americans) 1763 (England)	pain, fever	Salicylic acid	
Digitals purpurea (foxglove)	1250 (Wales) 1785 (England)	cardiac stimulant	Digoxigenin	

Poppy photo by Réginald Hulhoven/Wikimedia Commons, CC BY-SA 1.0. Quinine photo by Wouter Hagens/Wikimedia Commons, CC BY-SA 3.0. Ephedra photo by Lazare Gagnidze/Wikimedia Commons, CC BY-SA 4.0. Coca photo by Geoff Gallice/Wikimedia Commons, CC BY 2.0. Chondrodendron photo by Alex Popovkin/Flickr, CC BY 2.0. Salix alba photo by Ladislav Luppa/Wikimedia Commons, CC BY-SA 3.0. Digitalis purpurea drawing by Elisabete Ferreira/Wikimedia Commons, CC BY-SA 3.0.

Many active plant products contain either a basic amine or a carboxylic acid, and isolation procedures rely on liquid–liquid extractions with sequential basic and acidic aqueous solutions. The isolation of ephedrine outlined in what follows is a typical procedure for isolating an amine:

Step 1: The *Ephedra sinica* (ma huang) plant material is harvested, dried, and powdered, and the powder is partitioned between a two-phase mixture of benzene and cold aqueous sodium carbonate. In basic solution, the amine of ephedrine is not protonated, and the compound resides in the organic (benzene) layer. The basic aqueous layer contains various water-soluble compounds, debris, and any organic acids, which are deprotonated at high pH and, therefore, water-soluble.

Step 2: The two layers are separated, and the benzene solution is extracted with dilute HCl. This treatment protonates the amine, making ephedrine charged, and it is partitioned into the aqueous layer. Any unwanted non-basic organic compounds remain in the benzene layer.

Step 3: The aqueous layer from the previous step is combined with chloroform, and then potassium carbonate is added to "free base" the amine. This converts ephedrine back to its amine form, rendering it organic-soluble once again. The organic layer is collected, and the chloroform is evaporated to provide solid ephedrine with minimal impurities.

Step 4: The purest forms of solids are usually crystalline materials, and many amines crystallize easily as salts. The ephedrine in this case was treated with oxalic acid and crystallized as its oxalate salt. Nagai obtained a 2% yield (based on plant material) of the pure ephedrine salt following this procedure (**Figure 1.1**).

Although historically plants were the earliest sources for medicines, in the 1900s, bacteria and molds became important sources as well. The type of isolation procedure outlined in **Figure 1.1** works well for many compounds containing basic amines, as long as they are stable to extraction conditions with strong base and strong acid. In contrast to ephedrine, some sensitive compounds, such as penicillin, cannot be isolated in this manner. In fact, the development of penicillin was temporarily abandoned in 1932 because of the difficulty in the isolation of the pure compound.

Until the 1940s, nearly all drugs were isolated directly from plants as the natural products or were simple derivatives, such as esters, of a natural product, with most having some origin in herbal and medicinal plant lore. In the late 1900s, natural product isolation as a source of drugs became a less frequent approach; however, there was still interest in finding new drugs or leads from plant and animal sources, often through government initiatives. One example was the International Cooperative Biodiversity Groups project (ICBG) launched by the National Institutes of Health (NIH), the National Science Foundation (NSF), and the US Agency for International Development in 1993. This program supported partnerships between academic institutions, private enterprises, and local populations in countries where medicinal plants are found. The goal of this initiative was to identify new drugs to treat a variety of diseases by both random and targeted collection of plants and plant extracts, based on ethnobotanical information from local healers. A more recent example is the US National Cancer Institute (NCI) Program for Natural Product Discovery Prefractionated Extract Library, launched in 2019, whose goal is to reinvigorate natural product–based drug discovery research by providing, at no charge to those who want to use these samples for drug discovery, a very large library of partially fractionated samples of plant, marine invertebrate, and microbial samples. In addition, in 2014, the "Nagoya Protocol" was adopted by nearly 200 countries and provides a common agreement on how genetic resources (e.g., plants, microorganisms, etc.), as well as any benefit from their use, can be shared between nations. This treaty is also expected to result in a resurgence in interest in natural products as drugs.

1.2.2 Semisynthetic Drugs Are Prepared from Natural Products

A **semisynthetic compound** is one that is derived from a natural product but undergoes additional chemical modification, usually requiring only a few synthetic steps. The earliest reported semisynthetic drug is acetylsalicylic acid (aspirin), derived from

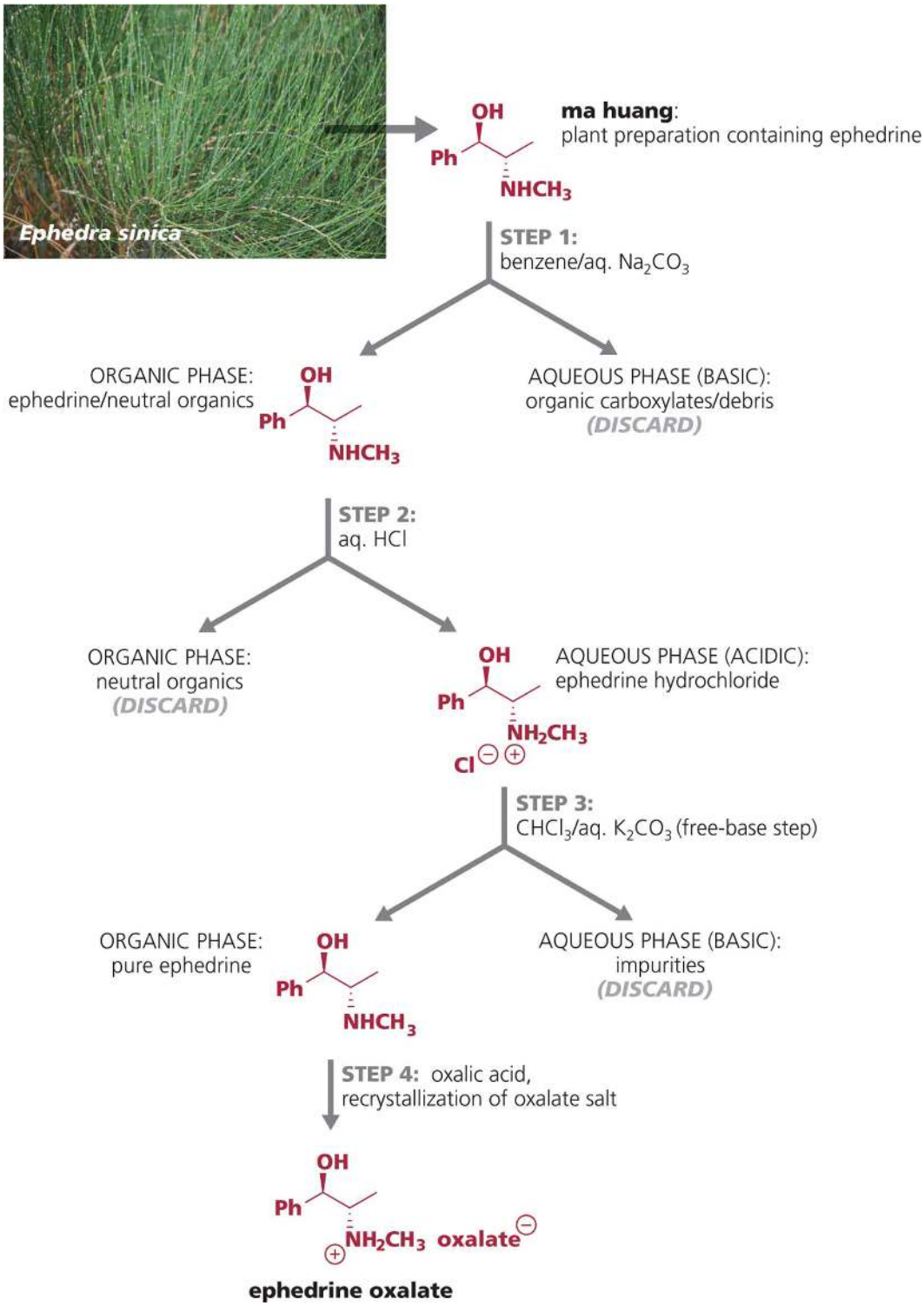


Figure 1.1 **Procedure for isolation of ephedrine from ma huang.** Step 1: The crude plant preparation is partitioned between benzene and aqueous base (ephedrine is shown in red). Step 2: The organic phase is extracted with aqueous acid. Step 3: Base is added to the aqueous phase to free-base ephedrine; the next step is extraction of the free base into organic solvent and evaporation to solid ephedrine. Step 4: Recrystallization as the oxalate salt. Photo by Lazare Gagnidze/Wikimedia Commons, CC BY-SA 4.0.

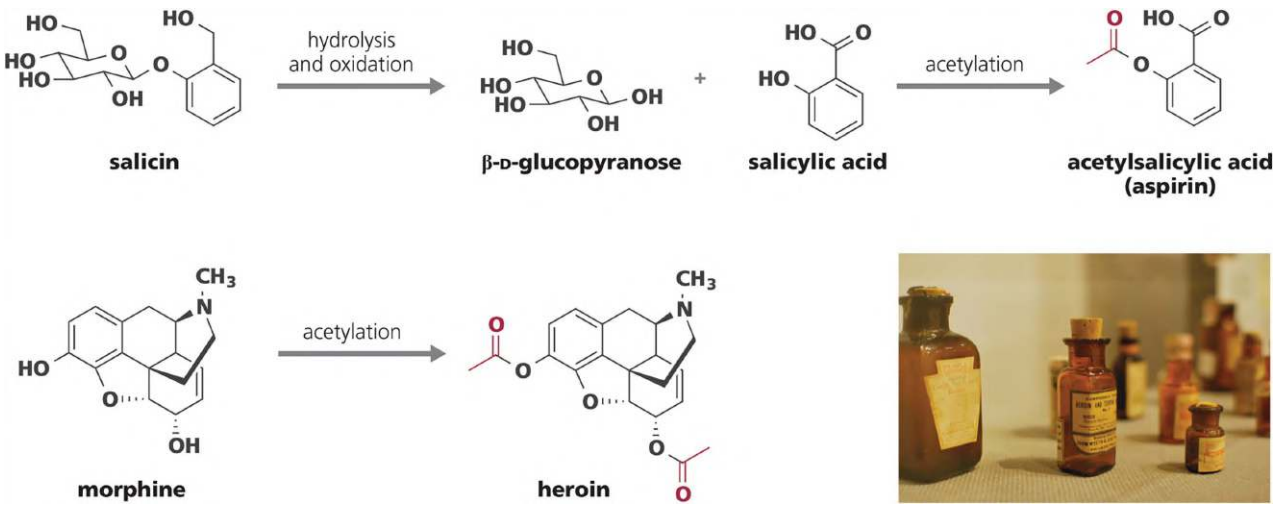


Figure 1.2 Semisyntheses of aspirin and heroin. The hydroxy group of salicylic acid (generated from hydrolysis and oxidation of salicin) and both hydroxyl groups of morphine react with acetic anhydride to form the acetate esters (shown in red). (Photo by Karen Neoh/Flickr, CC BY 2.0.)

salicylic acid by acetylation with acetic anhydride. In 1829, Henri Leroux isolated approximately 25 g of salicin from 1 kg of willow bark. Nine years later, Raffaele Piria converted salicin to β-d-glucopyranose and the oxidation product, salicylic acid. Salicylic acid is generated upon ingestion of salicin and is the active ingredient in willow bark. Salicylic acid, however, was too irritating to the stomach to be useful, and the acetate was first prepared by Charles Gerhardt in 1853 as a less-irritating form of the drug. Although Gerhardt did not pursue marketing the drug, in 1897, Felix Hoffmann, a chemist working for Bayer Pharmaceuticals in Germany, prepared some of the acetate for his father, who was suffering from arthritis. Acetylsalicylic acid was marketed by Bayer Pharmaceuticals in 1900 as aspirin (**Figure 1.2**). Aspirin is still widely used to alleviate pain and fever, as well as for the prevention of cardiovascular disease and stroke.

Another early example of a semisynthetic drug is heroin, which is the bis(acetate) of morphine. The drug was first synthesized in England in 1874 by C. R. Alder Wright by the reaction of morphine with acetic anhydride. It was later prepared by Hoffmann (the inventor of aspirin) and dubbed heroin after the German word for *heroic*. Heroin was mistakenly believed to be a less-addictive form of morphine and was marketed by Bayer as a cure for morphine addiction until 1910, when it became clear that it was rapidly converted to morphine and was, in fact, even more addictive.

1.3 OVERVIEW OF THE PRACTICE OF MEDICINAL CHEMISTRY IN THE MODERN ERA

During the last 100 years, medicinal chemistry has become a branch of science in its own right. The most significant advances were in the identification and structure elucidation of active compounds, the appreciation that drugs interact with specific proteins, such as enzymes and receptors, to elicit a beneficial effect, and in understanding the molecular interactions between a drug molecule and its target. In just the last 50 years alone, there have been tremendous advances in our knowledge of drug targets, such as enzymes and receptors, and in the development of new modalities beyond traditional small molecules. With combinations of improved spectroscopic techniques and computer graphics, as well as advances in imaging, we can now "see" a drug as it interacts with a specific enzyme or protein. This technology allows for the design of new drugs based on observable interactions, whereas in the past, the approach to developing new drugs was based on empirical observations.

1.3.1 Identification of Complex Structures Began in the Early 1900s

The early 1900s witnessed an explosion in the elucidation of structures of bioactive molecules. Earlier, substances were described by the methods by which they were isolated, by their appearance, and by simple characterization, such as melting point, rather

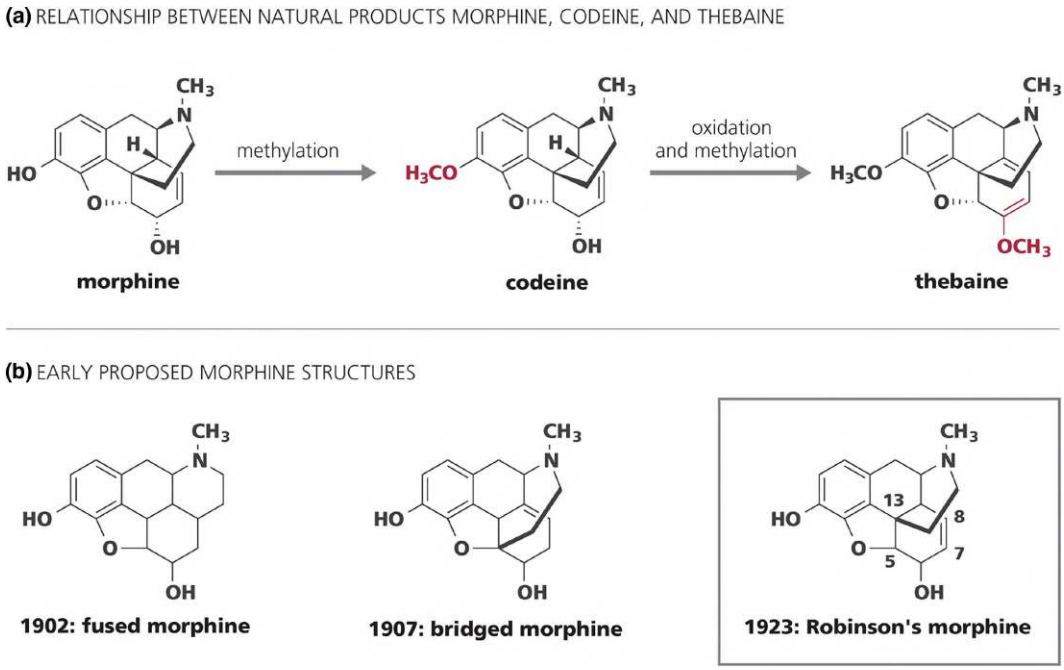
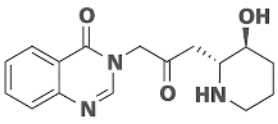


Figure 1.3 Elucidation of morphine structure. (a) Relationship between the natural products morphine, codeine, and thebaine; structural differences in each step are shown in red. (b) Initially proposed morphine structure containing fused ring system; structure modified to bridged structure; and correct structure [which modifies the ring fusion and double-bond positions compared to 1907 structure].

than their specific chemical structures. Without the benefit of modern spectroscopic methods, these complex structures were solved primarily by use of elemental analyses, degradation studies, and simple functional group manipulations to deduce which functional groups were present. For example, from 1902 to 1907, early work in Germany by four separate laboratories identified the empirical formula of morphine, as well as some features of the structure. Natural products isolated from the poppy plant included morphine ($C_{17}H_{19}NO_3$), codeine ($C_{18}H_{21}NO_3$), and thebaine ($C_{19}H_{21}NO_3$). Simple analysis of its acid–base properties showed that while morphine possessed a phenolic (acidic) alcohol, codeine and thebaine did not (**Figure 1.3a**). The presence of a methylated basic nitrogen in each was also indicated. Methylation of the phenol group of morphine provided codeine, and oxidation of the alcohol of codeine to a ketone, followed by the formation of the methyl enol ether, afforded thebaine. Other important common structural features identified were the presence of an alkene and the fact that degradation of each afforded a phenanthrene structure (based on three fused benzene rings). On the basis of knowledge that the three core structures were related and further studies of the properties and reactions of these compounds, a pentacyclic structure for morphine was proposed in 1902 (shown in **Figure 1.3b**). Soon after, in 1907, a bridged, rather than a fused, ring system was suggested. This structure was accepted as correct for nearly 20 years, until, in 1923, Sir Robert Robinson published the structure in which the bridge junction is at C13 rather than C5 and the alkene is between C7 and C8 rather than C8 and C9. Ultimate confirmation that Robinson's structure was correct came in the 1950s with Marshall Gates's independent synthesis of morphine.

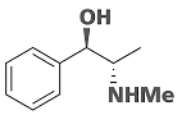
Overall, the elucidation of the structure of morphine took approximately 30 years and the work of many scientists from independent laboratories. **Figure 1.4** illustrates other medicinally active compounds (many listed in **Table 1.1**) that were isolated and whose structures were elucidated during this period.

Since the early 1900s, purification procedures have been greatly improved with the advent of modern chromatographic methods. Sources of natural products have also expanded beyond plant materials to include products from bacteria, fungi, and marine sources. In addition, the development of infrared spectroscopy (IR), nuclear magnetic resonance spectroscopy (NMR), and mass spectrometry (MS) allows the use of very small quantities of material to rapidly determine the presence of functional groups, molecular connectivity, and accurate masses, respectively. With these spectroscopic methods, as well as the advances in X-ray crystallography, complex structures are now identified much more quickly and with greater confidence. For example, the complex



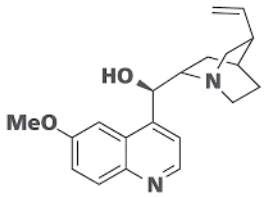
febrifugine

USE: fever
SOURCE: *Dichroa febrifuga*



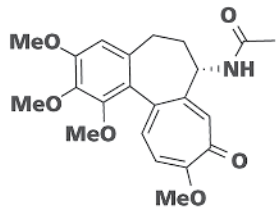
ephedrine

USE: bronchial dilation
SOURCE: *Ephedra sinica*



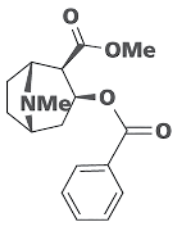
quinine

USE: malaria
SOURCE: *Cinchona*



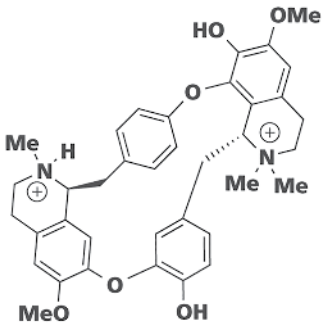
colchicine

USE: gout
SOURCE: *Colchicum autumnale*



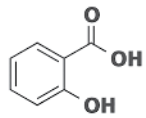
cocaine

USE: stimulant
SOURCE: *Erythroxylum coca*



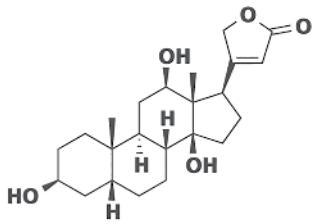
tubocurarine

USE: paralytic
SOURCE: *Chondrodendron tomentosum*



salicylic acid

USE: fever, pain
SOURCE: *Salix alba*



digoxigenin

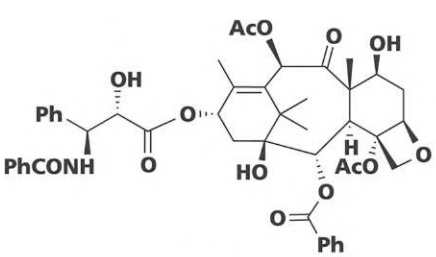
USE: cardiac stimulant
SOURCE: *Digitalis purpurea*

structures of taxol, isolated from yew trees in 1967, and of calicheamicin, isolated in 1987 from soil microbes, were solved within just a few years of their isolation and the characterization of their potent anti-proliferative activity. Mansukh Wani and Monroe Wall at Research Triangle Institute reported the structure of Taxol (Figure 1.5) in 1971 based on NMR, IR, and MS, along with X-ray crystallography, of a simple derivative. Scientists at Lederle Laboratories used classical degradation studies as well as spectroscopy and X-ray crystallography to solve the structure of the highly complex enediyne known as calicheamicin γ1 (see Figure 1.5) in 1992. Taxol (paclitaxel) is currently approved for the treatment of certain cancers. Calicheamicin γ1 is too toxic to be used as a stand alone agent, but a derivative was conjugated to an antibody, and that compound (gemtuzumab ozogamicin) is now approved to treat certain types of acute myeloid leukemia.

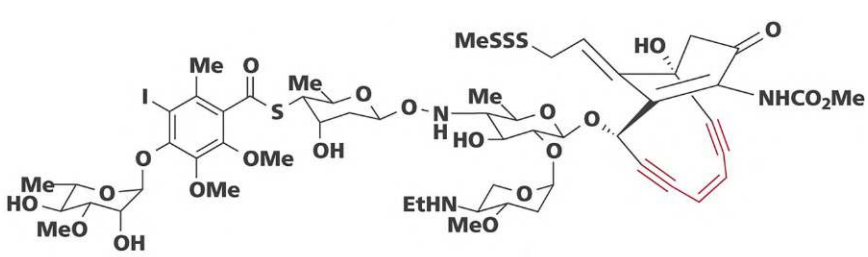
Figure 1.4 Examples of isolated active components. These structures were elucidated in the early 1900s, and their uses and plant sources are listed.

1.3.2 Advances in Total Synthesis Allow Access to Natural Products and Analogs

Once methods to reliably elucidate the structures of biologically active natural products were readily available, the synthesis of those structures and their analogs became feasible. The development of a **total synthesis**, the complete synthesis from simple starting materials, of a natural product is desirable and important for a number of reasons:



taxol

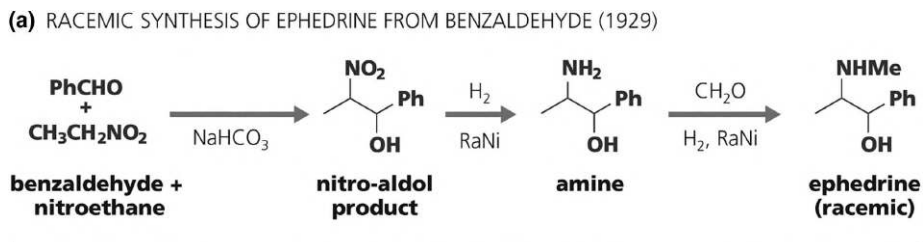


calicheamicin γ1

Figure 1.5 Structures of taxol and calicheamicin γ1. Taxol and calicheamicin γ1 are two recent examples of structurally complex bioactive natural products. Taxol binds and stabilizes microtubules and inhibits cell proliferation. The ene-diyne portion (shown in red) of calicheamicin γ1 is responsible for its ability to cleave DNA.

- Isolations can be time-consuming and expensive and may give variable results.
- Very large amounts of raw materials (kilograms) may be required to produce milligrams of the isolated product.
- Access to materials may be restricted due to either climatic or political conditions.
- Independent synthesis of the natural product serves to confirm the structure.
- Synthesis provides reliable access to new and unnatural analogs.
- Synthesis can be a source of large quantities of material.

There is, of course, no difference between a natural compound and a synthetic one. They are chemically identical in all respects, including stereochemistry. Often, analogs of natural products, which have variations in the functional groups, and sometimes in the core skeleton, maintain biological activity and in some cases provide an improvement over the original structure. Developing a synthesis of the natural product allows for a much larger and more diverse array of analogs than might be available directly from the natural product itself. For example, semisynthetic compounds, such as aspirin and heroin, are available by simple esterification of the alcohol functional groups present in the natural products. Since the number and reactivity of the functional groups in a given natural product can limit the number and diversity of analogs that may be prepared by semisynthesis, the modification of intermediates used in a total synthesis can be used to even further expand the analogs prepared. It can also provide optical isomers that are inaccessible from the natural product. For example, two syntheses of ephedrine, one providing the racemic form and one providing the natural (levo) stereoisomer, are shown in **Figure 1.6**.



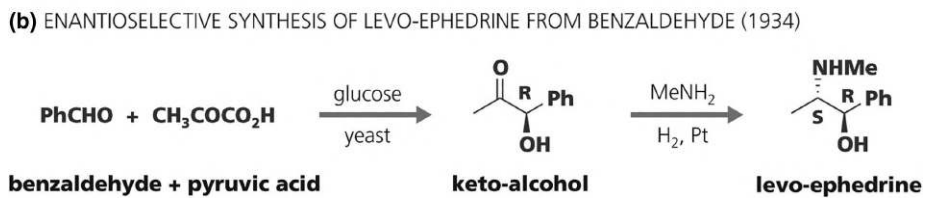
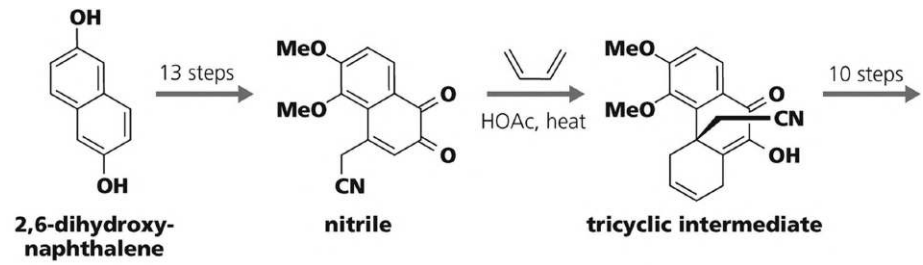


Figure 1.6 Syntheses of ephedrine. (a) Racemic synthesis from benzaldehyde. A nitro-aldol (Henry) reaction between nitroethane and benzaldehyde is followed by reduction of the nitro group to an amine; reductive amination with formaldehyde generates racemic ephedrine. (b) Enantioselective synthesis from benzaldehyde and pyruvate. Enzymes in fermenting glucose catalyze condensation of the starting materials to give a keto-alcohol with the *R*-configuration; reductive amination with methylamine affords levo-ephedrine selectively.



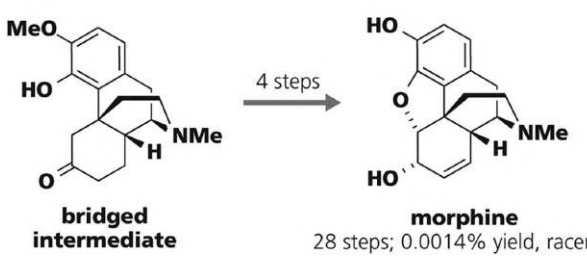
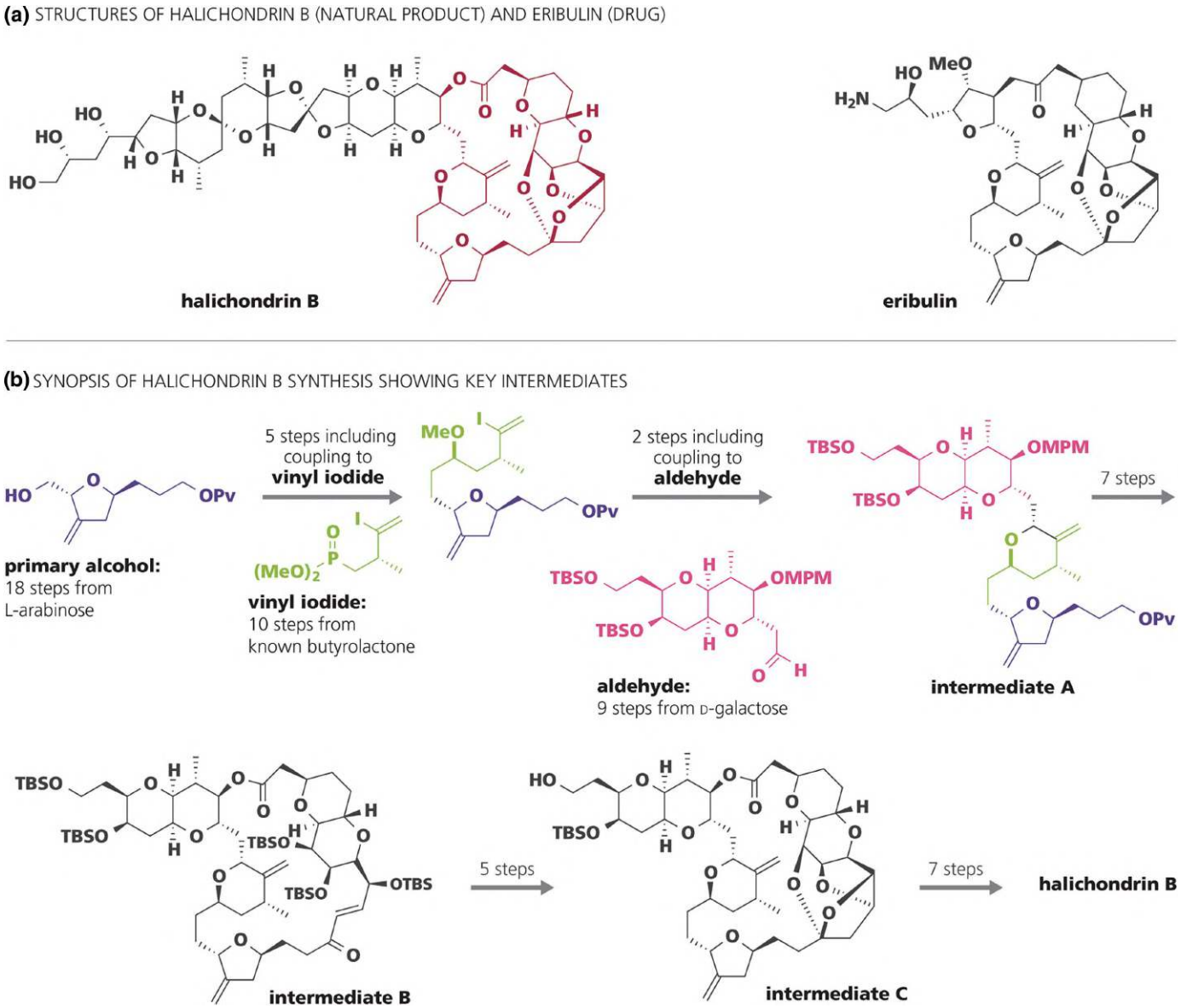


Figure 1.7 Outline of Gates's synthesis of morphine. Starting with 2,6-dihydroxynaphthalene, 13 steps adjust the oxidation level and add the two-carbon nitrile that will become the nitrogen-containing ring. A Diels–Alder reaction affords a tricyclic intermediate. Ten more steps are required to form the bridged intermediate, and then four final steps provide morphine.

Although the synthesis of ephedrine requires only a few steps, some total syntheses can be quite lengthy. For example, the first synthesis of morphine (by Gates) required 28 steps and afforded racemic material in only 0.0014% yield (see **Figure 1.7** for synopsis). A more efficient biomimetic synthesis was reported by Kenner Rice in 1980; it led to morphine in 18 steps with 11% yield. The first enantioselective total synthesis, one that did not rely on a step to resolve stereoisomers, was not accomplished until 1993. That synthesis, reported by Larry Overman, began with an enantioselective ketone reduction and proceeded in 24 steps with 5% yield.

A more recent example can be seen surrounding the efforts toward the total synthesis of halichondrin B, an anti-tumor product isolated from the marine sponge *Halichondria okadai* in 1986. Very limited quantities of the natural product were available from natural sources; therefore, a number of chemistry labs published their efforts to synthesize this highly complex molecule. The first total synthesis of halichondrin B was reported by Yoshito Kishi in 1992, with the longest linear sequence requiring 47 steps (see **Figure 1.8** for a synopsis; synthesis of intermediates not shown). Intermediates prepared during a total synthesis are routinely tested to determine which segments of the structure are essential for biological activity. Those efforts indicated that fragments of the right side of halichondrin B exhibited significant anti-proliferative activity. This

Figure 1.8 Total synthesis of halichondrin B and identification of eribulin. (a) Structures of halichondrin B and eribulin, a drug derived from an intermediate prepared during a total synthesis of the natural product. The moiety (right side) of halichondrin B from which eribulin was derived is shown in red. (b) Synopsis of halichondrin B synthesis. Three fragments, the primary alcohol (violet), vinyl iodide (green), and aldehyde (pink), are combined to give intermediate A. Seven steps convert A to intermediate B; five steps are required to form the spiroacetal intermediate C; and seven final steps prepare halichondrin B.



fragment was further modified and resulted in the anti-cancer agent eribulin, highlighting how the development of a synthesis of a complex natural product can lead to the generation of a related but unnatural anti-cancer agent.

In addition to the synthesis of natural products, the synthesis of medicines not derived from nature also developed in the early 1900s. One of the first synthetic drugs used was arsphenamine (Salvarsan), an arsenic-based compound (whose structure was not correctly elucidated until much later) developed by Paul Ehrlich and marketed by Hoechst to treat syphilis. It was used from 1910 until the 1940s, when it was replaced by much safer drugs. Another early synthetic drug, developed in 1935, was the antibacterial sulfanilamide, which is generated in the body after the administration of the synthetic dye prontosil (Figure 1.9). Currently, approximately 90% of approved drugs are small molecules. The majority of these are synthetic compounds, and about a quarter are natural products or semisynthetic derivatives of a natural product.

1.3.3 Medicinal Chemistry Currently Is Based on the Molecular Causes of Disease

In the not-so-distant past, diseases or disorders were often attributed to witchcraft, possession by the devil, or immoral behavior. It was only in more recent times that it was proved that certain diseases were caused by an infectious agent or an aberrant physiological process. For instance, the germ theory of disease, although suggested in some cultures for many hundreds of years, was not widely accepted until the 1800s. Studies of human physiology in the 20th century led to the basis of our current understanding of many conditions. Most recently, developments in DNA technology and sequencing of the human genome allowed the identification of genomic changes, and therefore the corresponding changes at the protein level, that are associated with disease.

Almost all drugs interact with a specific **macromolecular target**. The three major classes of targets are enzymes, receptors, and the nucleic acids DNA and RNA (these, and others, will be discussed in detail in Chapters 6–9). Ehrlich introduced one of the most important concepts in modern medicinal chemistry in 1897 when he proposed the side chain theory to explain how bacterial toxins might interact with antitoxins through a specific chemical character of the cells. In 1900, he replaced the *side chain* term with **receptor** and defined it as a specific chemical entity that could interact with drugs. He also introduced the concept of a **magic bullet** to target specific bacteria. This idea, that a drug could kill bacteria by specific interactions, without damaging the patient's cells, led to his development of Salvarsan, described earlier. The receptor concept was further advanced by John Langley, a British physiologist who studied the autonomic nervous system. He observed that a solution of jaborandi (an extract

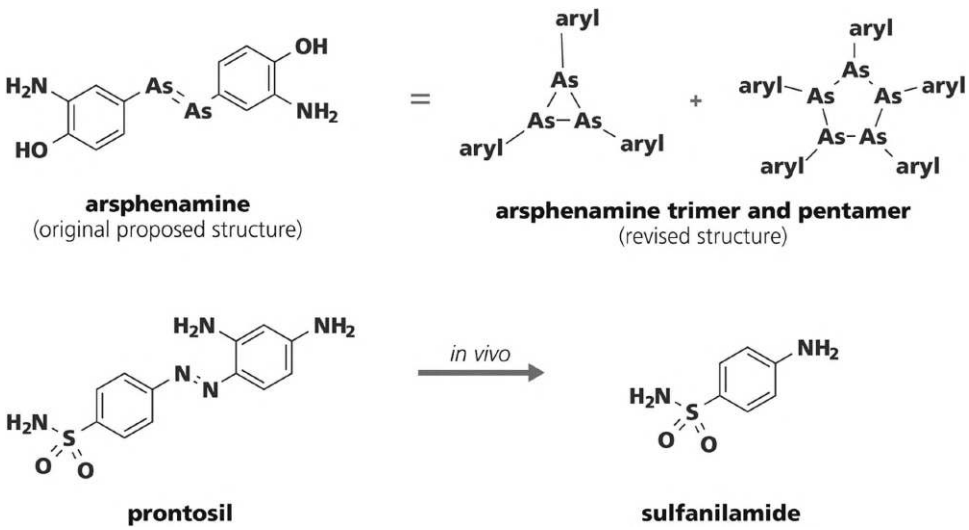


Figure 1.9 Examples of early synthetic drugs. Arsphenamine was used to treat syphilis: the proposed As=As structure (left) was revised in 2005 and shown to consist of a mix of trimer and pentamer structures. The antibacterial sulfanilamide is produced *in vivo* by cleavage of the azo bond (N=N) of the synthetic dye prontosil.

from a South American shrub), which contains the alkaloid pilocarpine, decreased the heartbeat of frogs when injected, but the heartbeat could then be increased by dosing with atropine, another natural product isolated from a plant called deadly nightshade (Figure 1.10a). Eventually, it was recognized that pilocarpine was a muscarinic receptor agonist (activator) and atropine was a muscarinic receptor antagonist (blocker): Both compounds affected the same receptor, but with opposite results.

Later in his career, in 1905, he observed a similar antagonistic action between nicotine and curare (Figure 1.10b). The injection of nicotine into a chicken led to a contraction of denervated leg muscle; injection of the plant extract curare, which contains tubocurarine, abolished the contraction. This work led to the concept that these drugs interact with a specific receptive substance on the muscle cells.

The receptor theory was not fully accepted until the 1930s, and with the late 1940s came the recognition that multiple subtypes of the same receptor could exist. This appreciation explained how similar compounds could exhibit different types of activities and portended a focus on selectivity for one receptor subtype over another, with the hope of developing drugs with fewer side effects. The first receptors were not actually isolated until the 1970s, and in just the last 40 years, there have been tremendous advances in our knowledge of drug targets, such as enzymes and receptors.

The history of enzymes parallels that of receptors, with Wilhelm Kühne's proposal of the term *enzyme* in 1878 to describe the component involved in the fermentation of sugar to alcohol; in 1907, Eduard Buchner named that specific enzyme zymase. Emil Fischer proposed the **lock and key hypothesis** to explain how a substrate interacts with an enzyme by fitting into a specific geometric shape as a key (substrate) does in a lock (enzyme). The identification and isolation of a number of enzymes followed throughout the 1900s, culminating with the report by David Phillips and colleagues of the first X-ray crystal structure of lysozyme in 1965, and followed a few years later (1969) by the first synthesis of an enzyme (ribonuclease) in a laboratory.

With combinations of technologies like X-ray crystallography, NMR spectroscopy, or cryo-electron microscopy together with computer graphics, we can now "see" a drug as it interacts with a specific enzyme or protein. In many cases, an X-ray crystal structure can allow for the design of new drugs based on observable

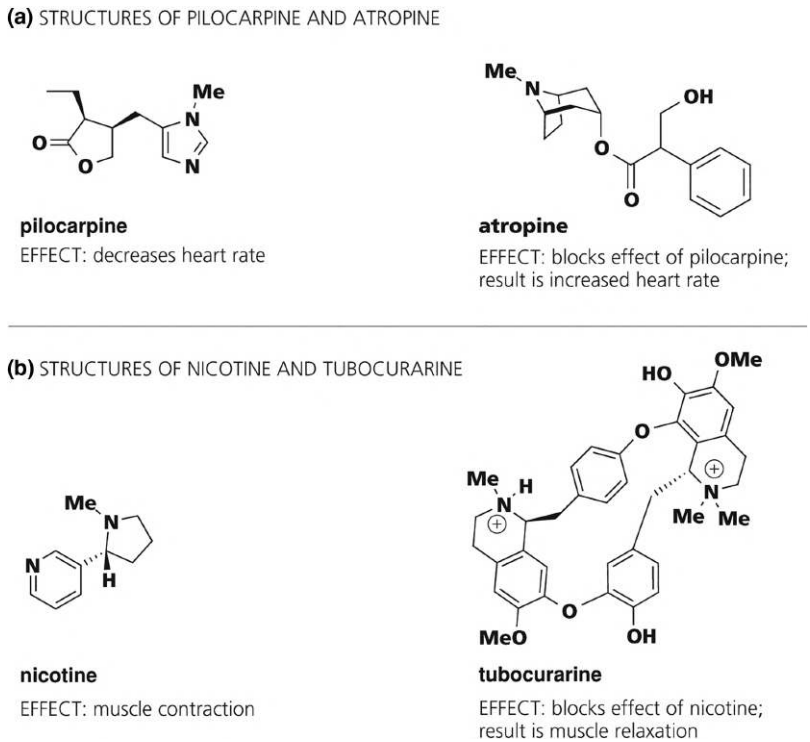


Figure 1.10 Antagonistic compounds. (a) Structures of pilocarpine and atropine, which led to Langley's proposal of receptor theory based on their effect on heart rate: pilocarpine decreases heart rate, while atropine blocks the effect of pilocarpine. (b) Structures of nicotine and tubocurarine, the active species in curare: curare blocks the effect of nicotine.

interactions. Although there are far fewer X-ray structures of receptors than of enzymes, the molecular modeling of receptors has also streamlined the process of drug discovery.

1.4 SUMMARY

While the remainder of this text will focus on current drug discovery practices, it is also important to review how the practice of medicinal chemistry began. Until the 19th century, medicines were, for the most part, either crude or semi-pure plant preparations. Even now, there is continuing interest in **ethnopharmacology**, the study of indigenous plants traditionally used to treat diseases by native populations. In the mid- to late 1900s, natural products continued to make important contributions as starting points for new drugs, particularly in both cancer and antibacterial drug discovery. With the rise of the scientific method and the development of modern chemistry came isolation and structural identification of the active components. This coincided with the development of synthetic compounds as medicines. New technologies, including synthetic biology and engineered biosynthesis, provide access to novel products with a natural origin. Continued research into the active agents and mechanism of action of traditional plant-based medicines, such as cannabis, will also provide new avenues for medicinal chemistry research.

In the 20th century, numerous advances were made in the understanding of mechanism and the molecular basis of disease. All this has led to the modern process of drug discovery, which is based on the idea of developing drugs to act with a specific biochemical mechanism to treat diseases. Most recently, advances in understanding diseases and diagnosis have provided access to precision medicine. For example, the genetic testing of tumors allows specific treatment protocols to be designed that will be most effective for a particular patient.

CASE STUDY 1 DISCOVERY OF ARTEMISININ

The discovery of the antimalarial drug artemisinin began in 1967 in China as a response to outbreaks of quinine-resistant malaria among the North Vietnamese during the Vietnam War. A research project (Project 523) was initiated at a number of institutes throughout the country to find new effective medicines, both by synthetic approaches and by screening traditional Chinese medicines. At the Institute of Chinese Materia Medica, Youyou Tu screened over 2,000 plant extracts and identified 640 that were active. About 380 extracts were screened in a mouse model of malaria infection. Extracts from *Artemisia annua* L., a species of wormwood known in China as qinghao, inhibited the growth of the disease-causing organism *Plasmodium falciparum* by 68%; however, the results were not reproducible in later experiments.

By consulting the ancient literature (*A Handbook of Prescriptions for Emergencies* by Ge Hong, AD 284–346), she found that the recommended extraction procedure was to "immerse a handful of qinghao with two liters of water, wring out the juice and drink it all." This information led to the idea that heating in the traditional extraction procedure was destroying the active component and might be responsible for the modest and irreproducible activity. The separation of the extract by typical acid extraction afforded acidic and neutral portions. The removal of the acidic portion of the extract left the neutral organic extract, which was nontoxic in assays and showed 100% inhibition of growth of the organism. Pure crystals were obtained, and the structure was confirmed by X-ray crystallography in 1979 as an unusual sesquiterpene lactone containing an endoperoxide (**Figure 1.11**).

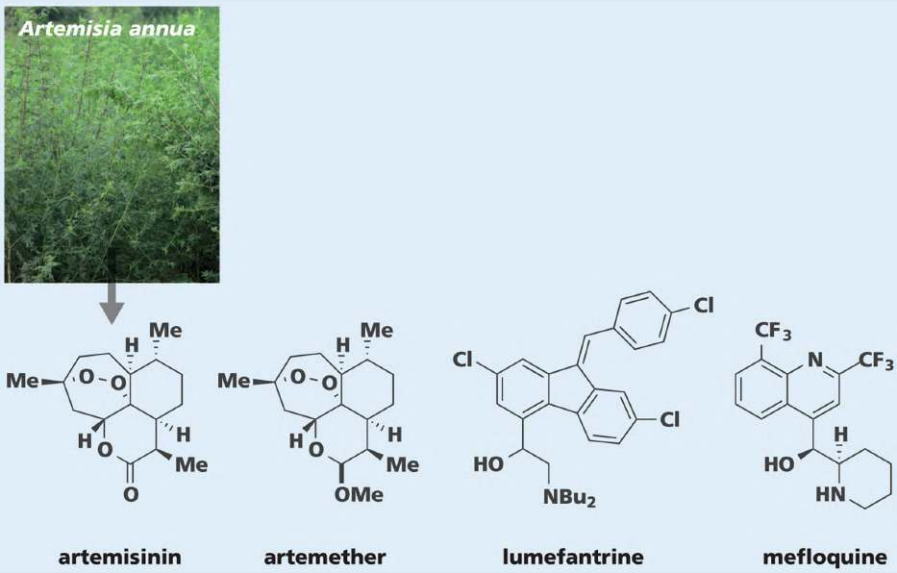


Figure 1.11 Structures of artemisinin and other antimalarial drugs. (Photo by Mark Tuschman, Novartis AG.)

Artemisinin acts rapidly but has a short half-life, and it is now used in combination with the synthetic drug mefloquine to treat malaria. The reduced derivative, artemether, is also commonly used in combination with the synthetic antimalarial drug lumefantrine.

This case study represents a relatively modern example of the discovery of a natural product as a new drug from an ancient remedy. Youyou Tu was recognized for this work with the 2015 Nobel Prize in Physiology or Medicine.

CASE STUDY 2 DISCOVERY OF TAXOL

The discovery and development of Taxol (paclitaxel) highlights the long and arduous process that it sometimes takes to get a drug approved. It involved natural product screening and structure elucidation, careful mechanistic studies that uncovered a unique mechanism of action, total synthesis and semisynthesis efforts to address an issue of limited supply, and eventual agreements between government and industry to move the drug forward.



Figure 1.12 Bark of the Pacific yew tree *Taxus brevifolia*. Photo courtesy of Susan McDougall

In 1962, a US Department of Agriculture botanist, Arthur Barclay, working under contract for the US National Cancer Institute (NCI), collected stem and bark samples from the Pacific yew tree (*Taxus brevifolia* Nutt) in Washington State (**Figure 1.12**). Extracts from these samples were found to be cytotoxic to KB cancer cells in a typical assay to predict anti-cancer activity. In 1965, through another NCI contract, a large sample of bark was sent to Monroe E. Wall at the Research Triangle Institute. The isolation and bioactivity of Taxol were reported by Wall and his collaborator, Mansukh Wani, at the 1967 American Chemical Society Meeting; however, its structure was not elucidated until 1971.

Initial enthusiasm for the bioactivity of Taxol and its potential as a drug was not high. Its potency in *in vitro* and *in vivo* leukemia assays was modest, it had very poor solubility in water, and it was only available in small quantities from the bark of the uncommon and slow-growing Pacific yew tree. However, additional *in vivo* testing in B16 melanoma models provided more promising results and led scientists to pursue it further. In 1977, the NCI asked Susan B. Horwitz of Albert Einstein College of Medicine to study the mechanism of Taxol. She received 10 mg of material, and in a short time, her lab observed that Taxol-treated cells gave rise to a unique phenotype of highly stable microtubule bundles. Further study revealed a unique mechanism of action: the promotion of tubulin polymerization and stabilization of microtubules. Phase I studies of Taxol were initiated in 1984, and phase II studies in 1985, despite significant limitations in drug supply. In 1989 and 1991, positive clinical results in ovarian and breast cancer were reported.

The promising results from clinical trials made solving the supply issue even more urgent. The Pacific yew is a slow-growing conifer found in old-growth forests in the Pacific Northwest. In the late 1980s, the NCI acquired ~120,000 lb of dried Pacific yew bark, which was derived from approximately 7,000 trees, and afforded only 18 lb of pure Taxol (0.015% yield). It was estimated that six 100-year-old trees would be needed to treat one patient, and the relative scarcity of tree bark, as well as the damage that bark removal caused the tree, was a serious concern. A significant effort by a number of synthetic organic chemistry labs around the world focused on developing a scalable synthesis of Taxol to alleviate the supply issue. The problem was eventually solved via semisynthesis. In 1988, Pierre Potier at the Centre National de la Recherche Scientifique (CNRS) in France reported that 10-deacetylbaccatin III was readily available from the needles of a very common shrub called *Taxus baccata* and could be converted to Taxol-like structures. Since needles could be harvested without killing the plant, unlike the bark, this was a much more desirable source plant. Robert Holton at Florida State University developed an improved method for semisynthesis using the general strategy shown in **Figure 1.13**. The key step was acylation of the sterically hindered C13 alcohol with a reactive β-lactam to introduce the side chain.

The NCI granted a Cooperative Research and Development Agreement (CRADA) to Bristol-Myers Squibb (BMS), and they took over the development of Taxol. The FDA approved Taxol (paclitaxel) on December 29, 1992, 30 years after the extract was first collected, 21 years after the structure was elucidated, and 8 years after the first clinical

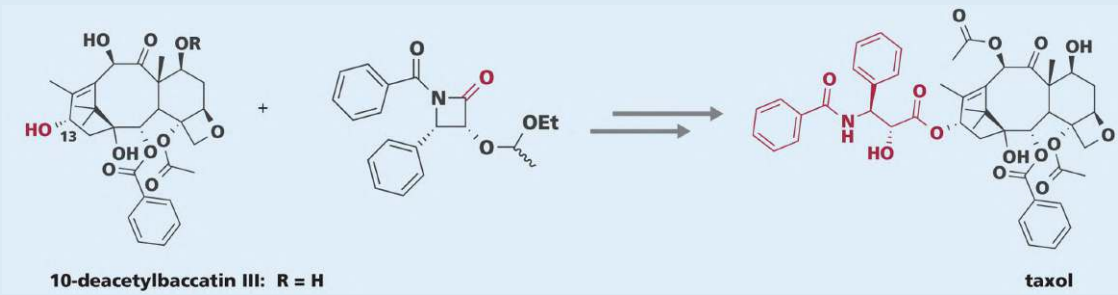


Figure 1.13 Key step in taxol semisynthesis. 10-Deacetylbaccatin III was found to be readily available from the needles of a common yew tree, and could be converted to Taxol-like structures. The Taxol side chain was coupled to 10-deacetylbaccatin III, by acylation of the C13 hydroxyl with a β -lactam. The C13 alcohol and the β -lactam carbonyl are shown in red in the starting materials, with the side chain of the coupled product also in red.

trials were initiated. Taxol is approved to treat a number of cancers involving solid tumors, including breast, ovarian, bladder, and lung cancers.

Taxol continues to inspire scientific work. The total synthesis of Taxol was reported by a number of academic groups, but its use to provide the quantities needed for development and marketing was never feasible. Analogs of Taxol (docetaxel and cazabitaxel) and other microtubule-stabilizing agents (ixabepilone) are important drugs to treat cancers.

1.5 REVIEW QUESTIONS

1. Describe two advances that were made in the field of medicinal chemistry during the period from 1700 to 1900.
2. Give two examples of semisynthetic drugs and the corresponding natural products.
3. What major advances were proposed by Ehrlich and Langley?
4. What are some reasons for the development of a total synthesis of a bioactive natural product?
5. Describe how recent advances in small-molecule NMR and MS have impacted the field of natural product drug discovery.
6. Describe the impact of X-ray crystallography on modern medicinal chemistry.
7. Review **Figure 1.1**.
   a. If the product of interest contained an acid functional group rather than a secondary amine, how would the flowchart change?
   b. Draw a modified flowchart of **Figure 1.1** to show how one could isolate salicylic acid in general terms.

1.6 APPLICATION QUESTIONS

1. The isolation of camptothecin was reported in 1966 (Wani MC, Wall ME et al. [1966] Plant anti-tumor agents. I. The isolation and structure of camptothecin, a novel alkaloidal leukemia and tumor inhibitor from *Camptotheca acuminate*. *J Am Chem Soc* 88: 3888–3890). Summarize the isolation procedure and important features of the structural identification.
2. Cocaine was first isolated in 1860 by Niemann, who also made note that it "numbed the tongue." About 20 years later, von Anrep recognized its use as a local anesthetic after trials on both animals and himself. By the late 1800s, cocaine was in use both as an anesthetic and as a stimulant, with Freud promoting its use as an antidote to morphine addiction. Freud, and many others, became addicted to cocaine, and a search for a better local anesthetic led to the synthetic caine anesthetics. The first of these was amylocaine in 1903, and the second was novocaine in 1905.
   a. Identify two structural features that are common to cocaine, amylocaine, and novocaine.
   b. On the basis of your knowledge of organic chemistry, propose a synthesis of novocaine, beginning with toluene and using *N*,*N*-diethylethanolamine.
3. Explain some of the challenges of using medications directly from plant sources (such as roots and leaves).
4. Research and identify two approved drugs that are natural products and that are commercially produced directly from isolation of the natural product.
5. Research and identify two approved drugs that are natural product derivatives and that are commercially produced through semisynthesis.
6. Research and identify two approved drugs that are natural products or close derivatives and that are commercially produced through total synthesis.
7. Research and identify two approved drugs that are natural products and that are commercially produced through fermentation.
8. The Nobel Prize in Physiology or Medicine has been awarded to several scientists specifically for their contributions to medicinal chemistry. Identify these scientists and describe their contributions.

FURTHER READING

A General Source for Active Components of Many Herbal Preparations

Varro T & Foster S (1999) *The Honest Herbal: A Sensible Guide to the Use of Herbs and Related Remedies*, 4th ed. The Haworth Herbal Press.

Aspirin

Bentley R (1873) *A Manual of Botany: Including the Structure, Classification, Properties, Uses and Functions of Plants*. J & A Churchill.

Brodie BC (1811) Experiments and observations on the different modes in which death is produced by certain vegetable poisons. *Philos Trans R Soc London* 101:178–208.

Mahdi JG (2010) Medicinal potential of willow: a chemical perspective of aspirin discovery. *J Saudi Chem Soc* 14:317–322.

Morphine

Gates M & Tschudi G (1952) The synthesis of morphine. *J Am Chem Soc* 74:1109–1110.

Knorr L & Horlein H (1907) Uber die haftstellen des stickstoffhaltigen nebenringes im kodein und uber die konstitution der morphiumalkaloide. XII. Mittellung: zur kenntnis des morphins von Ludwig Knorr. *Ber Dtsch Chem Ges* 40:3341–3355.

Pschorr R & Einbeck H (1907) Zur konstitution des morphins: uber die konstitution des oxymethylmorphimethins. *Ber Dtsch Chem Ges* 40:1980–1983.

Robinson R & Gulland JM (1923) The morphine group. Part I: a discussion of the constitutional problem. *J Chem Soc C* 123:980–998.

Halichondrin

Aicher TD, Baszek KR, Fang FG et al. (1992) Total synthesis of halichondrin B and norhalichondrin B. *J Am Chem Soc* 112:3162–3164.

Damle NK & Frost P (2003) Antibody-targeted chemotherapy with immunoconjugates of chalicheamicin. *Curr Opin Pharmacol* 3:386–390.

Dong C-G, Henderson JA, Kaburagi Y et al. (2009) New syntheses of E7389 C14–C35 and halichondrin C14–C38 building blocks: reductive cyclization and oxy-Michael cyclization approaches. *J Am Chem Soc* 131:15642–15646.

Maehle A-H, Prull C-R & Halliwell RF (2002) The emergence of the drug receptor theory. *Nat Rev Drug Discov* 1:638–641.

Top pharmaceuticals: from aspirin to Viagra and more. (2005) *Chem Eng News* 83:1–162.

Case Studies

Borman S (2007) Drugs from academia. *Chem Eng News* 85:42–46. https://www.acs.org/content/acs/en/education/whatischemistry/landmarks/camptothecintaxol.html

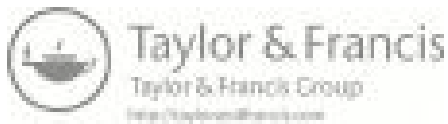
Horwitz SB (2004) Personal recollections on the early development of taxol. *J Nat Prod* 67:136–138.

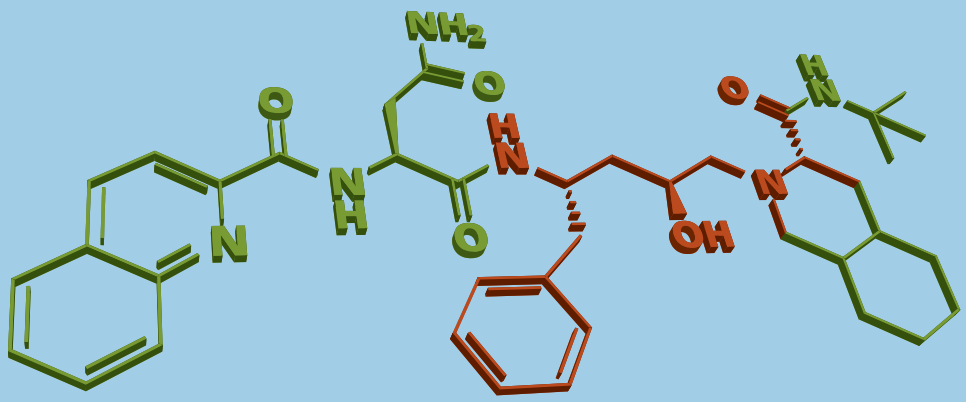
Kingston DGI (2001) Taxol, a molecule for all seasons. *Chem Commun*:867–880.

Miller LH & Xinzhuan S (2011) Artemisinin: discovery from the Chinese herbal garden. *Cell* 146:855–858.

Oberlies NH & Kroll DJ (2004) Camptothecin and taxol: historic achievements in natural products. *J Nat Prod* 67:129–135.

Tu Y (2011) The discovery of artemisinin (qinghaosu) and gifts from Chinese medicine. *Nat Med* 17:1217–1220.


Taylor & Francis
Taylor & Francis Group



Drug Discovery: Hit and Lead Identification

2

LEARNING OBJECTIVES

- Learn the drug discovery process.
- Understand the concept of hits and leads in the context of drug discovery.
- Become familiar with the types of assays used for hit characterization.
- Describe various sources of hits and leads.
- Explain how modern small-molecule libraries are constructed.
- Understand various methods to screen compound libraries.
- Learn common tools and strategies used to recognize high-quality hits.

2.1 INTRODUCTION

How do we go about finding a cure for cancer, or treatments for infectious diseases or heart disease, or remedies for any other conditions that affect us? As described in Chapter 1, until the early 1900s, most drugs were products of nature and were based on the historical and cultural use of plant or animal materials. Today, a number of different strategies are used to discover new drugs. Some new drugs are closely related to older drugs or historical plant products, while others arise from serendipitous observations of effects in humans. The most common strategy in the modern era, however, is to start at a much earlier stage with compounds that exhibit some effect in a biological assay and then, through a process of optimization, eventually develop a compound that exhibits all the appropriate properties of a drug candidate. This chapter will highlight the multiple different pathways to a drug candidate, with a focus on the strategies used to identify the earliest starting points.

2.2 OVERVIEW OF DRUG DISCOVERY PROCESS

Today, the most common method of drug discovery is a process that starts by identifying a small-molecule starting point (a hit) and then optimization of its biological, physical, and pharmaceutical properties to generate a lead, and eventually a drug candidate, through an iterative process of design, synthesis, and testing. For the sake of consistency and clarity, a **hit** is defined as a small molecule of confirmed structure and purity that exhibits biological activity at some defined threshold (typically <20 μM) in a particular biological test. Since not all hits are useful starting points for medicinal chemistry, the term **validated hit**, which connotes stricter criteria, is sometimes used

DOI: 10.1201/ 9781003465607-3

to describe a molecule that has reproducible, dose-dependent biological activity after its resynthesis and works through a desirable mechanism of action.

For example, in the development of a drug to treat hepatitis C, a *hit* may be any substance of a known structure that is 90% pure and is capable of decreasing viral load in infected cells by a defined level (such as 50%) at a concentration of 20 μM. A *validated hit* would be one of the hits that was identified earlier but had also been resynthesized and retested to confirm its activity. Its activity would be **dose-dependent**—meaning, there is a change in potency with a change in concentration: A concentration of 50 μM of the compound would show greater inhibition, and a concentration of 5 μM would show less inhibition. In some cases, additional characterization of the molecule's mechanism of action may also have been performed to verify that the activity was due to a viral infection–specific mechanism and not due to nonspecific effects.

Hit triage, the process of evaluating hits, eliminating some and validating others, is the first step toward identification of those compounds that are most likely to be optimized into drug candidates. Validated hits are often the starting points for medicinal chemistry activities, and they evolve through a **hit-to-lead process**, during which the goal is to characterize their attributes and liabilities and to identify and prioritize those that are most likely to be improved further during the optimization process. The product of this process, a **lead** (or **lead series** if a set of structural analogs is identified), is typically defined as a compound that exhibits the following qualities:

- Potency (often <1 μM)
- Some degree of selectivity toward designated or related biological targets or pathways
- A lack of promiscuity against unrelated targets
- Evidence of emerging structure–activity relationships (SAR)
- Appropriate pharmaceutical and physical properties, such as solubility, permeability, and stability
- Accessibility by either isolation, semisynthesis, or synthesis

Continuing with the theoretical drug to treat hepatitis C, the hit-to-lead process surrounding the validated hit that inhibited a viral enzyme at 20 μM would have been approached on multiple levels. Structural modification of one of the validated hits through preliminary medicinal chemistry activities may have led to a new compound with inhibition at 1 μM, which would have illustrated evidence of a correlation between structural modifications and activity and, at the same time, proved that analogs were synthetically accessible. This lead would have been evaluated in a number of assays to evaluate its physical properties (stability, aqueous solubility, permeability, for example), and those properties would be deemed acceptable. In addition, biological testing to further characterize its mode of action, as well as selectivity, would have been performed and proven that its mechanism was specific to the virus or viral infection processes and the lead had an acceptable selectivity index. When a drug is developed from a lead, chemists, biologists, and other scientists work together in the process of **lead optimization**, wherein modifications of the structure are made in an effort to improve activity and selectivity, reduce toxicity, and optimize physical and pharmaceutical properties. The process of lead optimization ultimately results in a **clinical candidate**, which is a compound with activity and safety levels sufficient for testing in humans and that is appropriate for development into a drug. The process of **drug development** focuses on characterization of the clinical candidate in a series of animal tests, as well as the development of the drug substance through optimization of its synthesis and the development of an appropriate formulation. **Clinical trials** prove the efficacy and safety of a candidate in human subjects and are required for regulatory approval by a government agency, such as the US Food and Drug Administration (FDA). If successful, these studies ultimately result in approval of the molecule as a **drug** (Figure 2.1).

It should be noted that a molecule might enter into this process at any stage. For example, natural products such as paclitaxel may enter into the process at the stage of lead or clinical candidate. Drugs that have found an unexpected use at the time of clinical trials may enter the pipeline as a clinical candidate and be **repurposed**, that

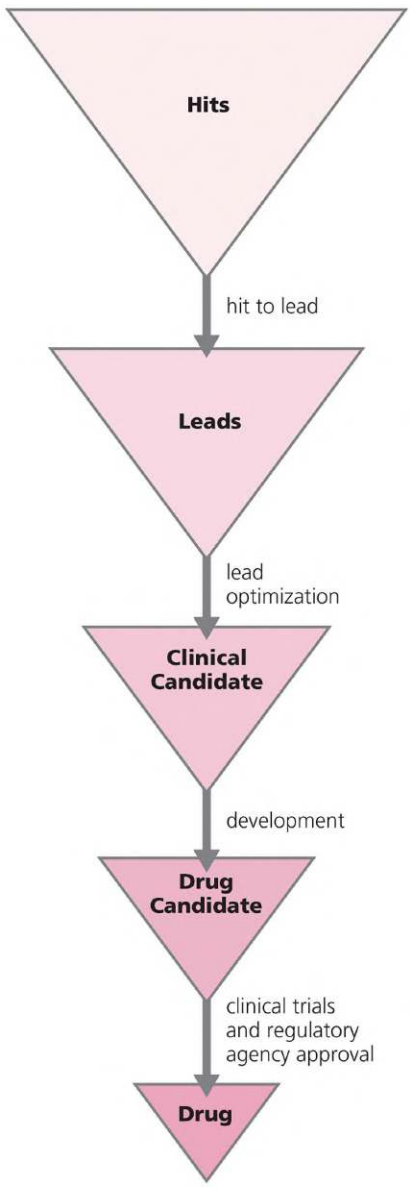


Figure 2.1 Overview of the drug discovery process, showing progression from a large number of hits to identification of a drug. The hit-to-lead process generates leads. The lead optimization process improves potency, safety, and properties of the lead compounds in order to generate clinical candidates. Clinical candidates undergo extensive testing in animals; formulations and syntheses are developed to provide a drug candidate, which is tested in clinical trials. Successful clinical trials and regulatory approval are required before a compound becomes a drug.

is, used for a different disease than was originally intended. However, most molecules require the entire progression of activities. Lead optimization to generate a clinical candidate and as well as medicinal chemistry strategies, the development process, clinical trials, and the approval process are described in Chapters 3–5. The earliest aspect of the drug discovery process is initial identification of hits and leads from a variety of sources, which is the subject of this chapter.

2.3 IDENTIFICATION OF HITS AND LEADS

The entire process of hit and lead discovery relies on the identification and characterization of bioactivity, regardless of the sources of hits and leads and where they enter the drug discovery pipeline. Prior to the mid-1900s, the discovery of medicines was predominantly based on the determination of bioactivity in either animals or humans (*in vivo*) or, in the case of antibacterial drugs, on the inhibition of bacterial growth. With the knowledge that specific macromolecular targets such as enzymes and receptors are involved in diseases, discovery shifted to testing for activity in enzyme or receptor assays (*in vitro*). **Drug targets** are the macromolecules, most often proteins or nucleic acids, that the drug affects in the body.

2.3.1 Various Types of Assays Are Used during the Hit and Lead Identification Process

In the modern era, an essential component of hit and lead identification is the availability of a biological assay that detects and measures the effects or activity of a small molecule in a biological system. These assays range in complexity and relevance from those that measure an effect, such as death of a cancer cell at one extreme, to those that work at the molecular level and measure the interactions between two proteins that may initiate a cascade of further events. In the latter case, the outcome measures a phenomenon that is quite remote from the final desired event; in the former case, the death of a cancer cell is the exact desired final outcome of a potential drug. Drug discovery efforts typically start by using one or several assays to identify and characterize small molecules with specific effects. While a variety of different technologies have been applied to these assays, most fall within four categories:

1. **Enzymatic or biochemical assays** typically monitor the formation of the product of an enzymatic reaction. In many cases, the enzyme itself is isolated and purified, and the assay measures its inhibition (or activation) by detecting, directly or indirectly, the amount of product formed. The nature of enzymes allows a thorough characterization of the kinetics of the reaction and of any inhibitors or activators identified.
2. **Binding assays** detect interaction between two molecules. Under physiological conditions, this binding event often initiates a cascade of biological responses, and therefore, measuring binding can identify molecules that either mimic the same response or perhaps inhibit the downstream events. The binding event can occur between a small molecule and a single macromolecule, like an enzyme, or nucleic acids, or between two macromolecules, such as two partner proteins or DNA and a transcription factor. It is not necessary for the macromolecule binding partner(s) to be isolated and purified for the binding event to be measured.
3. **Functional assays** generally measure downstream effects of a binding event. Examples include those that measure changes in levels of certain ions or endogenous molecules, such as cyclic adenosine monophosphate (cAMP), transcription events, and protein levels.
4. **Phenotypic assays** measure changes in phenotype or an effect in a complex system, regardless of the specific molecular target. They can be performed in cells and even in higher organisms, such as yeast, zebrafish, and flies. Examples include assays that monitor whether a molecule can cause the death of a cancer cell through a specific mechanism, such as apoptosis; can protect cells against

infection by virus or bacteria; or can repair a defect in genetically altered zebrafish. In the broadest sense, most animal models can be considered phenotypic screens. For example, the historical examples of atropine and nicotine described in Chapter 1 measured changes in heart rate or contraction of a muscle after the application of a certain compound. These assays reflected a specific phenomenon that was the result of a complex mechanism.

The specific strategies and assays used to identify hits and leads will depend on a variety of factors. These may include the specific disease, the level of understanding of the disease and its causes, and (when available) the nature of the drug target, as well as the available knowledge surrounding that target and any of its ligands.

2.4 SOURCES OF HITS AND LEADS IN DRUG DISCOVERY: COMPOUNDS WITH KNOWN BIOACTIVITY

Traditional sources for hits and leads include plant metabolites and microbial products, as well as the natural ligands for enzymes and receptors found in our bodies. Others include bioactive compounds reported in the scientific literature, in addition to clinical observations that identify unexpected activity in known bioactive compounds. While drug discovery based on natural products still occurs today, it is less common than in the past. Examples include the first statins for lowering cholesterol, such as lovastatin; cancer drugs, such as paclitaxel; and immunosuppressants, such as sirolimus. An analysis of small-molecule drug approvals in the recent past shows that natural products or natural product derivatives still account for some new drug approvals. Of the majority that are synthetic drugs, many can trace their origin or inspiration to a natural product. Examples include the cholesterol-lowering agent atorvastatin, which evolved from knowledge of the important structural features of the natural products lovastatin and compactin, and ixabepilone, a semisynthetic cancer drug that is an analog of the natural product epothilone.

2.4.1 Plant-Based Sources Are the Starting Points for Many Drugs

Plants were the earliest source of medicines, and today many of those same active compounds are still used. In some cases, the active component of the plant material is an approved drug, but more often, that compound is the lead for a less toxic or more potent derivative. Ephedrine, described in Chapter 1, is a natural product from the plant *Ephedra sinica* with stimulant properties and was the starting point used to develop the drugs amphetamine, methamphetamine, and methylphenidate. While methamphetamine is no longer an approved drug due to its addictive properties, amphetamine and methylphenidate are used to treat attention deficit hyperactivity disorder (ADHD) and narcolepsy (**Figure 2.2**).

In a similar fashion, meperidine and fentanyl were developed on the basis of the structure of morphine, the analgesic isolated from the opium poppy. A more recent example is paclitaxel, isolated from the Pacific yew tree, whose activity was first detected in a phenotypic screen of an extract. Paclitaxel itself is used as an anti-cancer drug and also provided the starting point for analogs such as docetaxel and cabazitaxel, both drugs approved to treat certain cancers.

As described in Chapter 1, screening of crude plant extracts typically is a slow process involving heating in an organic solvent, followed by fractionation by liquid–liquid extraction into increasingly polar solvents. Purification of each subfraction by multiple chromatographic steps ultimately provides pure compounds. Bioassays, most often phenotypic assays, are done at each level of purification to identify the active fractions. Once pure compounds are obtained, it may take considerable effort to elucidate novel structures.

The modern techniques that are common today to rapidly assay for activity against pure macromolecular targets (biochemical and binding assays) are not readily applicable to crude extracts that often contain tens to hundreds of compounds. Some

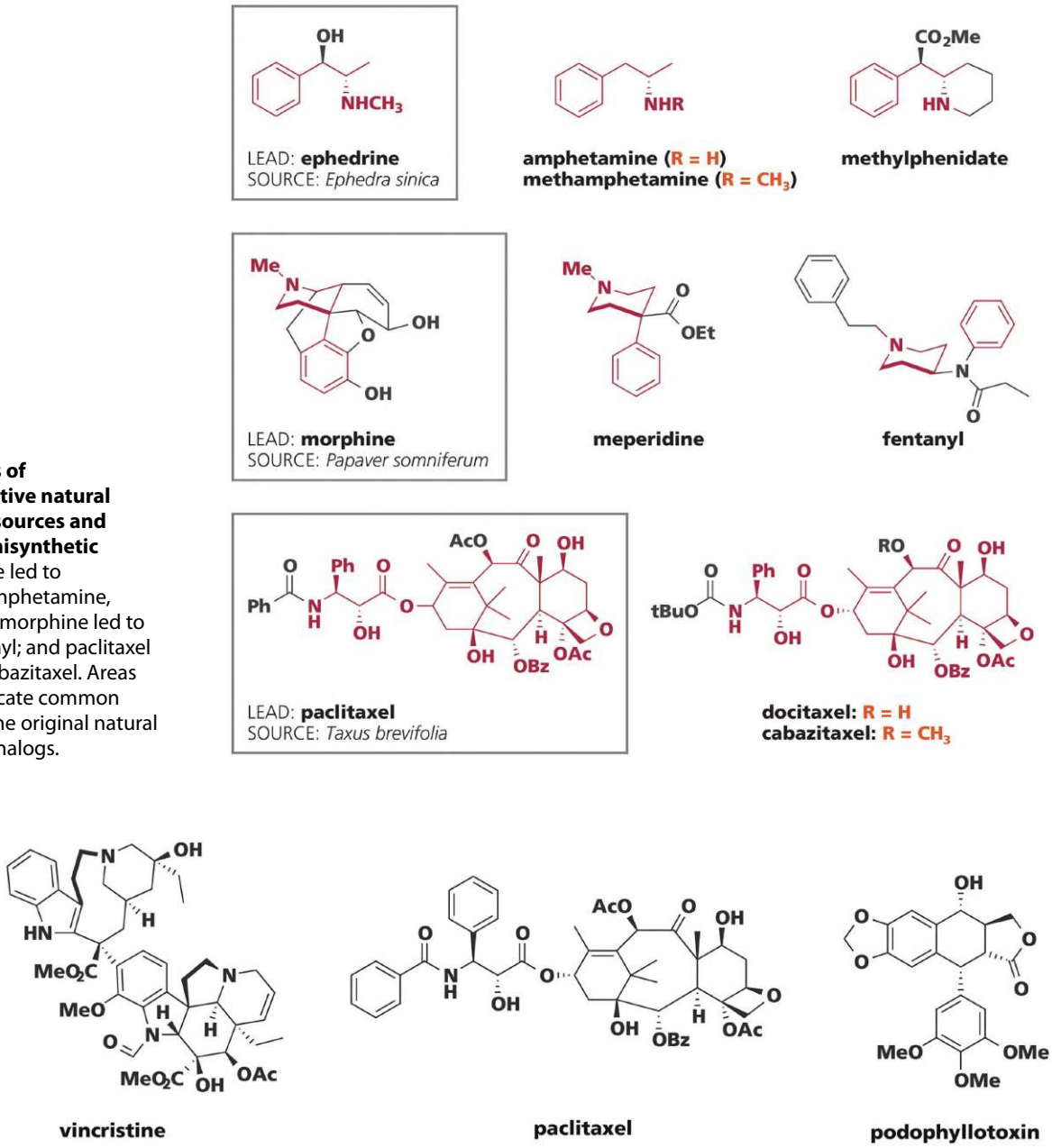


Figure 2.2 Structures of pharmacologically active natural products from plant sources and their synthetic or semisynthetic derivatives. Ephedrine led to amphetamine, methamphetamine, and methylphenidate; morphine led to meperidine and fentanyl; and paclitaxel led to docitaxel and cabazitaxel. Areas highlighted in red indicate common structural features of the original natural product and derived analogs.

Figure 2.3 Structurally diverse anti-tumor natural products. Vincristine, paclitaxel, and podophyllotoxin all bind to the protein tubulin, although at different sites.

problems include a lack of knowledge of the concentrations of active components and the need to remove unwanted compounds, such as tannins and chlorophyll, that cause interference with the assay. In addition, rediscovering known compounds is common, isolating or generating sufficient material for all the testing required can be challenging (or impossible), and government regulations required to get access to and/or use the source materials may be complicated and time-consuming. However, there has been increasing recognition that there are also many advantages to natural products, including high structural diversity, druglike properties, and evolutionary selection for biochemical specificity. Figure 2.3 shows some of the diverse structures of compounds, all plant-derived natural products, including paclitaxel, vincristine, and podophyllotoxin, that exhibit anti-tumor activity by binding to the same macromolecular target, tubulin.

Newer techniques to access these natural products include the fields of combinatorial biosynthesis and synthetic biology, which have developed to generate analogs of natural products by genetic manipulation of the biosynthetic machinery of the

organism. In addition, advances have been made in the preparation of pre-fractionated samples of plant extracts (e.g., partially purified extracts) that are adapted to the newer screening methods, allowing for rapid testing of natural products as mixtures of a few compounds rather than hundreds of compounds.

2.4.2 Microbes and Higher Animals May Be Sources for Hits, Leads, and Drugs

In addition to plants, microbes are an important source of pharmacologically active compounds, with the first compounds having been isolated from fermentation broths in the early 1900s. Penicillin G is the product of a mold, and the cephalosporin beta-lactam antibacterial drugs are also microbial products. Cephalosporin C was isolated from a fungus, and its structure was modified for improved stability and a broader spectrum of activity to afford compounds such as the third-generation cephalosporin cefotaxime (Figure 2.4). Doxorubicin, a metabolite from a soil fungus, has been used for many years as an anti-cancer agent but is cardiotoxic. The synthetic analog mitoxantrone is slightly less potent but has significantly lower cardiotoxicity. A more recent example is that of the epothilones, a class of compounds isolated from a slime bacterium that grows in soil. Several of the natural epothilones have undergone clinical trials, and they have provided the inspiration for semisynthetic analogs, such as ixabepilone, an approved cancer treatment. In this case, the only change in structure is from the lactone of epothilone B to the lactam of ixabepilone, which makes the drug less susceptible to degradation.

Higher animals, such as insects, amphibians, and even mammals, can also be sources of drugs. Magainins and cecropins are antimicrobial peptides from frogs and

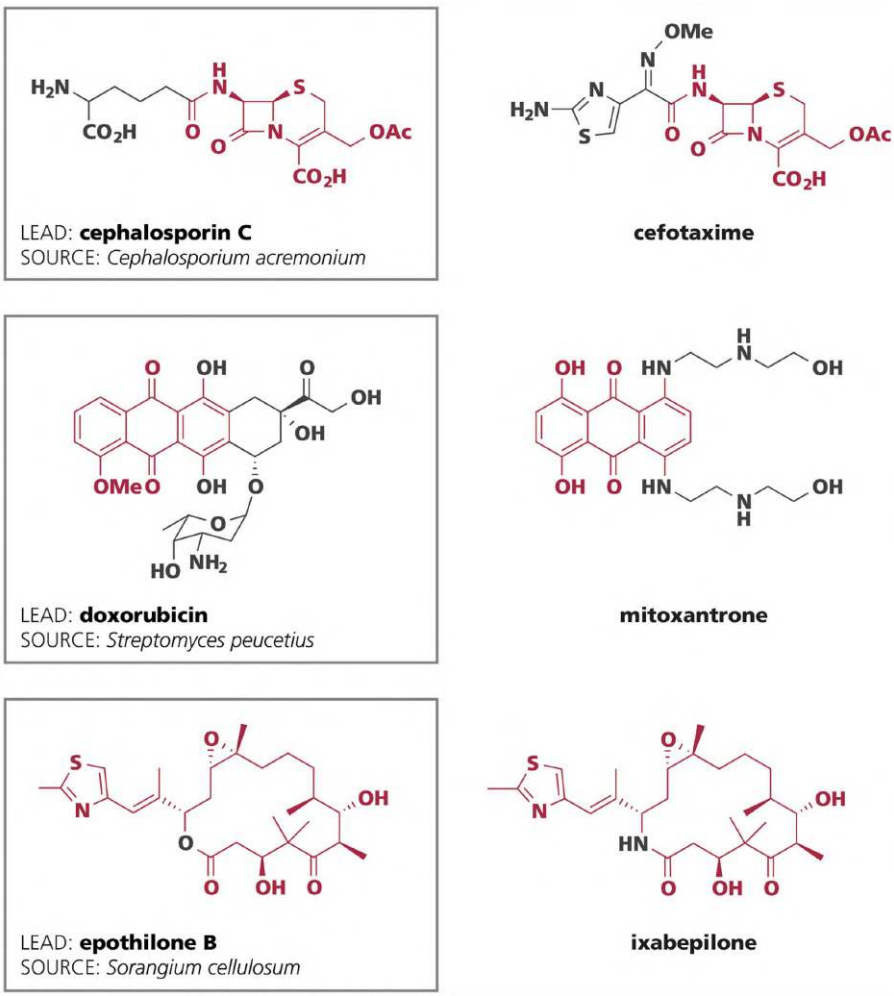


Figure 2.4 Examples of pharmacologically active natural products from microbial sources and their synthetic or semisynthetic derivatives. Cephalosporin C was the lead for cefotaxime, which has improved stability and a broader spectrum of activity. Doxorubicin, an approved anti-cancer agent, led to mitoxantrone, a compound that exhibits much less cardiotoxicity. Epothilone B led to ixabepilone, a more stable analog, which is the only approved agent in this class. Areas highlighted in red indicate common structural features of the original natural product and derived analogs.

silk moths, respectively. Although there have been no drugs derived from these peptides to date, they have been used as structural starting points for potential new antibacterial compounds. Ziconotide, the ω-conotoxin peptide found in cone snails, is an example of a peptide from an animal source that is in use as an analgesic. Finally, Premarin, a mixture of conjugated estrogens that is prescribed for symptoms of menopause, is isolated from horses.

Natural products will continue to be a source of hits and leads. In many of the historical examples described, the natural product already exhibited many of the properties required of a lead or even a clinical candidate and entered the process at a relatively late stage (see **Figure 2.1**). Some of the newer methods for natural product discovery allow a much earlier entry. In these cases, liabilities uncovered during hit-to-lead or lead optimization activities can be addressed with the synthesis of improved analogs that are inspired by natural products. Disadvantages of relying on natural products as a source include the limited number of natural products with the desired biological activity, the limited quantities, and the potential challenge of modifying the structure of complex natural products.

2.4.3 Hits and Leads Are Often Based on Natural Ligands

With a greater understanding of physiology came an additional source of hits and leads for drug discovery and less of a reliance on natural products. Once the physiological effects of macromolecules, such as enzymes, receptors, and ion channels, and their relationship to disease became understood, so did the concept and importance of **natural ligands**. These are the physiologically relevant molecules, often peptides or small molecules, that bind to or interact with a macromolecule to initiate a biological response. For drug discovery projects that target the biological response mediated by a macromolecule–natural ligand interaction, the natural ligand is an excellent starting point as a hit or lead. This strategy for drug design uses the structural features of the natural ligand to develop new compounds that bind at the same site. Once bound, the new molecule may either give a similar effect as the natural ligand or inhibit the effect of the natural ligand. The former is desirable when the disease is caused by a deficiency of the ligand, and the latter is desirable when the disorder is caused by overactivity of the enzyme or receptor.

The anti-ulcer drug cimetidine, which blocks the H_2 subtype of histamine receptors, is one example of a drug that was developed by this strategy. The natural ligand for the H_2 receptor is histamine, and binding between histamine and the H_2 receptor stimulates gastric acid secretion. Cimetidine, a synthetic analog of histamine, binds to the H_2 receptor and prevents histamine from binding, which ultimately leads to a reduction in stomach acid secretion (**Figure 2.5**).

In applying this strategy to develop enzyme inhibitors, the natural substrate is the original inspiration, and it is often modified to afford a novel inhibitor that is recognized by the active site of the enzyme but is not acted upon. For example, when a protease enzyme cleaves the amide bond of a peptide substrate, a **peptidomimetic** is designed that mimics the substrate but does not contain a cleavable bond. Human immunodeficiency virus (HIV) protease is an essential enzyme unique to the virus, and its inhibition results in the production of non-infectious viral particles. The HIV protease inhibitor saquinavir was designed, in part, on the basis of the Phe-Pro segment of the substrate (the GagPol precursor protein), which is recognized by the enzyme.

The advantage in using a natural ligand or enzyme substrate as the starting point for hits and leads is that the derived structures are typically already active toward the target. In addition, a significant amount may be known about the ligand and how it binds to the receptor or enzyme, allowing rapid progress in drug design. A disadvantage to this approach is that it will typically generate compounds that are similar in structure to the natural ligand, and the derived analogs may not be selective toward one form of macromolecule vs. other forms that utilize the same natural ligand. Obviously, this approach is limited by knowledge of the natural ligand and is not useful when the substrate or ligand is unknown. Today, this approach has expanded beyond enzymes and receptors and is applied to natural ligands that constitute one of the partners in protein–protein interactions that initiate some biological signal.

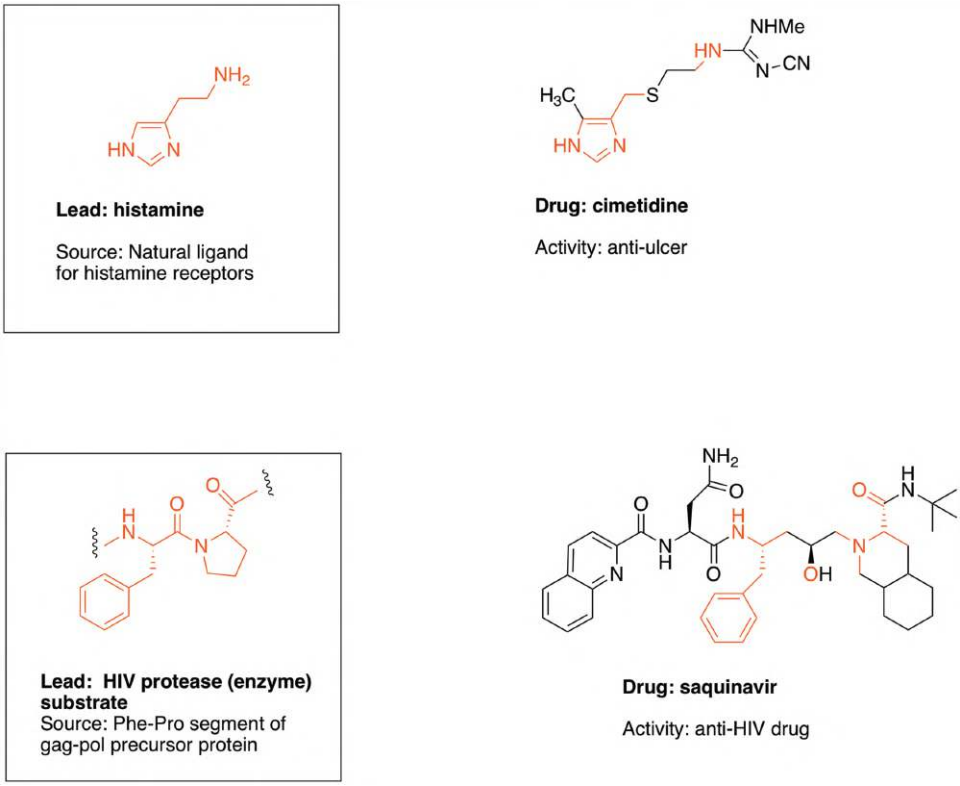


Figure 2.5 Examples of drugs derived from natural ligands, such as neurotransmitters or enzyme substrates. The natural ligand of the H_2 receptor, histamine, led to cimetidine, an anti-ulcer agent. The Phe-Pro segment of the Gag-Pol substrate of the enzyme HIV protease led to saquinavir. Areas highlighted in red indicate common structural features of the original natural ligand and derived drug.

2.4.4 Compounds with Reported Bioactivity Are Important Sources of Hits and Leads

While natural ligands can be the inspiration for new structures with improved profiles, compounds that have been previously reported in the literature to exhibit activity against a certain target or class of targets are also useful starting points. Similar to the strategy based on natural ligands, this approach relies on previous knowledge of the small molecule's biological activity and often results in compounds that already exhibit activity toward the desired target. In other cases, selectivity for one target can be manipulated into selectivity toward another when starting with a nonselective bioactive compound.

One example of a hit discovered in this fashion is azidothymidine (AZT), which was the first drug approved to treat AIDS (**Figure 2.6**). This nucleoside, where the 3-hydroxy group of natural nucleosides is replaced with an azido group, had been synthesized in the 1960s by Jerome Horwitz at the Michigan Cancer Foundation as a potential cancer drug, but it was ineffective in an animal model. Later, a publication in 1974 reported antiviral activity against a mouse retrovirus, attracting only academic interest, since at that time retroviruses were not known to cause any human diseases. In the 1980s, HIV was identified as a retrovirus, and scientists working with AZT at Burroughs Wellcome tested it for activity at the National Cancer Institute. Other known 3-deoxynucleosides were also the subject of scientific attention, including eventual drugs, such as dideoxyinosine (ddI).

Another example of hit identification based on reports in the literature includes some of the kinase inhibitors used to treat cancer. An early Bcr-Abl tyrosine kinase inhibitor drug, imatinib, was approved in 2001 to treat certain types of cancer. On the basis of an X-ray structure of imatinib bound to the enzyme target, the structure was modified to give nilotinib, which is more potent than imatinib.

The literature is a rich source of hits and leads for drugs, and this emphasizes the importance of published research. In some cases, the compounds may be identified but their value as drugs may not be realized until many years later. AZT and ddI, both originally synthesized in the 1960s as potential anti-cancer agents, are two such

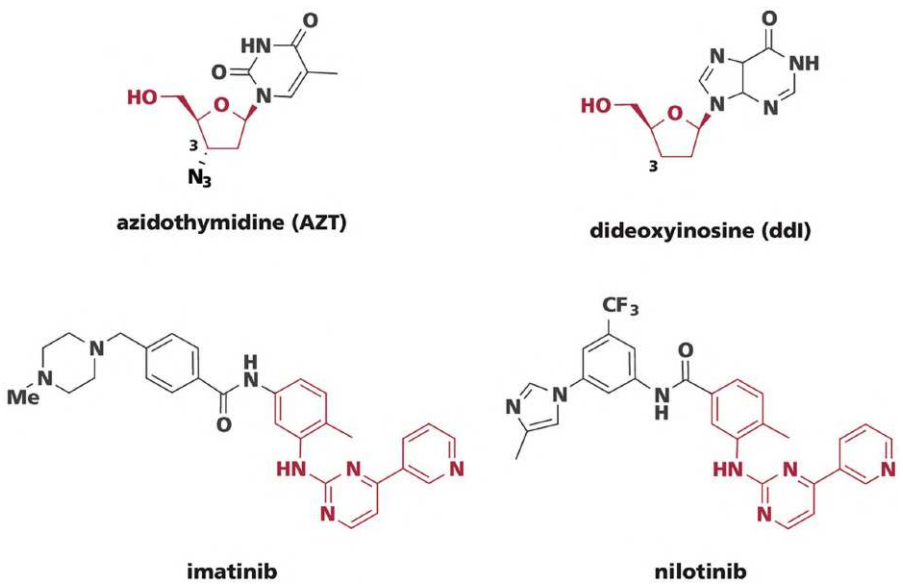


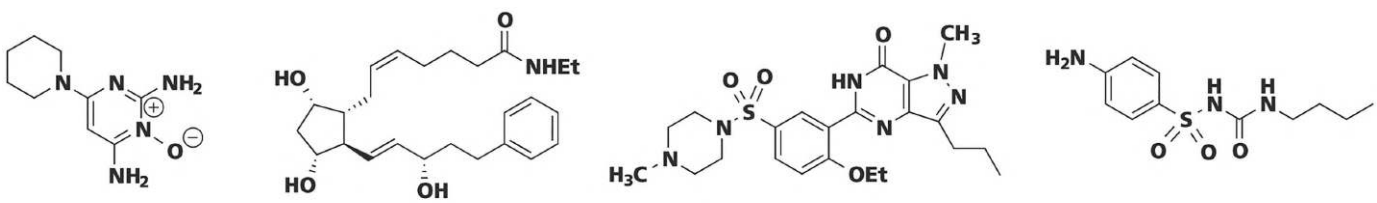
Figure 2.6 Structures of known bioactive compounds that led to new drug approvals. AZT, originally reported as an anti-cancer agent in the 1960s, was approved to treat AIDS in 1987. Other previously reported 3-deoxynucleosides, such as ddI, also became approved AIDS drugs. Imatinib, a kinase inhibitor approved to treat certain cancers, became the lead for nilotinib, another approved anti-cancer agent. Common features are shown in red.

representatives. In other cases, new information in an emerging area may allow one hit or lead to be rapidly modified to generate an improved analog.

2.4.5 Clinical Observations May Uncover Unexpected Activities of Known Bioactive Compounds

While most of the examples discussed previously focus on deriving new compounds with similar activity to that of the original source, serendipity has also played a role in developing new drugs. A number of well-known drugs or drug candidates that were being tested for one indication were found to have unexpected, beneficial side effects that eclipsed their original intended use, with some examples shown in Figure 2.7. Perhaps one of the more widely known drugs that fall into this category is minoxidil. Minoxidil was patented in 1967 by Upjohn as an anti-hypertensive agent and approved in 1979 to treat high blood pressure. Some patients taking it noticed that they had unusually thick hair growth on their arms and backs. The drug was reformulated as a cream that could be applied to the scalp and remarketed as a hair-growth promoter. A recent, similar instance is bimatoprost. It was originally used as eye drops to treat glaucoma but had the side effect of promoting eyelash growth and is now approved for that cosmetic use.

Another important, early example is the anti-diabetic drug carbutamide. This sulfonylurea was originally developed as an analog in the class of sulfonamide antibacterial drugs. During clinical trials, it was noted that carbutamide caused hypoglycemia.



minoxidil

ORIGINAL USE: hypertension (1979)
CURRENT USE: hypertension and hair growth promoter (1988)

bimatoprost

ORIGINAL USE: glaucoma (2003)
CURRENT USE: glaucoma and eyelash growth (2008)

sildenafil

ORIGINAL INTENDED USE: angina
CURRENT USE: erectile dysfunction (1996)

carbutamide

ORIGINAL INTENDED USE: antibacterial
CURRENT USE: anti-diabetic (1950s)

Figure 2.7 Examples of drugs identified by clinical observations. Minoxidil was approved to treat hypertension, and a side effect was hair growth. Bimatoprost was approved to treat glaucoma, and a side effect was eyelash growth. Sildenafil was in clinical trials for treatment of angina, and a side effect was relief of erectile dysfunction. Carbutamide was in clinical trials for antibacterial activity, and a side effect was hypoglycemia.

This observation led to its eventual use to treat diabetes; carbutamide was the first member of the important family of anti-diabetic agents known as sulfonylureas.

Sildenafil (Viagra) is probably the most remarkable example of a drug being developed on the basis of side effects observed during clinical trials. Originally developed by Pfizer as a potential treatment for angina, it has revolutionized the treatment of erectile dysfunction, an unexpected use that was noted during its clinical trials for cardiovascular disease. This success of sildenafil led to the development of numerous other drugs (for example, vardenafil and tadalafil) that work through the same mechanism.

Drugs that are repurposed on the basis of clinical observations often progress through clinical trials rapidly because they have already passed the numerous safety and regulatory hurdles required for administration to humans. Furthermore, if a beneficial side effect is noticed by patients, it is often robust enough to be significant and a true effect rather than an idiosyncratic response. Thus, the path to approval and market may be much shorter than typical. For this reason, "repurposing" approaches are also applied at earlier drug discovery stages. In these cases, known drugs are evaluated for a different indication during the preclinical stage. This was an important strategy during the COVID-19 pandemic, when rapidly identifying efficacious drugs was urgent.

2.5 SOURCES OF HITS AND LEADS IN DRUG DISCOVERY: SYNTHETIC LIBRARIES

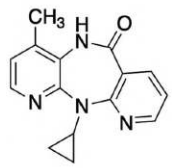
When the basis of a drug discovery project was a natural product, a natural ligand, or a known bioactive compound, the starting point already exhibited some level of desirable biological activity, as well as some of the necessary properties of a clinical candidate. With advances in molecular biology, protein science, and physiology came a greater understanding of the molecular basis of disease. This understanding led to a shift in drug discovery strategies from relying on compounds that already had known bioactivity as starting points to searching for new compounds that exhibited specific biological activity. Today, a common practice is **library screening**, which typically involves testing many compounds simultaneously in order to identify starting points for drug discovery. These compounds are often members of synthetic libraries (but can also be natural products, as described earlier) that number between a few hundred to millions of compounds. Unlike the previous approach, the vast majority of compounds tested will be inactive. For the purposes of this discussion, synthetic libraries will be classified into two general categories: legacy libraries and diversity libraries.

2.5.1 Legacy Libraries Contain Compounds from Prior Discovery Campaigns

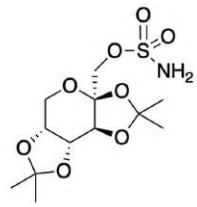
Legacy libraries consist of compounds that have been synthesized in the past, often as part of a medicinal chemistry optimization effort, and retained for testing in the future. These libraries, from both industry and academia, have historically provided some important hits and/or drugs.

The chances of developing a drug are quite small. On average, for every 10,000 compounds that are synthesized as part of a medicinal chemistry or drug discovery effort, only one becomes a drug. Samples are saved and stored for possible future use and are often included in legacy libraries. These libraries can be particularly useful when molecules for one target are synthesized and retained and then later are tested against a related target. This strategy was applied by some pharmaceutical corporations to rapidly identify inhibitors of HIV protease, a member of the family of aspartic proteases. Many companies had previously generated compounds to target another aspartic protease called renin that is involved in blood pressure regulation. Most of these compounds did not advance and became part of a legacy library. When the HIV protease was characterized as an aspartic protease, screening compounds originally synthesized to inhibit the enzyme renin led to potent hits very quickly. Compounds from legacy libraries that target a specific protein class are often used as a focused library set.

While there were reasonable expectations for activity in the renin/HIV protease example, in some cases, activity is observed for a compound that would not ordinarily



nevirapine
inactive as muscarinic antagonist (1960s)
Active: anti-HIV drug (1980s)



topiramate
synthesized as synthetic intermediate (1979)
active: epilepsy drug (1996)

Figure 2.8 Examples of hits obtained from legacy libraries. Nevirapine is an anti-HIV agent that was originally prepared as a muscarinic antagonist. Topiramate, a synthetic intermediate, became a library sample and was eventually approved to treat epilepsy.

have been expected to have that particular type of activity. Nevirapine is a compound that chemists at Boehringer Ingelheim originally prepared as a potential blocker of muscarinic receptors and was included in their legacy library. Although it was inactive in that assay, several years later, it was found to be active in an HIV reverse transcriptase assay and to have anti-HIV activity. This compound was the first nonnucleoside reverse transcriptase inhibitor approved. Another example is topiramate, which was originally prepared at McNeil Laboratories as a synthetic intermediate. A sample was added to the company library, and it was screened for various types of activity. It was selected for screening in a phenotypic assay for seizures, where it exhibited anticonvulsant activity. Topiramate was ultimately approved as an anti-epileptic and antimigraine drug almost 20 years after its original synthesis (Figure 2.8).

2.5.2 Diversity Libraries Are Large Sets of Compounds That Sample Broad Chemical Space

In addition to legacy libraries, **diversity libraries** are also frequently used as a source of hits and leads. Several recent analyses have indicated the enormity of the chemical universe of small molecules that could be used as starting points for drug discovery efforts. By some estimates, limiting the number of heavy (that is, non-hydrogen) atoms to 11 and specifically C, N, O, and F, the number of potential druglike compounds is over 110 million when stereoisomers are considered. Removing the limitation on atom numbers and atom types, the number of potential druglike molecules soars to 10^{60}. However, the number of compounds that have been synthesized and are available for testing is far smaller. Therefore, efforts have been, and continue to be, directed toward the synthesis of unique compounds that sample new, previously unexplored chemical diversity space to address new biological targets vs. the historical targets that legacy libraries represented.

Diversity libraries began to be intentionally constructed in the early 1990s relying on combinatorial chemistry techniques, typically to prepare thousands to tens of thousands of compounds within a specific library. With the advent of new techniques to prepare **combinatorial libraries** (combinatorial synthesis) in the 1990s, numerous libraries were produced by pharmaceutical and biotechnology firms, academic groups, and companies whose main purpose was the synthesis of libraries. These libraries were often based on specific chemistries and were designed such that the resultant molecules, often with the same scaffold containing diverse substituents, were novel and predicted to exhibit druglike properties.

These combinatorial libraries are characterized by the rapid synthesis of large numbers of compounds. Note that the term *combinatorial library* has come to have many meanings. For the purposes of this discussion, it is defined as a relatively large library (>100 compounds, typically thousands) of structurally related compounds, regardless of the method of synthesis. This strategy has been recently adapted to prepare libraries of millions, and even billions, of compounds by **DNA encoding**. Another strategy that emerged in the early 2000s was the synthesis of libraries based on **diversity-oriented synthesis** (DOS), focusing on diversity of the scaffolds from one common intermediate and efficiency of synthesis.

A particular application of such libraries is the preparation of "**focused libraries**." In these cases, large libraries are synthesized that are predicted to be active against a

specific target class, for example, certain enzyme classes like kinases (see Case Study 2). The libraries are designed to contain a core structure with known effects but contains diverse substituents on the periphery. These libraries are particularly useful for the identification of hits for a new target that is a member of, or has similarity to, a well-known class for which a number of active compounds have already been identified.

2.5.3 Combinatorial Libraries Are Sets of Closely Related Analogs That Are Prepared by Parallel or Split-Pool Synthesis

Combinatorial chemistry was initially designed for the synthesis of peptides in the mid-1980s and was applied to traditional small molecules in the early 1990s. It is the systematic, repetitive, covalent connection of a set of different building blocks to each other to yield a large array of diverse substituents around a common scaffold. In each individual step, the bond formed is the same, but the specific substituents will differ. The number of compounds generated depends on the number of building blocks used and the number of reaction steps. In a tripeptide example, using 20 amino acids in each coupling step would generate 20^3, or 8,000, unique tripeptides. This tactic leads to large chemical libraries that have been synthesized on a variety of materials, such as resins or silica chips (solid phases), or in a solution. In a way, combinatorial chemistry mimics nature in the use of these building blocks. Peptides, DNA, and oligosaccharides are all prepared in nature from the building blocks of amino acids, nucleotides, and monosaccharides, respectively. Indeed, most of the early libraries used amino acids as building blocks, although the chemistry rapidly became more sophisticated.

To produce combinatorial libraries, a variety of strategies and technologies can be employed. The final products can be prepared either as a mixture or as isolated individual products; solid-phase, fluorous-phase, or solution-phase synthetic methods can be used; and anywhere from tens to millions of compounds can be prepared at one time. When a solid-phase approach is used, the initial building blocks are linked to a resin. The linked resin is then reacted with sets of building blocks successively until the reaction sequence is completed. Cleavage from the resin affords the final products. The two tactics commonly used in combinatorial synthesis are split-pool synthesis and parallel synthesis. Descriptions of each, making use of solid-phase technology, are given in what follows.

In a **split-pool synthesis**, compounds are typically assembled on beads of a resin by a repetitive sequence of combine, divide, and couple. This process is illustrated in Figure 2.9.

Three pots containing resin, represented by spheres, are coupled to building blocks A, B, and C in three separate reactions to give resin-A, resin-B, and resin-C as starting points. These are combined (pooled) to give a mixture of the three monomers (first box in Figure 2.9) The mixture is then divided (split) into three equal portions to give three new pots, theoretically containing equimolar amounts of each monomer.

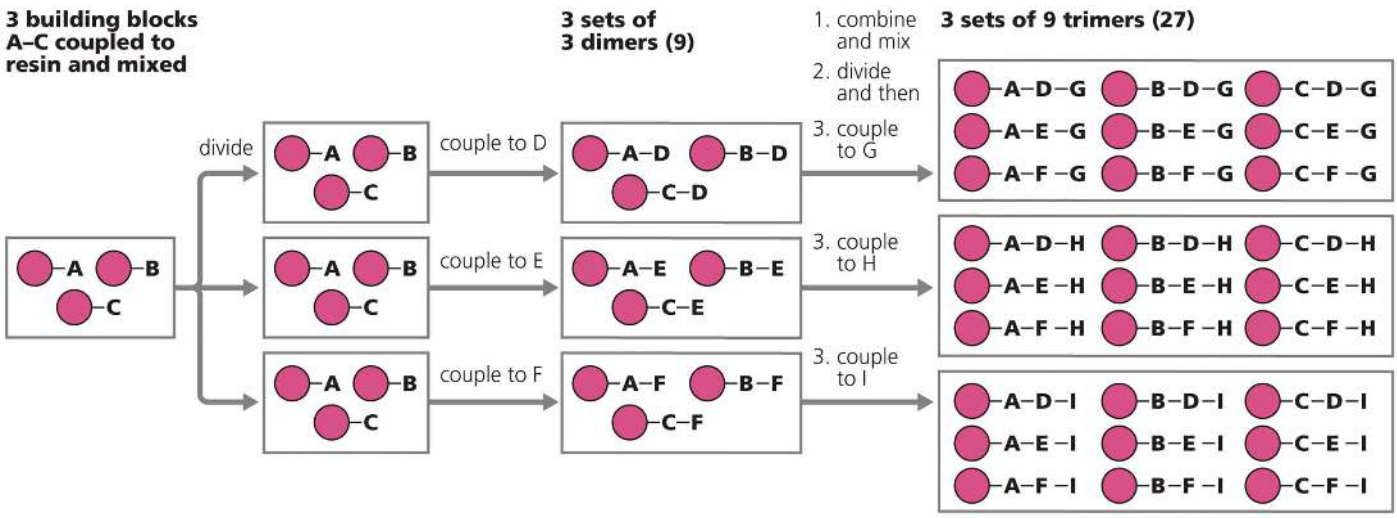


Figure 2.9 Diagram of split-pool combinatorial synthesis on a solid phase. The synthesis begins with three building blocks to prepare 27 compounds. A mixture of the first three building blocks, A–C, linked to a solid support (resin, represented by spheres) is divided, and then each sample is coupled individually to the second set of building blocks, D–F, to give three sets of three (nine) dimers. These three mixtures are combined, divided into three, and each mixture coupled to a third set of building blocks G–I to give three sets of nine trimers (27 trimers).

A coupling reaction is then carried out in each pot (a unique reagent for each pot) to add the second set of building blocks, D, E, and F. Each resulting pot contains a mixture of three unique dimers (nine compounds total). These are then combined and divided into three pots, which are coupled to building blocks G, H, and I to give three pots, each containing nine unique trimers, for a total of 27 chemically unique trimers. The process is continued for as long as desired to generate mixtures of final products.

In any split-pool synthesis, the final mixtures can be screened for biological activity. If they were prepared on a solid phase, compounds could be assayed either while still attached to a resin or after removal from the solid support. **Deconvolution** of the library, or isolation and identification of the active component, is not trivial, and a number of methods have been developed to facilitate this process. These include co-synthesis of either a DNA sequence or a peptide that can be identified by degradation; synthesis and testing of sub-libraries where one component is fixed in each sub-library, allowing for the identification of the active species by a process of elimination; tagging of pools of libraries; or separating individual beads into wells for identification by mass spectrometry and/or NMR.

One powerful technique for library deconvolution is the use of labeling to create **DNA-encoded libraries** (DEL). This technology was first described in 1992 by Sydney Brenner and Richard Lerner, who synthesized a 10^5-member peptide library as a proof of concept. At each synthetic step, a DNA oligomer is co-synthesized and serves as a barcode for the ultimate identification of active hits. During a DNA-encoded synthesis, the following steps occur:

- Each building block is attached to a linker that terminates with a short stretch of a specific DNA sequence.
- Each linker–building block is further extended by a specific short DNA sequence that is a tag for that individual building block.
- These are combined to give a mixture of tagged building blocks and split into groups for the next reaction cycle.
- The second set of building blocks is reacted, followed by attachment of the second specific short DNA sequence, which identifies the second building block.
- This process continues with each reactant or reaction being tagged with a unique DNA sequence.
- At the end of all steps, the compounds attached to the DNA tags are assayed.
- Any hit compounds, as well as their DNA tags, are cleaved from the linker.
- The unique DNA tag sequences are amplified by polymerase chain reaction (PCR) and decoded to reveal the identity of the component of each synthetic step.
- Finally, since the original hits are identified while bound to the DNA, the result is confirmed by testing the isolated small molecule "off-DNA."

An example of a split-pool synthesis of a DNA-encoded library is shown in **Figure 2.10**. Scientists at GlaxoSmithKline synthesized a focused library of over 4 billion triazines in a search for compounds that inhibit the enzyme aggrecanase, which may be useful to treat osteoarthritis. Four synthetic steps, plus those to attach a DNA oligomer code at each step, were carried out to provide the triazines, which were diversified at four positions. From this extensive library, several hits were identified. One hit example, shown in **Figure 2.10**, was modified to give a lead compound.

In a **parallel synthesis**, deconvolution is not necessary since each compound is prepared separately ("spatially addressable"), rather than as part of a mixture. In this type of combinatorial synthesis, each reaction is run on an individual compound (rather than one reaction being run on multiple compounds), and compounds are synthesized in separate vessels, on various phases, or in a solution. For example, a 96-well microtiter plate can be used, with different reagents added to each well. **Figure 2.11** shows a schematic of a spatially addressable parallel synthesis using solid support chemistry on a 4 × 4 plate. Four different building blocks (A–D) are linked to a resin (red sphere) and each placed in separate columns of wells of a microtiter plate.

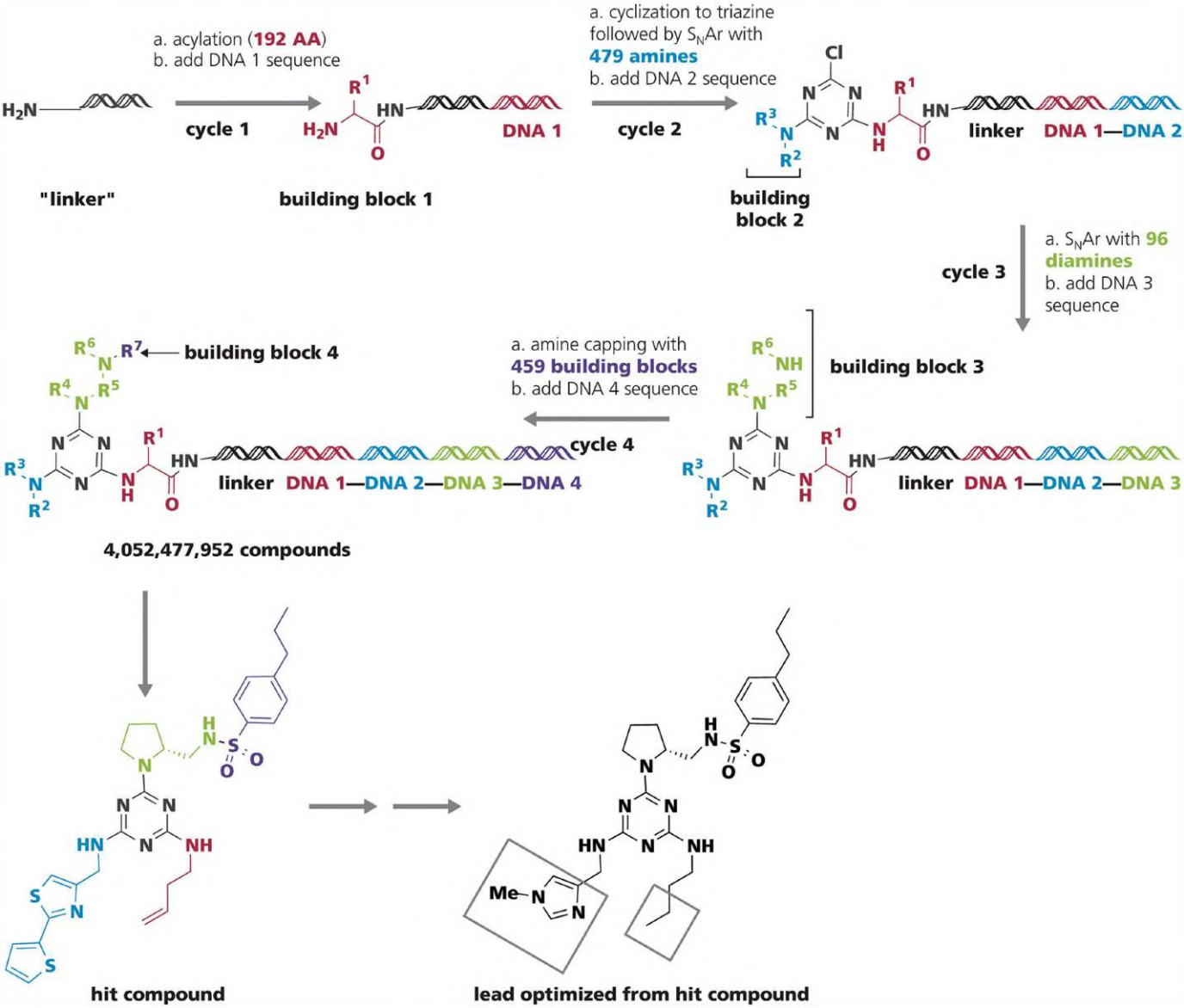


Figure 2.10 **DNA-encoded library synthesis of 4 billion triazines.** Each synthetic step is one color-coded cycle. Cycle 1: Amide linkage of 192 amino acid building block 1 (red) to linker (black) and attachment of DNA sequence 1 (red). Cycle 2: Cyclization of core triazine (common to all in library), followed by aromatic substitution with 479 amine building block 2 (blue) and attachment of DNA sequence 2 (blue). Cycle 3: Aromatic substitution with 96 diamine building block 3 (green) and attachment of DNA sequence 3 (green). Cycle 4: Amine capping with 459 building block 4 (purple) and attachment of DNA sequence 4 (purple). Cleavage from linker gives hit compound and its corresponding DNA sequence, which is used to determine its structure; a lead compound was generated from the hit by modification of two structural elements (in boxes).

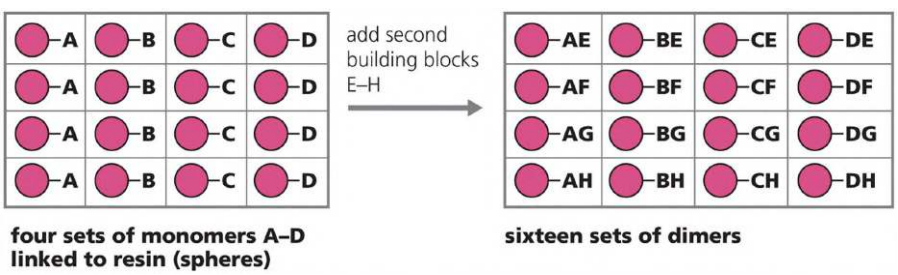


Figure 2.11 **Diagram of a parallel combinatorial synthesis.** Spheres represent resin linked to four different building blocks, A–D. Each column in a 4 × 4 plate contains one specific building block. Reaction with a second building block (E–H), one to each row, gives 16 compounds, each in a separate location.

A second set of four building blocks (E–H) is then added, a separate building block for each row of wells, to give 16 new compounds. Each compound can be identified by its location on the plate.

Parallel synthesis requires more discrete reactions to provide the same number of compounds than a split-pool synthesis, but it can provide larger quantities of each, and compounds are prepared individually rather than as mixtures. Although libraries prepared this way typically produce fewer compounds than those from a split-pool synthesis, a major advantage is the fact that the identity of the active compound is known immediately from its specific location.

In practice, various platforms for parallel synthesis have been applied and developed. For example, in Mario Geysen's multipin technology, the first building blocks are attached to polyethylene pins that fit into the wells of a microtiter plate, where the second coupling steps are carried out. Another early technique was Richard Houghten's tea bag method, where the initially linked building blocks are sealed in polyethylene mesh bags and labelled. They can be combined into one container for deprotection and then separated out into separate containers for coupling to the next building blocks.

The use of large synthetic combinatorial libraries became widespread beginning in the 1990s. While traditional combinatorial libraries provide a large number of compounds for testing in a relatively short period of time, they are, by design, focused on the synthesis of one molecular scaffold. A newer strategy involves the synthesis of diverse molecules, rather than those aimed at a specific biological target or a particular molecular scaffold, with the goal of providing unique structural templates.

2.6 METHODS FOR SCREENING COMPOUND LIBRARIES

The traditional approach for assaying compounds involved evaluating each compound individually, usually in a phenotypic assay, which was a time-consuming process. Advances in protein biochemistry, molecular biology, imaging, miniaturization, robotics, and computer science have led to capabilities for rapid screening. Furthermore, the synthesis and availability of large diversity libraries have led to strategies for screening large numbers of compounds so that many compounds are assayed nearly simultaneously. These capabilities allowed a shift away from strategies that relied on natural products, natural ligands, known bioactive compounds, and serendipity as starting points to high-throughput screening of large diversity libraries.

Screening, testing or assaying a specific collection (library) of small molecules, is currently the most common method to identify hits and leads. In this strategy, several different tactics can be applied depending on the specific project and target. Often, an unbiased set of small molecules (usually up to millions of compounds) is tested in an assay to identify those that have the desired biological effect. Other sets to assay may be natural products, crude plant or microbial extracts, or even virtual compounds. This section describes three main categories: high-throughput screening, fragment-based screening, and virtual screening. In all instances, once tested, the data are analyzed, and those compounds with the desired activity at a certain threshold define a set of small molecules that are considered hits.

2.6.1 High-Throughput Screening Assays Large Numbers of Compounds by a Variety of Different Techniques

High-throughput screening (HTS) has become a mainstay of drug discovery efforts and now accounts for about 30% of recent clinical candidates. HTS typically relies on a miniaturized *in vitro* assay and is performed robotically in microtiter plates. Capacity ranges from medium-throughput (tens of thousands of compounds) to high-throughput (hundreds of thousands of compounds) to ultrahigh-throughput (millions of compounds in a few weeks). In most instances a diversity library containing a random assortment of compounds is screened; however, when there is considerable knowledge about a specific biological target, a **focused screening** approach can be taken. In this case, the compounds assayed are part of a focused library and are not random but,

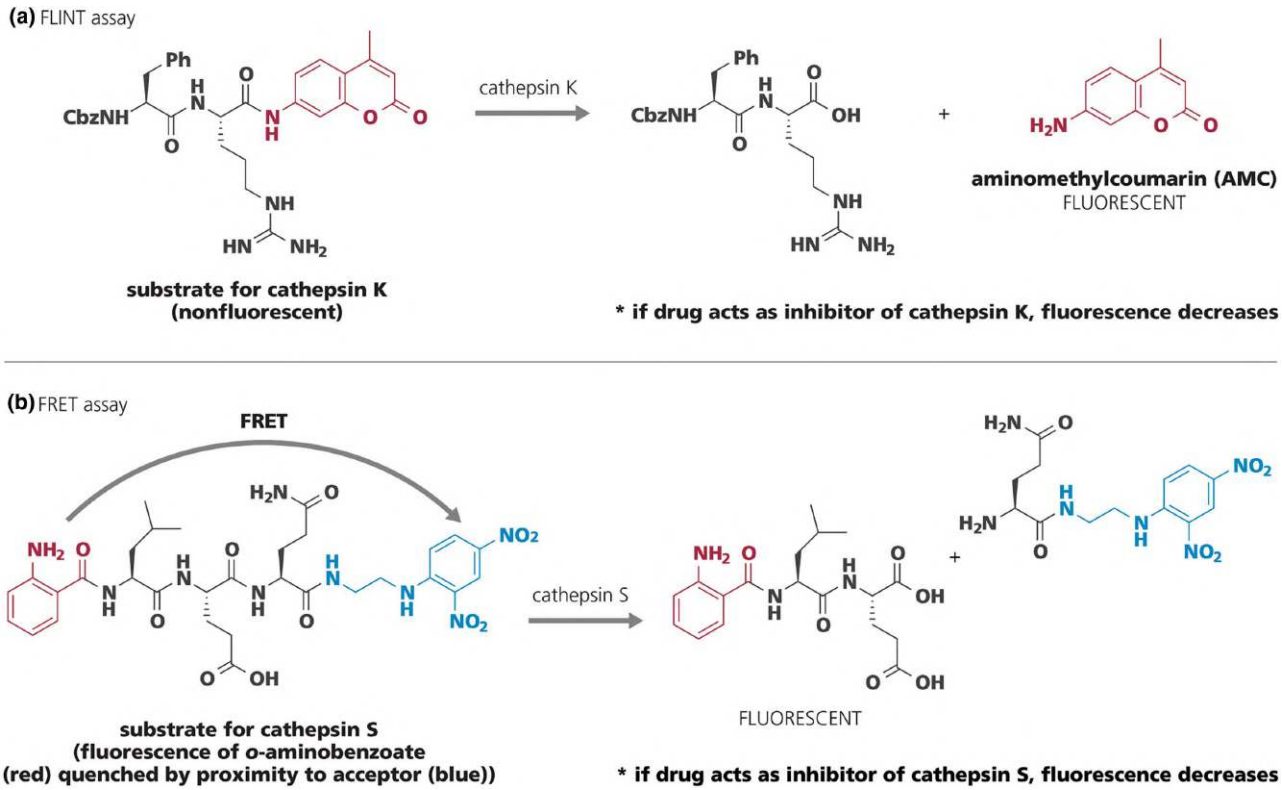


Figure 2.12 Two examples of assays that monitor the production of a fluorescent product: (a) In a FLINT assay, the non-fluorescent substrate is cleaved by the enzyme cathepsin K to generate two products, one of which (AMC, shown in red) is highly fluorescent. (b) In a FRET assay, the fluorescent donor portion of the substrate (red) is in close proximity to the acceptor portion (blue), and no signal is observed due to FRET. Cleavage by the enzyme generates two products that are no longer in proximity, and therefore, the fluorescence of the donor is not quenched and can be detected.

in fact, are known or predicted to have some activity against the biological target or a related target. Often, the number of compounds assayed in a focused screen is smaller than in a random screen.

HTS assays often screen for inhibitors or activators of some defined macromolecular target protein (enzyme or receptor), but they can also test for modulators of entire pathways or identify molecules that correct defects in cells or organisms. Many times the output of an HTS assay is a biological result at one concentration (for example, percent inhibition at a concentration of 10 μM), but sophisticated protocols can also be employed so that HTS outputs generate complete dose–response curves, providing further efficiencies in the process.

All assays require the detection of a signal that provides information about the results of the assay. For example, in an enzymatic or biochemical assay, the product of the reaction must be detected and quantified in some manner. Similarly, in a phenotypic assay, the phenotype must be observed and quantified. Numerous detection methods and technologies have been developed and applied to HTS assay development that are rapid, robust, efficient, cost-effective, and safe. Two examples of common methods used to detect the product of an enzymatic assay in HTS are described in detail. One example is a fluorescence intensity (FLINT) assay, where enzyme-mediated cleavage of a substrate generates a fluorescent molecule that can be detected; inhibition of the enzyme would result in a decreased signal. **Figure 2.12a** illustrates one example of a substrate developed specifically to detect the activity of the enzyme cathepsin K. The substrate itself is not fluorescent, but the cleavage product, aminomethylcoumarin (AMC), is highly fluorescent. Therefore, when the enzyme is active, the substrate is cleaved, AMC is formed, and high fluorescence can be detected; when the enzyme is inhibited, little or no fluorescence is observed. Importantly, the degree of inhibition is directly related to the amount of AMC formed and, therefore, can be related to the amount of fluorescence detected, which, in turn, reflects the potency of the inhibitor.

Similarly, a fluorescence (or Förster) resonance energy transfer (FRET) assay also relies on optical differences between the substrate of the enzymatic reaction and its product. In this type of assay readout, the substrate contains a fluorescent donor and

acceptor. In the example shown in Figure 2.12b, the *o*-aminobenzoic acid moiety (red) in the substrate is fluorescent, and *N*-(2,4-dinitrophenyl)ethylenediamine (blue) is an acceptor moiety. Within the context of the substrate, the donor–acceptor pair is in close proximity. Upon excitation at a certain wavelength, the acceptor captures the emission energy from the donor, and no signal is observed. When the substrate is cleaved by the action of the enzyme, in this case cathepsin S, the acceptor and donor moieties are no longer in the same molecule. The distance between them increases, the efficiency of the capture decreases, and therefore, a strong fluorescence signal is detected.

Other technologies that are often used include assays that detect changes in proximity of two molecules; assays that detect specific molecules, such as products of enzymatic reactions, or molecules or ions produced in response to receptor binding; and assays that detect physical changes (for example, mass or optical properties) observed upon binding of a small molecule to a macromolecule. The overall process for HTS, using an example of a fluorescence assay, is outlined in Figure 2.13.

Even though HTS can be used to screen tens of thousands, or even millions, of compounds, most often as individual molecules, but sometimes as mixtures, the false positive rate can be as high as (or even higher than) 50%. This lack of reproducibility is due to many factors, including poor quality of compounds in the screening library, lack of assay reproducibility, and assay interference mechanisms, among many others. For this reason, HTS campaigns require a significant post-HTS effort aimed at confirming active compounds (discussed later in this chapter). The advantages of the strategy of screening synthetic libraries include the fact that the small molecules identified may be distinct from the natural ligand and synthetically accessible, in contrast to natural products, whose availability may be limited.

2.6.2 Fragment-Based Screening Detects Direct Binding to a Protein Target

High-throughput screening based on fluorescence or related detection methods has the advantage of rapidly screening large numbers of compounds, typically of molecular weight between 300 and 600, and detecting those that are highly active (1 μM–1 nM). However, two significant disadvantages are the large number of false positives identified and the fact that the mechanism of action of the hit, and its mode of binding, may not be known. **Fragment-based screening** (FBS) is based on the concept of identifying smaller fragments (up to about 300 MW) that can bind to a specific protein target even with only modest (millimolar) affinity. By identifying compounds that bind directly to a specific protein target, FBS avoids many of the false positives that are generated in typical HTS approaches. In some instances, the specific binding site on the target protein can even be localized. Once hits are identified, however, they are of low potency. In order to optimize them, when several fragment-sized hits that bind to different sites are identified, they can then be combined; alternatively, single fragments can be grown into hits that occupy multiple binding sites, thereby increasing potency. The general process is outlined in Figure 2.14.

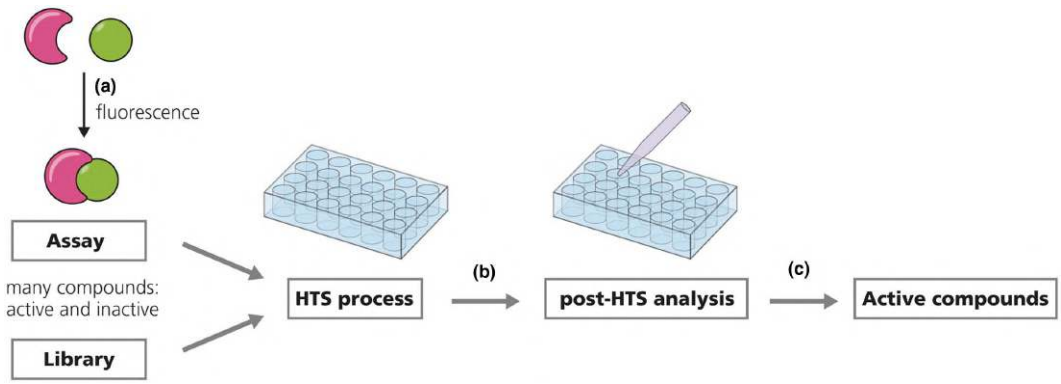


Figure 2.13 Schematic of HTS process based on an assay using fluorescence-based detection: (a) Assay that detects fluorescence is identified for use. (b) Many compounds in a library are assayed in microtiter plates by HTS. (c) Post-HTS analysis identifies hits (active compounds) on the basis of observed fluorescence in individual wells.

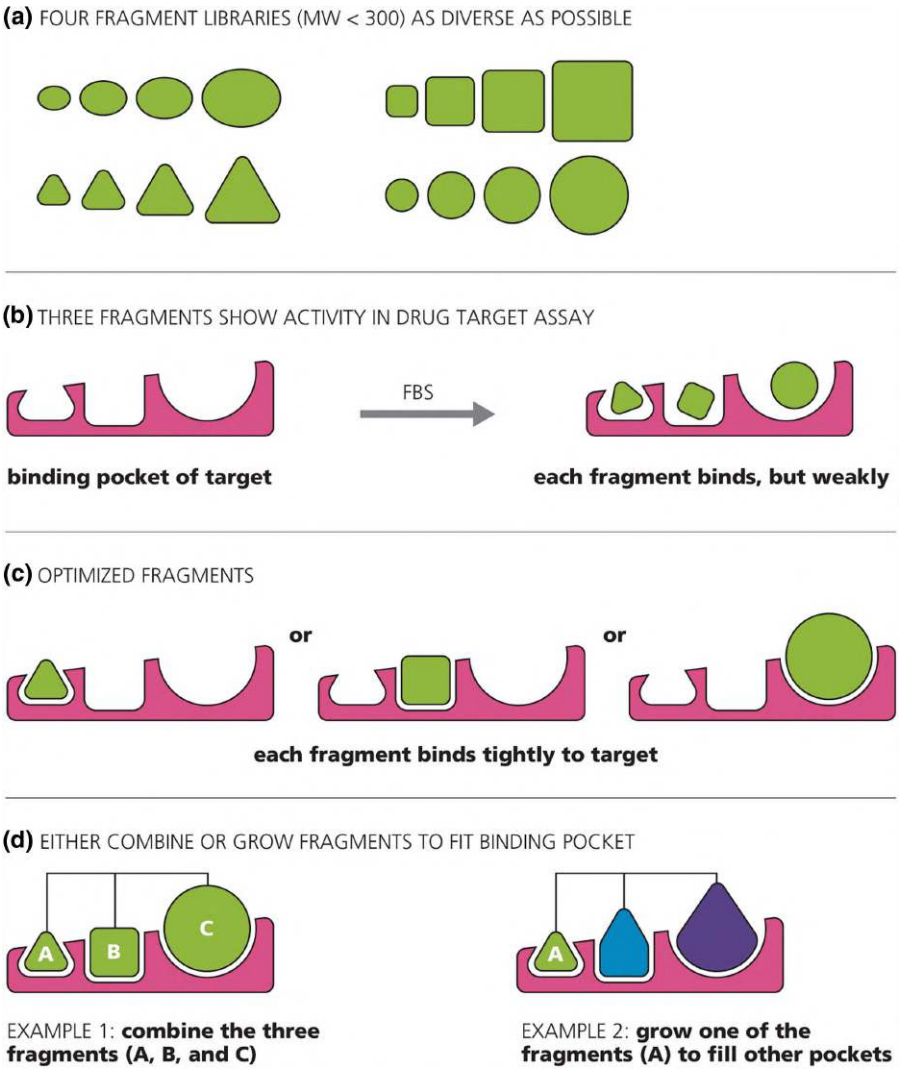


Figure 2.14 Schematic of fragment-based screening and optimization: (a) Four diverse fragment libraries (depicted as ovals, triangles, rectangles, and circles), each containing multiple members (depicted by size differences), are screened against a protein target. (b) Three of the fragments show weak (>1 mM) binding and are considered hits. (c) Each of those three fragments is optimized to bind more tightly to the target. (d) Fragments may be combined or grown to give a lead compound; shown are the three optimized fragments combined (*left*) and one optimized fragment (A) grown (*right*).

One requirement of fragment-based screening is the availability of an isolated enzyme or protein and, often, high-resolution structural data (such as NMR or X-ray) of the target macromolecule. The method to detect binding to specific sites on that target usually involves some biophysical change. Fragment-based screening begins with smaller collections of compounds than HTS, but any hits identified are validated by the knowledge that they bind directly to the target of interest. The process of fragment-based screening was pioneered at Abbott Laboratories in the 1990s using NMR techniques. The idea was based on changes in two-dimensional NMR spectra that are apparent when a compound binds to a macromolecule target. Comparison of spectra in the presence and absence of one or more fragments identifies those that exhibit binding to the target. Other common detection methods include X-ray crystallography, where protein crystals are soaked with various fragments and structures are solved that contain bound fragments, and surface plasmon resonance (SPR), which can detect weak binding to a protein. While fragment-based screening has a number of advantages, there are also some limitations having to do with the conversion of low-potency binding fragments into higher-potency drug hits and leads. In order to combine or grow these low-affinity fragments into high-quality leads, a significant amount of medicinal chemistry work is often required to identify the optimal linker length, position, and structure in order to generate a molecule that binds as expected.

Since the introduction of FBS in the 1990s, this technique (also called FBDD, for fragment-based drug discovery) has been very successful in generating compounds that have advanced to clinical trials. The first drug to reach the market using this

strategy was vemurafenib, an anti-tumor drug that targets a mutant form of the protein kinase B-Raf, found in 60% of all melanomas. Vemurafenib, approved in 2011, was developed at Plexxikon in a program that used X-ray crystallography FBS as part of the optimization process. Although the ultimate target was a specific kinase, B-Raf, the project was initiated with a traditional biochemical screen for inhibitors of five diverse kinases using a library of 20,000 fragment-like (MW = 150–350) compounds to identify those that inhibited at least three of the five kinases. This strategy was based on the idea that compounds that inhibited multiple kinases were likely to bind to the active site of any kinase and also reduced the number of fragments that advanced into the X-ray-driven, fragment-based screen. From that effort, 238 compounds were identified. At the time of screening, conditions for crystallizing the target kinase, B-Raf, were not yet optimized, so initial crystal structures were solved with a similar kinase enzyme called PIM-1. Over 100 co-structures with bound fragments were solved and showed how one fragment, 7-azaindole, bound to the ATP binding site (albeit with modest affinity ~200 μM). Synthesis of substituted 7-azaindoles led to the 3-aminophenyl analog with improved, but still modest, binding affinity (100 μM), and further optimization led to the B-Raf kinase inhibitor vemurafenib. This is an example where an identified fragment was grown into the final drug (**Figure 2.15a**).

Another example of a molecule developed by FBS is that of the matrix metalloprotease-2 and matrix metalloprotease-9 (MMP-2 and MMP-9) inhibitor ABT-518 (**Figure 2.15b**). Scientists at Abbott identified two fragments using NMR that bound to distinct sites on the protein, each with weak binding. One fragment, a biphenyl nitrile (shown in blue), was linked to the other fragment, a hydroxamic acid (shown in

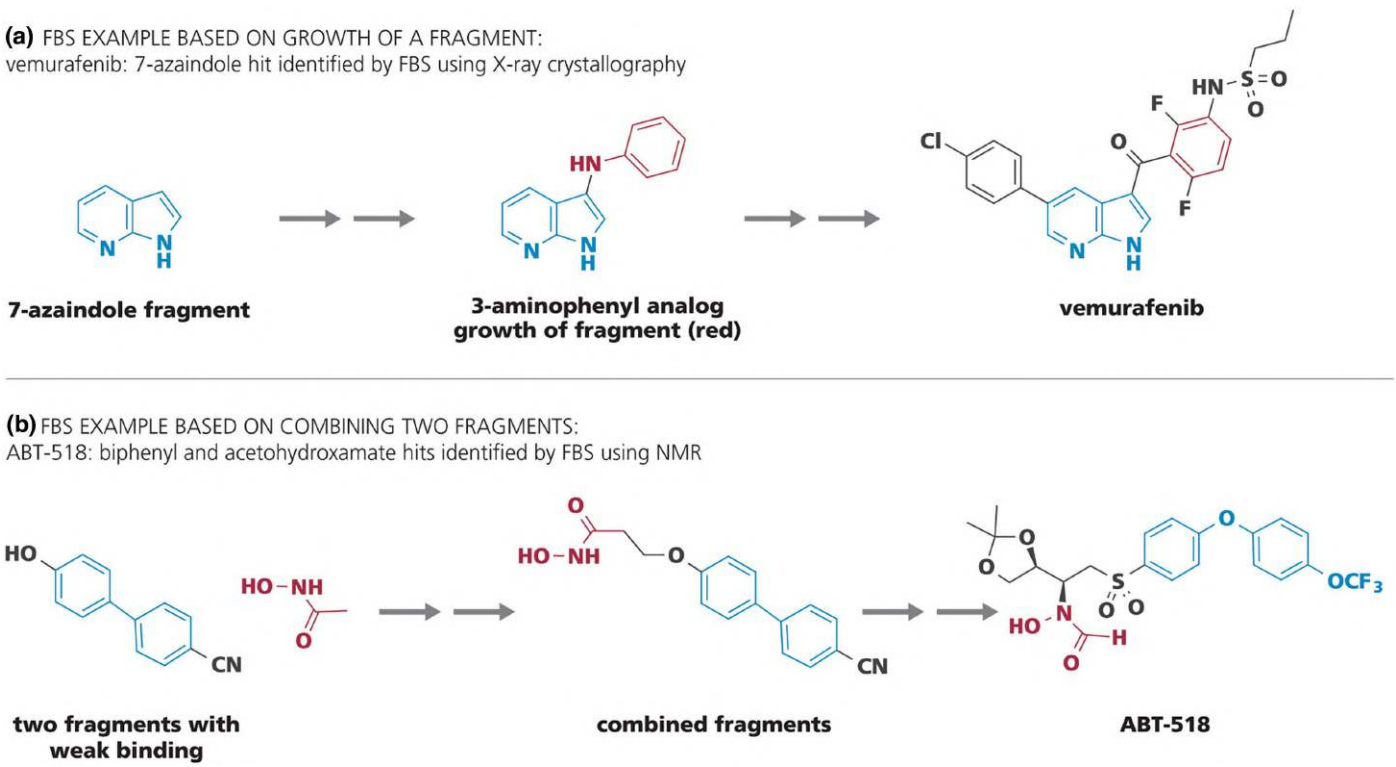


Figure 2.15 Two examples of drugs or clinical candidates based on hits identified by fragment-based screening: (a) 7-Azaindole (blue) was identified by FBS and shown by X-ray crystallography to bind to the ATP site of the kinase PIM-1. The fragment was grown to generate a 3-aminophenyl (shown in red) analog and eventually optimized to generate vemurafenib, a drug approved to treat melanoma. (b) Two fragments, a biphenyl nitrile (blue) and a hydroxamic acid (red), were identified by FBS for inhibitors of MMP-2 and MMP-9. They were combined and further optimized to generate ABT-518, which entered clinical trials as an anti-cancer agent.

red), to provide a compound with improved binding (57 μM) compared to the initial fragments. Further rounds of optimization provided the clinical candidate ABT-518, which entered development as an anti-tumor agent. This is an example where fragments were combined to generate leads that eventually evolved into a drug candidate.

2.6.3 Virtual Libraries Can Be Screened by Computational Algorithms, Including Artificial Intelligence (AI)

Virtual screening (VS), also called *in silico* screening, is defined as assaying by computer and predicting activity, rather than experimentally screening genuine compounds to identify potential hits. The advantage of this approach is that a virtual library of millions of compounds can be screened in a very short period of time at a much lower cost than other methods. However, subsequent testing in the laboratory is always required after the initial virtual screening activities are complete in order to confirm activity. For comparison, screening a library of a million compounds could take several weeks by HTS, using the most sophisticated equipment, but only a few days with VS. Libraries of compounds are generated computationally and can consist of known compounds or even theoretical compounds. Some examples of electronic files of available small-molecule libraries that can be used for VS include ZINC (University of California–San Francisco, UCSF), PubChem (National Center for Biotechnology Information, or NCBI), and DrugBank (University of Alberta and the Metabolomics Innovation Centre), all of which may be accessed for free. Other libraries that also contain virtual compounds include the GDB-17 database, which contains over 100 billion molecules with up to 17 atoms containing the heavy atoms C, N, O, S, and halogens. Compounds screened are often first computationally filtered for those that fit defined requirements for druglike properties and may be further computationally profiled for potential liabilities. These liabilities include the chances of cytochrome P450 interactions (an indication of toxicity) and hERG liabilities (an indication of cardiac toxicity), among many others. Further filtering may incorporate knowledge of the binding site, to eventually generate a selection of virtual hits. Virtual screening can be categorized according to two approaches.

Structure-based virtual screening is an approach that can be applied when a three-dimensional protein structure, co-structures of the protein and a binding partner, or homology models are known. These structures have often been determined by X-ray crystallography, NMR, or other means. A newer approach allows the use of protein structures that have not actually been solved in the laboratory. AlphaFold is an artificial intelligence (AI) system developed by Google's DeepMind that predicts protein structures with relatively high accuracy based on the amino acid sequence. While there are approximately 2,400 solved X-ray crystal structures of drug targets, there are now about 200 million structures predicted by AlphaFold available in a free database. With either type of structure, the process known as **docking** involves computationally fitting sets of compounds into a model of the binding site of a protein. Compounds are first minimized so that only low-energy conformations are represented, and then they are docked into the active site and scored to compute how well each one fits on the basis of shape, size, and key interactions. Those with the best scores receive higher priority. Since the docked compounds are virtual, they are either purchased or synthesized and then must be assayed in the laboratory. The initial VS process can be automated so that millions of molecules can be docked efficiently. The advantages of this method include relatively low cost and rapid screening as compared to other methods. It is particularly useful to identify a small library of compounds when screening of larger libraries is not feasible. The disadvantages are a high false positive rate and the fact that sometimes the virtual compounds may not be readily available.

Ligand-based virtual screening is a strategy that can be used when the structure of the protein target is not available but ligands that bind to that target have been described. The structure of the ligand(s) is described computationally to account for properties like size, polarity, H-bond donor and acceptor capabilities, and hydrophobicity. A model of the ligand, based on these descriptors, is then compared to the same descriptors of each member of a very large library of compounds in order to identify compounds that mimic those of the ligand. The source

of the compound libraries from which the descriptors are generated may be real (actual compounds that have been previously synthesized), such as those found in ZINC, or they may be simulated (never previously synthesized), such as many of those generated in GDB17. Hits are compounds that have similar properties to the known ligands but have a unique structure.

The success of this strategy, and the quality of the hits generated according to ligand-based VS methods, depends on the level of input. The simplest level would be one-dimensional descriptors of the small molecule, such as molecular weight or lipophilicity, where no structural information is required; however, this is rarely used. A more common approach involves a two-dimensional model of a structure or substructure of compounds with known activity. Structures may be searched for descriptors such as number of aromatic bonds or the presence of functional groups, such as alcohols or amines. Alternatively, when a limited number of compounds with known activity are available, **molecular fingerprints** may be used. These are a series of binary digits (bits) that are used to encode the structural features of a molecule, such as aromatic rings or other functional groups. In this type of search, a fingerprint is calculated for the template molecule and then used to search virtual libraries for compounds with a similar fingerprint, with the goal of identifying hits that have features in common with the ligand model but have unique scaffolds holding those features in place. Fingerprints may be calculated by computational methods that identify certain structural features, such as particular atoms or rings. A value called the **Tanimoto coefficient** is calculated to measure the similarity between compounds. Highly similar compounds have a Tanimoto coefficient close to 1; dissimilar compounds have a Tanimoto coefficient close to 0. This type of analysis can lead to the identification of structures with different backbones but with similar features, such as functional groups, and potentially similar biological activity.

Finally, three-dimensional ligand-based virtual screening incorporates a three-dimensional model of the compounds that puts specific functional groups or properties, such as a hydrogen bond donor or lipophilic group, in a specific orientation with respect to each other. This model is referred to as a pharmacophore and will be described in more detail in Chapter 3. Three-dimensional conformers of all library members are generated and compared to the three-dimensional model of the ligand. This type of effort requires more sophisticated calculations. If structural information on the target is available, a final step could be the introduction of the target enzyme or receptor and the docking of the three-dimensional conformers that have been identified.

One specific example of structure-based virtual screening, shown in Figure 2.16, illustrates this process. In this case, virtual screening was used to identify novel activators of peroxisome proliferator–activated receptors (PPARs), which are targets for disorders involving glucose and lipid metabolism. The screening began with several

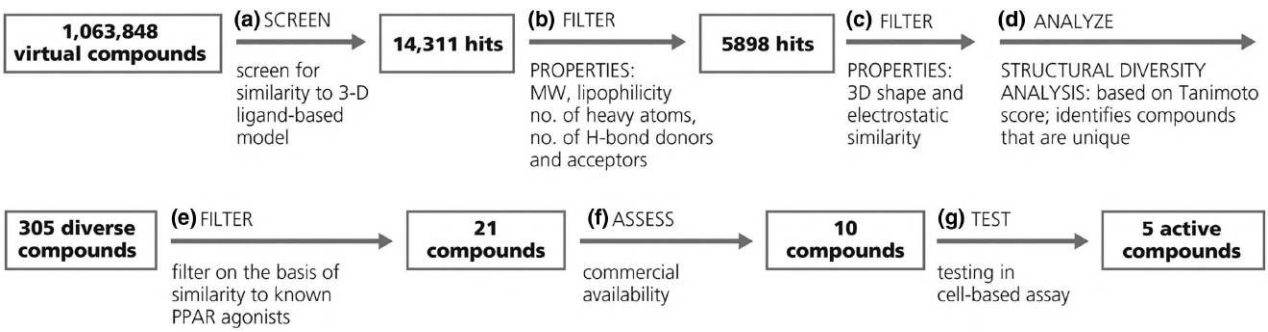


Figure 2.16 Virtual screening example used to identify five active compounds from a virtual library of >1 million compounds. Steps: (a) Screen for similarity to three-dimensional ligand-based model. (b) Filter for specific chemical properties. (c) Filter for similarity to compounds known to bind to target (those hits were kept); eliminate chemical labile compounds. (d) Filter for diverse structural classes (remove those too similar to known active compounds). (e) Filter to remove compounds similar to known PPAR-binding compounds. (f) Identify commercially available compounds. (g) Test in cell-based assay.

X-ray crystal structures of ligand–PPAR complexes found in the Protein Data Bank (PDB). Computational programs were used to derive three-dimensional models based on the bound conformation of the small molecules. Over 1 million (1,063,848) virtual compounds were compared to the known activators, which led to the identification of 14,311 virtual hits. The hits were filtered to remove those that had unacceptable physicochemical properties, and 5,898 virtual compounds remained.

The 5,898 hit molecules were screened further for three-dimensional shape similarity and electrostatic similarity to compounds known to bind to the PPAR target but distinct from those used in the first step, and the compounds that were most similar in shape and electronic characteristics were prioritized for further advancement. By use of various software programs, those compounds that contained chemically labile groups were eliminated. To make sure the final compound set represented a wide range of structures, a computational diversity analysis was performed, which identified 305 unique compounds. Final refinement of those 305 compounds included the removal of compounds that were highly similar in structure (two-dimensional similarity) to known PPAR-binding compounds, as well as those with reported PPAR activity. Of the 21 remaining compounds, 10 were purchased and tested, and 5 new structures were active in an assay for the activation of PPAR activity.

A relatively new approach to hit identification uses **artificial intelligence (AI) and machine learning (ML)**. Artificial intelligence is software that can perform human-like tasks, such as reasoning and learning. Machine learning is a subset of AI involving algorithms that allow computers to learn to perform tasks, identify relationships, and find patterns without the need for human instructions. These techniques are used in a number of areas of drug discovery, including drug synthesis, identification of potential drug candidates, and prediction of pharmacokinetic properties, and here we focus on its application to hit identification. AI algorithms can be applied to virtual screening, both target-based and ligand-based, by relying on large sets of data (e.g., protein structure, data sets of compounds and their bioactivity, among others) as *input*, which are used to train a model to identify relationships and generate *output* (e.g., small-molecule starting points and their predicted bioactivity). Examples of model types include neural networks, support vector machines, and random forest. This process can also mine medical literature and symptom data for drug repurposing. An alternate to similarity searching is the use of machine learning tools, such as deep neural networks (DNN), or proteochemometric (PCM) models that use both ligand and target structure. Prediction of physical properties and toxicity is also feasible, based on the availability of large data sets of these properties.

AI/ML and other computational methods can also be applied to **target fishing**, a strategy to identify an unknown biological target of a molecule. In these cases, instead of docking multiple small molecules into one target, a small molecule that exhibits some phenotypic effect is docked into multiple targets to identify a small molecule–target pair with high binding energy, leading to drug discovery starting points for the identified target. Ligand-based target fishing relies on public databases, such as ChEMBL, PubChem, and DrugBank, for structural information on bioactive compounds. Structure-based target fishing relies on 3D target structural and binding site data. A given ligand, or pharmacophore model, is docked into multiple targets that are available from protein databases, such as the Protein Data Bank (PDB), or from AlphaFold.

The virtual screening of libraries, as well as HTS and FBS, is now widely used for the identification of hits in modern drug discovery programs. All rely on the availability of a robust assay and the accessibility of high-quality small-molecule libraries (whether virtual or real). The identification of a high-quality hit is arguably one of the most important aspects of a successful drug discovery program.

2.7 HIT TRIAGE PROVIDES HIGH-QUALITY HITS FROM SCREENING

The previous sections focused on sources of compounds and the identification of hits by specific assays and screening methods. However, some of the screening methods described can suffer from high false positive rates. In addition, the number of

hits generated is often high, and for practical purposes, all hits cannot be pursued. Therefore, it is important to identify high-quality hits as early as possible in the hit-to-lead process. These activities are especially important for hits identified from high-throughput screening and are defined as **hit triage**. The hit triage process helps identify validated hits that are most likely to generate high-quality leads and includes the following components:

- Elimination of false positives
- Elimination of hits that are active via undesirable mechanisms
- Evaluation of physical and pharmaceutical properties (e.g., solubility, permeability, toxicity) to identify those starting points most likely to generate high-quality leads

2.7.1 False Positives and Nuisance Compounds Need to Be Eliminated from Primary Hit Lists

With the large number of compounds that are assayed by HTS and identified as active, there are inevitably some false positives and compounds that are undesirable. *False positives* include compounds that interfere with the assay detection method or that react with and/or deplete assay components. In addition, many compounds are indeed active against the target but work through undesirable mechanisms, such as the formation of aggregates, redox activity, or promiscuous covalent reaction. It is also not uncommon that small impurities in library compounds contribute to biological effects. One of the first priorities during the hit triage process is to identify and eliminate these compounds.

False positives are typically eliminated by the incorporation of control assays, independent re-assay of purified compound samples, and other protocols specific for the particular HTS. Active compounds that work through undesirable mechanisms are usually active across multiple assays and have come to be called **pan-assay interference compounds** (PAINS). Their promiscuous activity, through a variety of mechanisms, is considered artifactual. While a large number of specific compounds have been identified, a smaller subset of substructures is considered to contain the most frequent offenders. In practice, as a first step in hit triage, active hits identified by screening are often filtered through computational chemistry programs to eliminate these undesirable starting points.

There are a variety of mechanisms by which PAINS compounds exert their promiscuous activity. Some interfere with binding interactions by forming **aggregates**, which can then inhibit many enzymes or bind to proteins nonspecifically. It has been reported that compounds such as these may account for up to 95% of HTS hits, and they are particularly problematic in the screening of plant extracts. These aggregates may be 100–1,000 nm in diameter and form a physical association with the enzyme that may be reversed by the addition of a detergent, such as Triton B. Many compounds in this class exhibit detergent-like activity and denature proteins, causing them to appear to be inactivated. Some examples of common aggregate-forming compounds are shown in **Figure 2.17a** and include phenolic compounds, such as the flavonols quercetin and rhamnetin, as well as polyphenolic tannic acid.

Another frequently found features of PAINS compounds are **reactive functional groups**, like peroxides, sulfonyl halides, primary alkyl halides, dicarbonyl compounds, epoxides, azides, α,β-unsaturated systems, and isocyanates, among others. These groups, and structural motifs such as rhodanines and hydroxyphenylhydrazones, can react nonspecifically with assay components by forming covalent bonds. **Redox-active compounds** are those that can generate free radicals under assay conditions, initiating a redox cycle, thereby appearing to inhibit enzymes and/or proteins. This mechanism is especially prevalent when proteins being assayed contain active site cysteines or require reducing conditions. Examples of redox-active compounds include quinones and catechols (**Figure 2.17b**). Cheminformatics methods are used to eliminate PAINS compounds that not only look for known problematic substructures but also identify if a particular compound is active in many assays (using databases such as PubChem), suggesting that it is promiscuous.

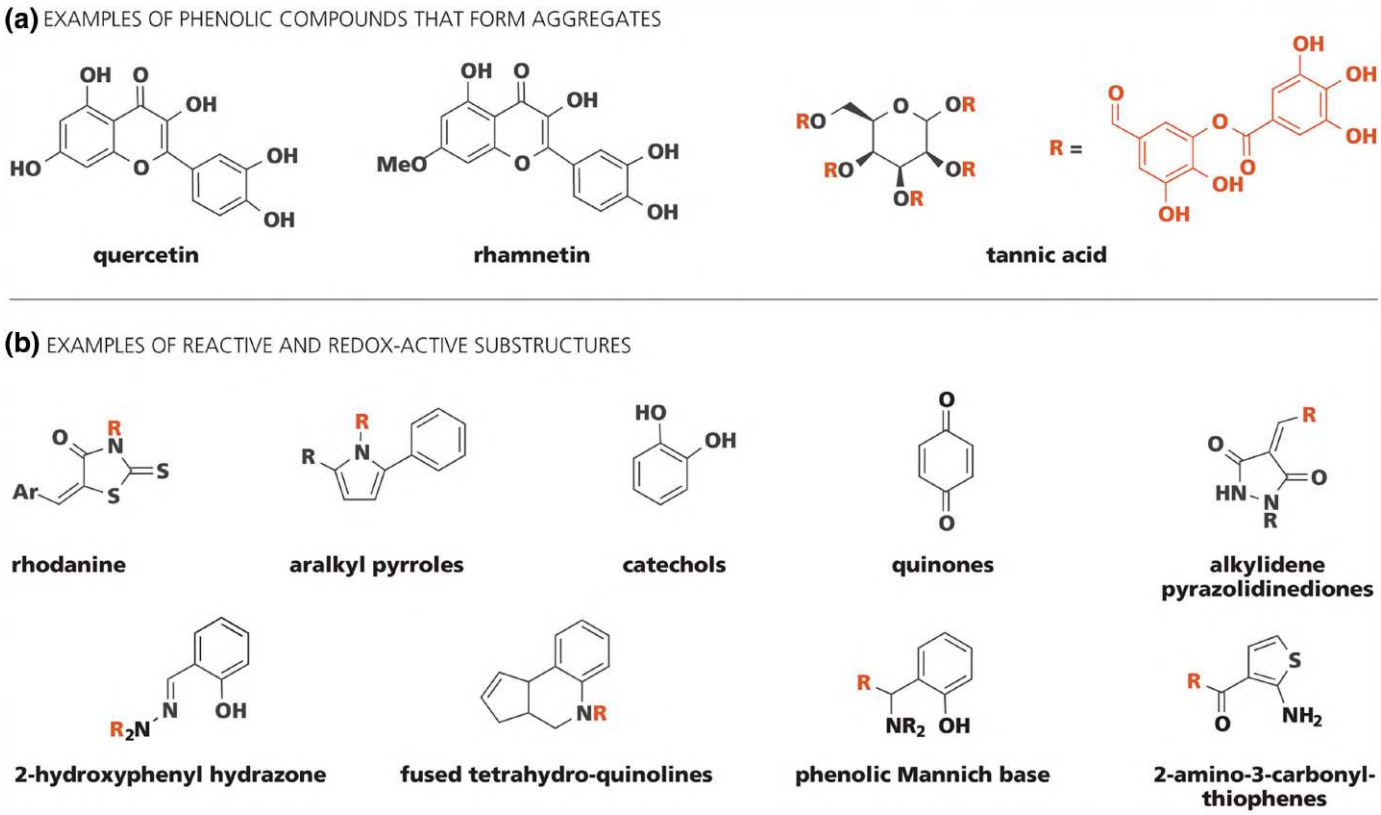


Figure 2.17 Examples of PAINS compounds: (a) Examples of phenolic compounds that form aggregates and cause false positives in HTS assays. (b) Examples of reactive and redox-active substructures identified in HTS libraries.

Triaging of hits must always be done with an awareness of the target class and the specific goals of the project in mind. For example, active site cysteine protease and serine/threonine protease inhibitors require a reactive functional group, although those may be undesirable for other targets. The AIDS drug AZT (also known as zidovudine, or ZDV) contains an azide functional group, and the anti-cancer drug ixabepilone contains an epoxide function; both are potentially reactive functionalities that would be eliminated by some filters. There is also a growing appreciation that careful installation of reactive functionalities in certain drugs can lead to very effective enzyme inhibitors. By some estimates, approximately 25% of known drugs would be eliminated by these filters.

2.7.2 Hits Identified in Screening Campaigns Need to Be Prioritized

Once undesirable compounds are eliminated, the process of hit prioritization takes place. These activities aim to identify the attributes and liabilities of the remaining compounds so that the most promising can be advanced further. Several characteristics are considered to identify druglike hits. The term **druglike** refers to "molecules that contain functional groups and/or have physical properties consistent with the majority of known drugs" (Walters & Murcko). This concept will be described in detail in Chapter 3, but certain parameters, such as molecular weight, lipophilicity, number of rotatable bonds, and number of hydrogen-bonding groups, fall into a specific range when drugs are evaluated in comparison to the larger chemical structure universe. Additional attributes evaluated often include lack of activity against other biological targets, patentability, and synthetic accessibility. Once high-quality hits and/or leads have been identified by any of the methods described in this chapter, there is still much work to be done to convert that molecule to a drug. This process of lead optimization is described in Chapter 3.

2.8 SUMMARY

In this chapter, current practices used in the initial steps of drug discovery, hit and lead identification, have been introduced. Identification requires both an assay and a source of hits. Assays used for hit identification range from biochemical and binding assays for a specific macromolecular target to functional and phenotypic assays. Historically, hits, leads, and drugs were based on previous observations of biological activity, either in humans or in animals. Drugs are still discovered in natural sources, such as plants and microorganisms that have known bioactivity, and from literature reports of bioactive compounds. In some unusual cases, serendipity has played a role, such as the observation of unexpected activity during clinical trials or after a drug is on the market. However, with a greater understanding of physiology and the molecular basis of disease, and with advances in molecular biology, robotics, and computation, modern drug discovery can now begin at a much earlier stage by identifying small molecules that have a defined activity toward a particular macromolecular target or pathway. These small molecules can come from knowledge of the natural ligand and its structural modification. However, an unbiased approach relies on the screening of libraries of compounds (synthetic, virtual, and even natural product libraries) to identify a small number of active compounds. The specific screening approach depends on the information and technology available and can encompass high-throughput screening of millions of compounds, virtual screening of tens of millions of virtual compounds, and fragment-based screening based on structural information of the target. Once hits are identified, triage and further prioritization are essential components of the process to ensure that only the most promising structures will move ahead.

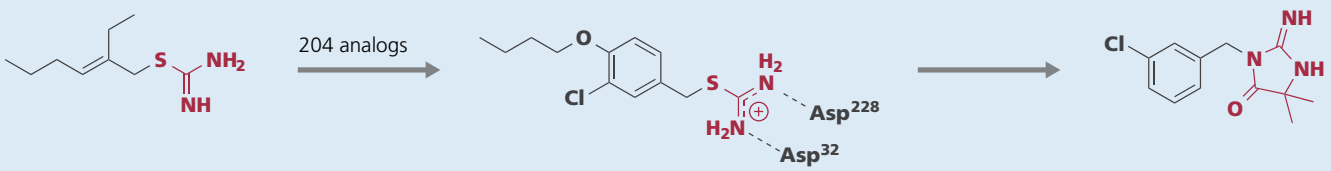
CASE STUDY 1 EXAMPLE OF FRAGMENT-BASED SCREENING

In this case study, a fragment-based drug discovery approach to the aspartic protease β-secretase (BACE-1) inhibitor is described. The approach led to the discovery of a potential Alzheimer's drug that advanced to clinical trials.

The approach used a combination of fragment-based NMR screening, X-ray crystallography, computational chemistry, and chemical modification of hits. Using ^{15}N-labeled BACE-1 protein, over 10,000 fragments were screened at concentrations of 100 μM–1 mM, as mixtures of 12 each, based on ^{15}N heteronuclear single quantum coherence (HSQC) NMR. This NMR technique, which correlates nitrogen and hydrogen signals, is commonly used in FBDD. It allows the observation of changes in the ^{15}N-labeled target (BACE-1 in this case) that occur when a compound binds. From this assay, nine structural classes of hit compounds were identified, binding with a K_d (dissociation constant) in the 30 μM–3 mM range. Of these classes, the authors focused on an isothiourea hit and prepared 204 analogs of this compound, leading to compound **1**, with a tight binding profile (low K_d) and low molecular weight, as preferred for drug-like compounds (**Figure 2.18**). This compound was optimized, in part, by studying X-ray structures of analogs. An X-ray structure of **1** co-crystallized with BACE-1 identified an important hydrogen-bonding network between the isothiourea portion and aspartates 32 and 228 in BACE-1.

One liability of the isothiourea scaffold was its instability. Modification of that portion of the molecule, inspired by other moieties with a similar pK_a (6–10) that maintained the hydrogen-bonding network, led to iminohydantoin **2** (see **Figure 2.19**). Although the binding of **2** (K_d = 200 μM) to the target enzyme was poorer than that of **1** (K_d = 15 μM), improved stability made this a more desirable lead compound.

Analysis of the binding site of the enzyme, where **2** was co-crystallized with BACE-1, identified binding pockets that could be occupied more efficiently by growing the hit at the site currently occupied by the *gem*-dimethyl groups, as seen in **Figure 2.19**. This ultimately provided **3**, with aromatic rings in place of the *gem*-dimethyl groups, as a lead. Attributes include low micromolar enzyme inhibition as well as appropriate lipophilicity (assessed by cLogP), a druglike quality. Further optimization can be seen in **Figure 2.19**. Appending a 3-pyridyl group to one of the phenyls afforded additional binding interactions in a pocket of the enzyme, as shown in the solved X-ray structure, resulting in even greater potency. An advanced analog in this series, verubecestat, entered phase III clinical trials as a drug to treat Alzheimer's disease.



isothiourea hit
K_d = 550 μM
MW = 186

compound 1
K_d = 15 μM
MW = 273

compound 2
iminohydantoin
K_d = 200 μM

Figure 2.18 Progression of initial isothiourea hit to analog 1. An isothiourea lead was identified by FBS and evolved to compound **1**, with a higher MW but tighter binding to the target. X-ray structure showed its interaction with key aspartic acid residues in the enzyme active site. Iminohydantoin compound **2** was designed to maintain interactions with the aspartic acids but improve stability. It has decreased binding compared to **1**, but improved stability.

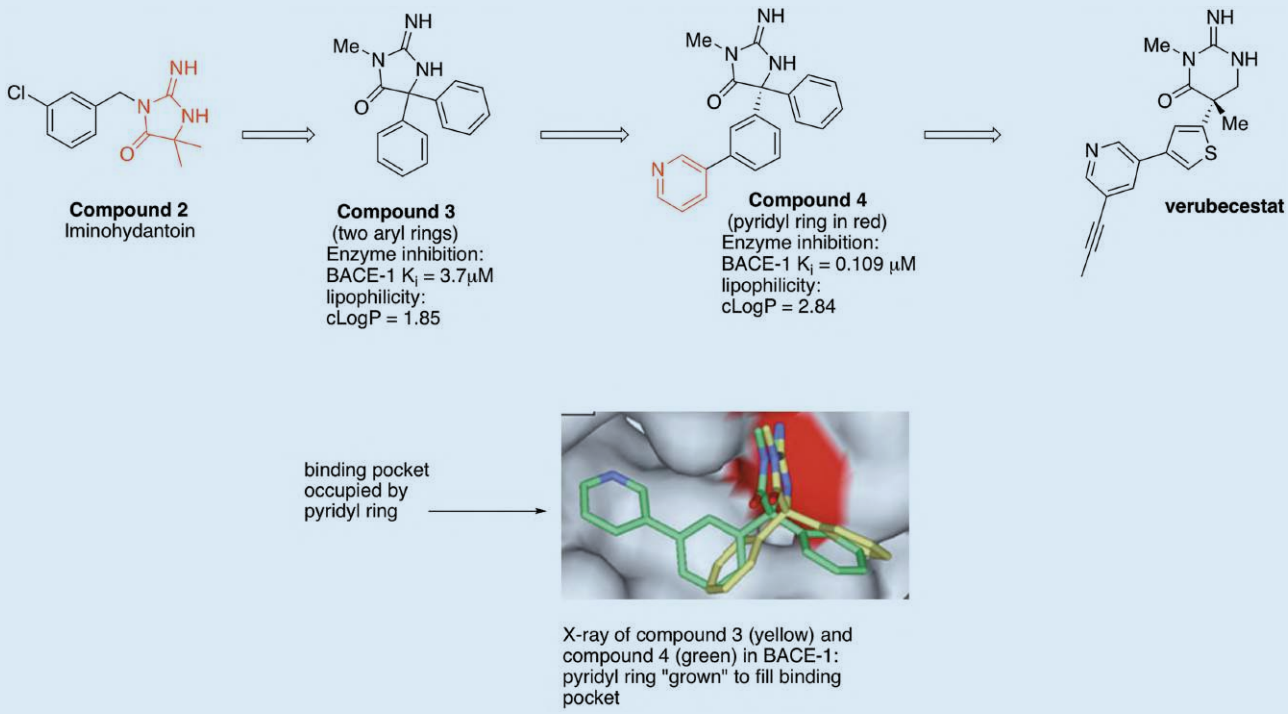


Figure 2.19 Optimization of iminohydantoin 2 to generate verubecestat. Compound 2 was optimized based on analysis of X-ray co-structures to generate compound 3, which had much higher affinity. Further optimization to incorporate an additional pyridine ring, shown in red (4), afforded even higher potency. Verubecestat, the final compound, entered clinical trials for Alzheimer's disease. X-ray structures of compounds 3 and 4 are shown: The pyridine ring fills an additional binding pocket.

Source: Wyss, D. F.; Wang, Y.-S.; Eaton, H. L.; et al. (2012). Top Curr Chem 317:83–114. With permission from Springer.

CASE STUDY 2 HIT IDENTIFICATION USING A FOCUSED LIBRARY

This case study describes an example of the synthesis and evaluation of a focused library for the identification of novel potential anti-tumor compounds. The authors in this study synthesized and screened a focused library in search for inhibitors of HDAC as potential anti-tumor compounds.

Histone deacetylases (HDAC) are enzymes that catalyze deacetylation of lysine residues on histones and other proteins. They are involved in epigenetic processing, which is necessary for cell proliferation. There are 11 members of the HDAC family. Pan-inhibition of HDACs is the mechanism of several commercial drugs used as anti-tumor agents, including vorinostat, belinostat, panobinostat, and romidepsin (described in Chapter 10). Recent efforts have targeted the inhibition of specific HDACs vs. pan-inhibition, based on the idea that more selective inhibition will lead to less toxicity. The authors chose to focus on inhibitors of HDAC1 and HDAC2 with slow-binding kinetics based on data that shows such mode of inhibition leads to prolonged pharmacological effect, in this case, anti-proliferative activity. There was also interest in inhibitors of HDAC2 as potential therapeutics for neurodegenerative disease. The goal of the project was to design and prepare a focused library and screen for the inhibition of HDAC1 and HDAC2 and to investigate the kinetics of binding.

The library was designed on the basis of four elements common to known inhibitors: (1) an *o*-aminoanilide core that binds to zinc in the active site, (2) a "cap" group that interacts with residues at the entrance to the active site, (3) a linker connecting the zinc-binding group (ZBG) to the cap, and (4) a "foot-pocket recognition group" (FPRG) found in some inhibitors that is an aryl group connected to the *o*-aminoanilide core. The basis of the synthesis is summarized in **Figure 2.20** and used CuAAC click chemistry (copper-assisted azide-alkyne cycloaddition) to assemble a focused library of 768 triazoles, where the triazole core is the linker.

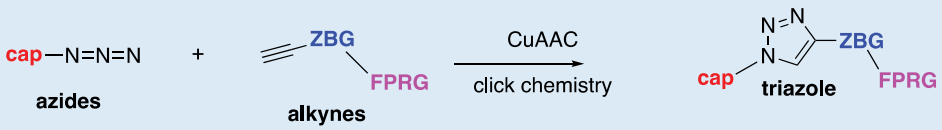
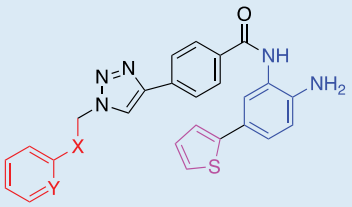
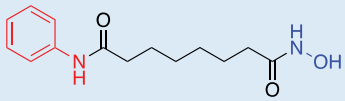


Figure 2.20 Key step in the synthesis of a triazole library using azides containing a "cap" group (red) and alkynes containing the zinc-binding group (blue) and foot-pocket recognition group (pink).



KPZ532: X = CH_2, Y = N
KPZ540: X = SO, Y = CH
KPZ560: X = SO_2, Y = CH



vorinostat

Figure 2.21 Structures of the three most active compounds from the focused library: KPZ532, KPZ540, and KPZ560. Vorinostat is shown for comparison. Compounds are color-coded as in Figure 2.20: red = cap, blue = ZBG, pink = FPRG.

The building blocks included 62 different aryl and alkyl azides and 12 different aryl alkynes. The azides contain the "cap" group, and the alkynes contain both the ZBG and the FPRG groups. The key step is a 1,3-dipolar addition of the azide to the alkyne, catalyzed by copper to afford the triazole core (linker).

Assay for inhibition of HDAC was performed using a commercial assay kit that measures the bioluminescence of a known substrate for HDAC. In this assay, three compounds inhibited HDAC2 activity by more than 75% at a 100 nM concentration, which is a higher inhibition than the drug vorinostat, a nonselective HDAC inhibitor. Shown in **Figure 2.21**, these compounds (KPZ532, KPZ532, and KPZ532) all contain an *o*-aminoanilide core as the ZBG, and a 2-thiophene group as the FPRG.

To confirm activity of the top three compounds, all were resynthesized, purified, and screened against all 11 isoforms of HDAC. The three compounds also showed inhibition of HDAC1, with little inhibition of HDACs3–11. Compounds KPZ532, KPZ540, and KPZ560 were analyzed for their kinetics of inhibition, using a fluorescence-based assay, and were all shown to be slow-binding inhibitors, which implies a longer lifetime of inhibition. Molecular modelling using virtual docking of KPZ560 into the HDAC2 core reveals important binding interactions (**Figure 2.22**). In particular, the oxygen and amine of the o-aminoaniline bind to zinc (ZBG), and the thiophene (FPRG) is in the "foot pocket" formed by hydrophobic groups Tyr 29, Met 35, Phe 114, Ile 144, and Cys 156. The phenyl "cap" interacts with Tyr 209, Phe 210, Pro 211, and Leu 276 in surface loops.

Further validation of these hits in cell-based assays showed the inhibition of MCF7 breast cancer cells, as well as enhanced neurite growth of N2a cells and increase in dendritic spine density in mice. These results are promising, both for potential cancer treatment and for neurodegenerative disorders.

This example shows the value in relying on previous data—in this case, pharmacophores for HDAC inhibition—to design a relatively small focused library of compounds. Also important was the synthetic strategy which allowed the rapid assembly of the library with diverse building blocks. Three hits with the desired activity (HDAC1/2 inhibition) were identified, with potential to lead to new entities to treat both cancer and neurodegenerative diseases.

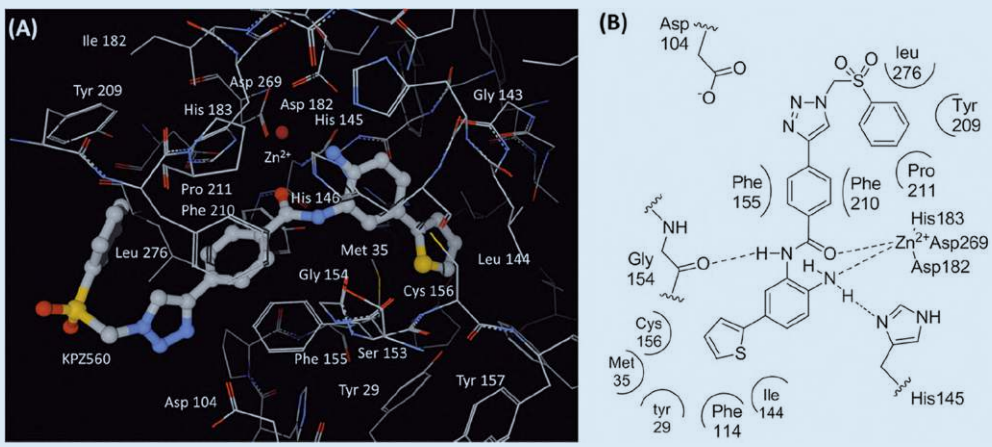


Figure 2.22 *Left*, **view of the inhibitor KPZ560 (gray) docked into the active site of HDAC2.** *Right*, schematic diagram of inhibitor with specific side chain interactions.

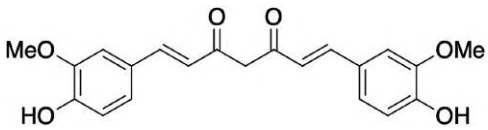
Source: J. Med. Chem. (2023), 66, 15,171–15,188. https://doi.org/10.1021/acs.jmedchem.3c01095.

2.9 REVIEW QUESTIONS

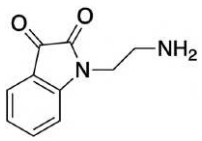
1. What are three different naturally occurring sources of hits and leads? Give an example of a drug that is directly or indirectly derived from each.
2. Explain the major differences between split-pool combinatorial synthesis and parallel synthesis.
3. How does FBS differ from conventional HTS, and what are the benefits and limitations of each?
4. Describe the pros and cons of using a phenotypic assay vs. a biochemical assay.
5. Explain the difference between a hit and a lead compound.
6. What are the advantages of a drug discovery program that is based on clinical observations?
7. Describe the pros and cons of using a natural ligand as the starting point for a drug discovery effort.
8. When would structure-based virtual screening be most appropriate to apply to hit identification? What information is necessary? What are the disadvantages of using this strategy?
9. Under what circumstances is a ligand-based virtual screening strategy appropriate? What information is required? What are the disadvantages of using this strategy?
10. Name three categories of compounds that are often false positives in high-throughput screens.
11. Give an explanation of how a Tanimoto coefficient is used.

2.10 APPLICATION QUESTIONS

1. Atorvastatin is one of the most widely used cholesterol-lowering agents. Its design was inspired by the natural products compactin and lovastatin. Draw the structures of the three compounds. What structural features do they have in common?
2. Research a drug whose structure is based on a natural ligand of (a) an enzyme substrate and (b) a receptor. Draw the natural ligand and the drug and compare the structures. Highlight the similar features.
3. The combinatorial, solid-phase synthesis of a benzodiazepine library was published in 1995 (Plunkett MJ & Ellman JA [1995] Solid-phase synthesis of structurally diverse 1,4-benzodiazepine derivatives using the Stille coupling reaction. *J Am Chem Soc* 1 17:3306–3307). (a) Design a library of 50–75 final products by choosing specific reagents to use at each step. (b) In general terms, how would you design a library of more than 5,000 compounds?
4. The use of DNA-encoded libraries has become accepted practice. Search J. Med. Chem. to find an example of the use of a DEL for hit identification.
   a. Outline the synthetic steps used in the synthesis.
   b. Describe what type of assays were used.
   c. Draw the structure of a hit identified from screening.
5. Shown is the structure of a well-known natural product. What is this natural product, and from where is it isolated? What medicinal properties have been attributed to it? If you found it to be active in a high-throughput screen, how would you evaluate it as far as its being a high-quality hit and why?



6. The compound shown was identified as a hit in a screen but is considered a poor-quality hit. Explain why. Design two analogs that, when tested, would determine whether the offending functional group was required for activity.



7. Modern phenotypic screening often relies on whole organisms, such as zebrafish or flies. Research why these particular organisms have been so useful. Find a research paper that describes a high-throughput screen using one of these organisms to identify a small-molecule hit.
8. You have solved an X-ray structure of a kinase enzyme with an inhibitor bound to the active site. You also have developed a biochemical assay that detects the reaction products of the kinase reaction. Describe two hit identification methods that would be most appropriate for the information you have, briefly describe the strategy, and explain why you would use it.
9. The drug target Bcl-2 and related family members are proteins that regulate cell death through interaction with other protein partners. Their structures have been solved by both X-ray crystallography and NMR. Describe two hit identification methods that would be most appropriate for the information you have, briefly describe the strategy, and explain why you would use it. What types of assays could be used to evaluate hits? Is there any hit identification strategy you would not apply? Why not?
10. Use one of the online tools (for example, http://chemmine.ucr.edu/help/) for calculating a Tanimoto coefficient to compare two drug molecules that work at the same target, for example, sumatriptan and rizatriptan. Compare them to the natural ligand for the 5-HT receptor. How similar is each to 5-HT and to each other based on the Tanimoto coefficient? Is this result surprising?

FURTHER READING

Brown DG & Bostrom J (2018) Where do recent small molecule clinical development candidates come from? *J Med Chem* 61:9442–9468.

High-Throughput Screening and Library Synthesis

Bleicher KR, Bohm H-J, Muller K & Alanine AI (2003) Hit and lead generation: beyond high-throughput screening. *Nat Rev Drug Discov* 2:369–378.

Deng H, O'Keefe H, Davie CP et al. (2012) Discovery of highly potent and selective small molecule ADAMTS-5 inhibitors that inhibit human cartilage degradation via encoded library technology (ELT). *J Med Chem* 55:7061–7079.

Geseru GM & Masara GM (2009) The influence of lead discovery strategies on the properties of drug candidate. *Nat Rev Drug Discov* 8:203–212.

Houghten RA (1985) General method for the rapid solid-phase synthesis of large numbers of peptides: specificity of antigen-antibody interaction at the level of individual amino acids. *Proc Natl Acad Sci USA* 82:5131–5135.

Kleiner RE, Dumelin CE & Liu DR (2011) Small-molecule discovery from DNA-encoded chemical libraries. *Chem Soc Rev* 40:5707–5717.

Reiher CA, Schuman DP, Simmons N & Wolkenberg SE (2021) Trends in hit-to-lead optimization following DNA-encoded library screens. *ACS Med Chem Lett* 12:343–350.

Tan DS (2005) Diversity-oriented synthesis: exploring the intersections between chemistry and biology. *Nat Chem Biol* 1:74–84.

Thompson LA & Ellman JA (1996) Synthesis and applications of small molecule libraries. *Chem Rev* 96:555–600.

Walters WP & Namchuk M (2003) Designing screens: how to make your hits a hit. *Nat Rev Drug Discov* 2:259–266.

Fragment-Based Screening

Erlanson DA, Fesik SW, Hubbard RE, Jahnke W & Jhoti H (2016) Twenty years on: the impact of fragments on drug discovery. *Nat Rev Drug Discov* 15:605–619.

Erlanson DA, Wells JA & Braisted AC (2004) Tethering: fragment-based drug discovery. *Ann Rev Biophys Biomol Struct* 33:199–223.

Virtual Screening

Bajorath J (2002) Integration of virtual and high-throughput screening. *Nat Rev Drug Discov* 1:882–894.

Markt P, Petersen RK, Flindt EN et al. (2008) Discovery of novel PPAR ligands by a virtual screening approach based on pharmacophore modeling, 3D shape, and electrostatic similarity screening. *J Med Chem* 51:6303–6317.

Sun H (2008) Pharmacophore-based virtual screening. *Curr Med Chem* 12:1018–1024.

Artificial Intelligence and Machine Learning

Borkakati M & Thornton JM (2023) Alphafold2 protein structure prediction: implications for drug discovery. *Curr Opin Struct Biol* 78:102526.

Galati S, DiStefano M, Martinelli E, Poli G & Tuccinardi T (2021) Recent advances in in silico target fishing. *Molecules* 26:5124–5142.

Identifying High-Quality Hits

Baell J & Walters MA (2014) Chemical con artists foil drug discovery. *Nature* 513:481–483.

Baell JB & Holloway GA (2010) New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J Med Chem* 53:2719–2740.

McGovern SL, Helfand BT, Feng B & Shoichet BK (2003) A specific mechanism of nonspecific inhibition. *J Med Chem* 46:4265–4272.

Walters WP & Murcko M (2002) Prediction of drug-likeness. *Adv Drug Deliv Rev* 54:255–271.

List of Publicly Available Resources

AlphaFold. https://alphafold.ebi.ac.uk

Database of compounds and bioactivity: PubChem (pubchem.ncbi.nlm.nih.gov/); ChEMBL. https://www.ebi.ac.uk/chembl/

Databases of drugs: CMC (Comprehensive Medicinal Chemistry); MDDR (MACCS-II Drug Data Report); WDI (World Drug Index)

Database of virtual compounds: ZINC. http://zinc.docking.org/

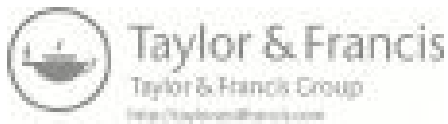
Examples of non-drugs: ACD. Available Chemical Directory

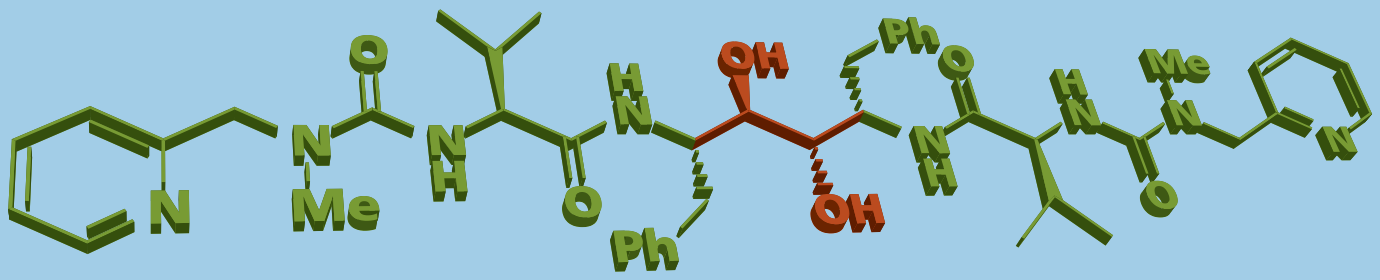
Hit filtering: ADMET, false positives, etc. http://www.cbligand.org/xie-lab/technology.php

Case Studies

Itoh Y, Zhan P, Tojo T et al. (2023) Discovery of selective histone deacetylase 1 and 2 inhibitors: screening of a focused library constructed by click chemistry, kinetic binding analysis, and biological evaluation. *J Med Chem* 66:15171–15188.

Wyss DF, Wang Y-S, Eaton HL et al. (2011) Combining NMR and X-ray crystallography in fragment-based drug discovery: discovery of highly potent and selective BACE-1 inhibitors. *Top Curr Chem* 317:83–114.


Taylor & Francis
Taylor & Francis Group



Lead Optimization: Drug–Target Interactions and the Pharmacophore

3

LEARNING OBJECTIVES

- Identify properties evaluated and improved during lead optimization.
- Recognize the molecular interactions involved in drug–target binding.
- Describe some historical and current methods to evaluate structure–activity relationships.
- Understand how pharmacophores are identified and utilized.
- Learn the principles of ligand-based and structure-based drug design.

3.1 INTRODUCTION

Once a bioactive lead has been identified, there is still a lengthy process involved in converting that lead to a drug used in humans. While there are notable exceptions where the initially discovered hit or lead becomes a drug, most modern drug discovery requires a process of **lead optimization**. Lead optimization involves chemical modification of the structure of a hit or a lead compound through an iterative series of design, synthesis, and testing. This process eventually generates a clinical candidate, which, after further development and study, is appropriate for clinical trials in humans (**Figure 3.1**). This clinical candidate is a compound that exhibits the desired potency, selectivity, efficacy, and physical and pharmaceutical properties, as well as an optimal safety profile. After testing in humans and clearing of regulatory requirements by the appropriate agency, the compound becomes an approved drug.

The multiparameter optimization of leads requires numerous compounds to be tested in a battery of assays, and the results of those assays inform the next round of design, synthesis, and testing. The process usually starts with testing in relatively simple *in vitro* assays and then, as compounds meet certain requirements, progresses to testing in increasingly complex secondary assays, and eventually to evaluation in animal (*in vivo*) models of disease. The end result is a compound that fulfills all the criteria required to become a clinical candidate that will eventually be tested in humans (**Figure 3.2**). However, lead optimization does not always begin with a primary assay but may be initiated at any point in the process where a liability in an advanced compound is uncovered and improvements in the molecule are required.

DOI: 10.1201/ 9781003465607-4

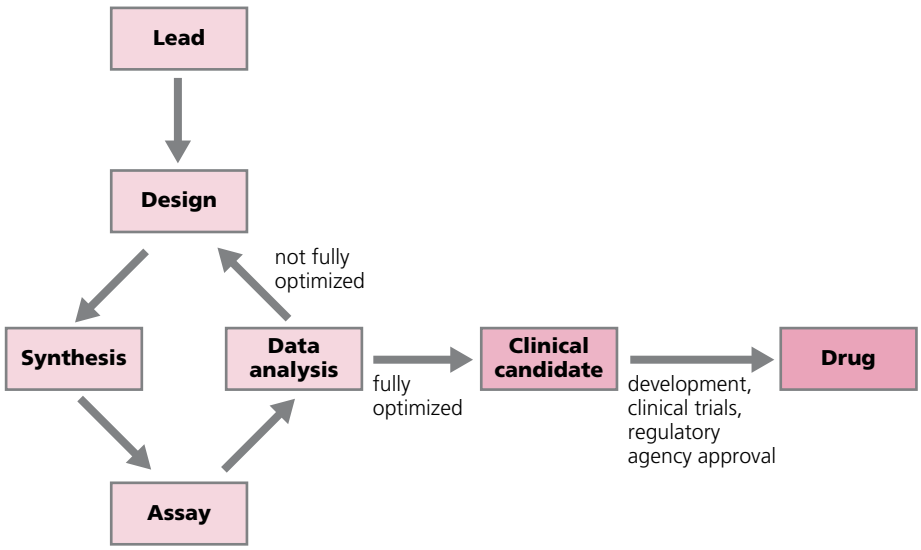


Figure 3.1 Iterative process of lead optimization, beginning with design of modified leads. After synthesis and assay, data are analyzed, and if acceptable, the modified lead becomes a clinical candidate. If unacceptable, further rounds of design, synthesis, assay, and analysis are undertaken until a fully optimized compound is generated. After further testing, clinical trials, and regulatory approval, the candidate becomes an approved drug.

While the entire process of testing is outlined in Figure 3.2, the focus can be divided roughly into three main activities:

1. Testing for biological potency and efficacy is one key component. The lead compound is optimized by modification and then tested for biological activity in primary and secondary assays. An important aspect of this phase is the optimization of drug–target interactions involved in binding.
2. A second component, often done in parallel, focuses on optimizing druglike properties. These druglike properties include **physical properties**, such as solubility and lipophilicity, that affect the ability of molecules to interact with biological systems. Also included are **pharmaceutical properties**, which are measured to predict or evaluate the effect of living systems on the small molecule. Success in these phases leads to compounds that are evaluated in animal models of disease.
3. The final component of drug development involves additional testing in animals, progression to human clinical trials, and eventually, drug approval.

The first of these three components, understanding the molecular basis of biological activity, with an emphasis on drug–target interactions, is the focus of the remainder of this chapter. Chapter 4 describes medicinal chemistry strategies for optimizing biological, physical, and pharmaceutical properties, and Chapter 5 outlines the final stages of the drug development process.

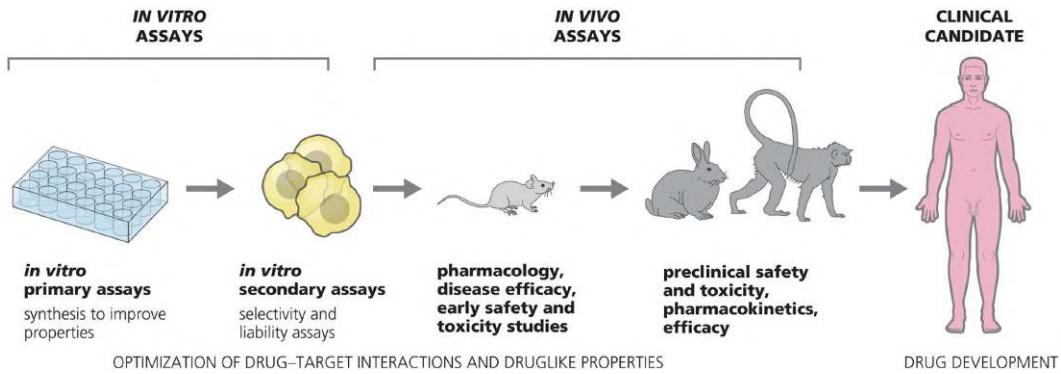


Figure 3.2 Progressive process of testing during lead optimization. Compounds are tested in primary assays, and those that pass a specified threshold advance to secondary and then *in vivo* (animal) testing during the lead optimization process. Compounds are optimized *in vivo* for toxicity, safety, pharmacokinetics, and efficacy to identify a compound appropriate for clinical trials. The optimized compounds become clinical candidates and are evaluated in a further battery of tests.

3.2 OPTIMIZATION OF DRUG–TARGET INTERACTIONS

In the early stages of lead optimization, the measurement and the optimization of biological activity are key activities. Additionally, ensuring high affinity for the desired target and low affinity for other targets mitigates safety issues. In many cases, a biochemical assay is used, one that involves measuring the direct effects of a small molecule on a biochemical macromolecule. In other cases, a more complex system is used to measure a downstream event that is the result of the small molecule binding to a macromolecule. Regardless of the measurement, the overall goal is to optimize binding at the desired molecular target before moving on to more complicated secondary assays. To do so requires understanding of how the structure of the small molecule interacts with the biological target.

3.2.1 Optimization of Binding Depends on Specific Molecular Interactions

Optimization of binding is directly related to improving potency and efficacy as well as selectivity. The affinity of a drug (D) to its protein target (P) is thought of as in equilibrium and is typically characterized by a dissociation constant $K_{dissociation}$ (K_d), which compares the concentration of the drug–protein (D-P) complex to the concentrations of the individual drug (D) and protein (P) components and measures how readily the drug–protein complex dissociates (Equation 3.1).

$$\mathrm{D\text{-}P} \rightleftharpoons \mathrm{D} + \mathrm{P} \text{ and } K_d = \frac{[\mathrm{D}][\mathrm{P}]}{[\mathrm{D\text{-}P}]} \qquad 3.1$$

On the basis of this equation, the smaller the K_d, the tighter the binding. Typical K_d values for drug–protein interactions are in the range of 10^{-7}–10^{-9} M.

This constant is directly proportional to the Gibbs free energy of binding (ΔG) according to Equation 3.2, where R is the gas constant and T is the absolute temperature.

$$\Delta G = -RT \ln K_d \qquad 3.2$$

ΔG is also a function of the binding enthalpy (ΔH) and the binding entropy (ΔS):

$$\Delta G = \Delta H - T\Delta S \qquad 3.3$$

Therefore, for an interaction between a drug and its biological target to occur, two properties must be considered and optimized. **Enthalpy** (ΔH) involves the energetics of the drug–target system and is related to the specific interactions between the compound and the target. More energetically favorable interactions result in tighter binding, and therefore, the interaction is likely to occur. The **entropy** term (ΔS) involves the disorder of the system. This term usually opposes the enthalpy term, since greater disorder is more favorable, but order is required to make specific interactions between the small molecule and the target. One goal of medicinal chemistry activities is to maximize the enthalpy (that is, stronger and more binding interactions) and minimize the entropy terms.

Drugs typically bind to their biological target through noncovalent forces, mimicking the interactions observed in proteins that dictate their secondary structure. Therefore, understanding the possible interactions of functional groups present in amino acids and proteins is critical to understanding drug–protein interactions. These interactions include electrostatic, hydrogen-bonding, hydrophobic, dipole, and cation–π interactions, among others, and are energetically favorable. Any gain in enthalpy through the formation of new interactions between the drug and the target usually comes with an entropic penalty. This compensation is based on the specific geometric requirements of most binding interactions, which result in an ordered state of the complex. Energy is often required to orient the drug and protein target to optimal orientations, and once the complex is formed, rotational and translational freedom is limited compared to the individual components in an unbound state, and this process is entropically disfavored.

The process of a drug binding to its protein target is a complex set of events. The small-molecule drug, surrounded by a shell of water molecules, must undergo conformational rearrangements, breaking any interactions it has with the surrounding water molecules to find its way into a binding site of the protein and make new and more energetically favorable interactions. This process is called **desolvation**. Similarly, solvent molecules occupy the vacant target binding site itself and must be displaced by the drug for drug–target binding to occur. When both the drug and the binding site are relatively nonpolar, this is usually a favorable process. However, when either (or both) is highly polar, and therefore interacts favorably with water, this process is energetically unfavorable and can be offset only if the new interactions between drug and protein target are so favorable that they overcome the desolvation penalty.

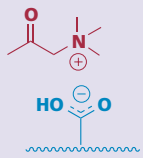
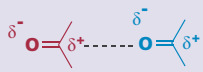
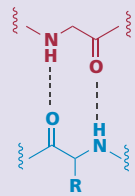
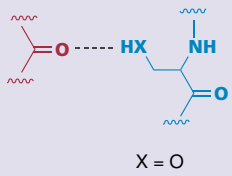
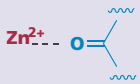
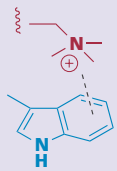
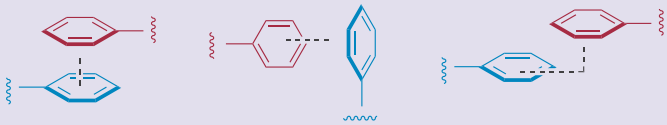
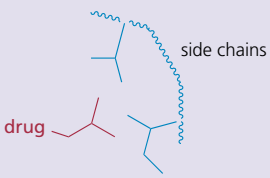
3.2.2 Specific Interactions Contribute to the Overall Binding of Drugs to Their Targets

The vast majority of drugs bind to particular amino acids (and sometimes metals) that comprise their target through a combination of specific binding interactions of varying strengths. Sometimes the interactions are covalent; however, most are noncovalent and depend on particular functional groups. These are described in the following and are illustrated in **Table 3.1**.

The strongest noncovalent forces occur between two ions and are termed **electrostatic (ionic) interactions**. These occur between oppositely charged species, such as a carboxylate anion and an ammonium cation. The strength is approximately 5–10 kcal/mol and decreases in proportion to the square of the distance between the two atoms. Amino acid side chains containing acids (such as glutamic acid or aspartic acid) or amino groups (such as lysine or arginine) often participate in electrostatic interactions with drug molecules containing complementary functional groups. **Dipole–dipole interactions** occur between functional groups that have permanent dipoles, such as carbonyl groups and carbon-halogen moieties. The partial positive charge on one atom is attracted to the partial negative charge on the other. The presence of amide carbonyl groups in protein backbones offers abundant opportunities for this type of interaction in drug–protein complexes. An important subclass of dipole–dipole interactions is **hydrogen-bonding interactions**. Classical H-bonding interactions take place between a hydrogen donor that is covalently bound to an electronegative atom, such as oxygen or nitrogen (N-H, O-H), and an electronegative atom acceptor, such as oxygen, nitrogen, or fluorine. Optimal hydrogen bonds have specific geometries (often linear) and bond lengths (~2.4–3.0 Å). The energy of a hydrogen bond can be extremely high, but ΔG usually ranges from 2 to 5 kcal/mol. The amide function of each amino acid in target proteins offers H-bond donors and acceptors, as do many functionalized amino acid side chains; therefore, this interaction is frequently observed in protein–drug complexes. Weak hydrogen-bonding interactions have also been observed between many unconventional donor/acceptor pairs. Examples include CH/NH$\cdots\pi$ systems, C-H$\cdots$O/N, and S-H$\cdots$O/N. These weak interactions may not have the same geometric constraints as classical H-bonds, but nonetheless, they can contribute to drug–target binding.

Ion–dipole interactions are observed between a charged species and a polar functional group. Common examples are a metal cation or protonated amine interacting with the oxygen atom of a carbonyl. These interactions are estimated to be between 0.5 and 2.0 kcal/mol in energy. Certain drug targets contain metals in their active sites, and these interactions are particularly important in those cases. Positively charged ions may also interact with π systems, and these **cation–π interactions** occur between the π system of an aromatic ring and a cation. The aromatic amino acids phenylalanine and tryptophan can contribute to the π system; the cation may be either a metal or a protonated amine. The strength of these interactions is estimated to be between 0.2 and 2.5 kcal/mol. Another important class of interactions is **π–π interactions**. These occur between π systems in aromatic groups, are frequently observed in drug–protein complexes, and are also the basis for interactions between nucleobases in DNA and RNA. They occur with various geometries, termed sandwich, T-shaped (or edge-to-face), and parallel displaced. The weakest noncovalent interactions

Table 3.1 **Typical noncovalent binding interactions between drug and target.**

Binding interaction	Example	Target residue, fragment or atom	Energy (kcal/mol)
ion–ion (electrostatic, ionic)		cation: Lys, Arg anion: Asp, Glu	5–10
dipole–dipole		backbone carbonyls	2–5
hydrogen bonding	 amide backbone  X = O side chains	backbone amide; Ser, Thr, Tyr (OH); Lys, Arg, His (NH)	2–5
ion–dipole			0.5–2.0
cation–π		cation: Lys, Arg π donor: Phe, Trp, Tyr	0.2–2.5
π–π	 sandwich T-shaped parallel displaced	Phe, Tyr, Trp, His	0.5–1.0
van der Waals	 side chains drug	Ala, Val, Leu,Ile	0.5–1.0

(0.5–1.0 kcal/mol) are **hydrophobic interactions** or **van der Waals forces**. These occur between nonpolar groups and involve temporary dipoles. The most common examples of these in drug–protein complexes occur between nonpolar alkyl groups in the drug and alkyl amino acid side chains in the protein. While these are weak interactions, they are strong contributors to drug–protein binding. In addition to enthalpic considerations, because they displace water from unfavorable interactions with nonpolar areas within the protein and drug–solvent shell (desolvation), these interactions are also entropically favorable. As many more high-resolution co-structures of small molecules bound to drug targets are being solved, some other types of interactions

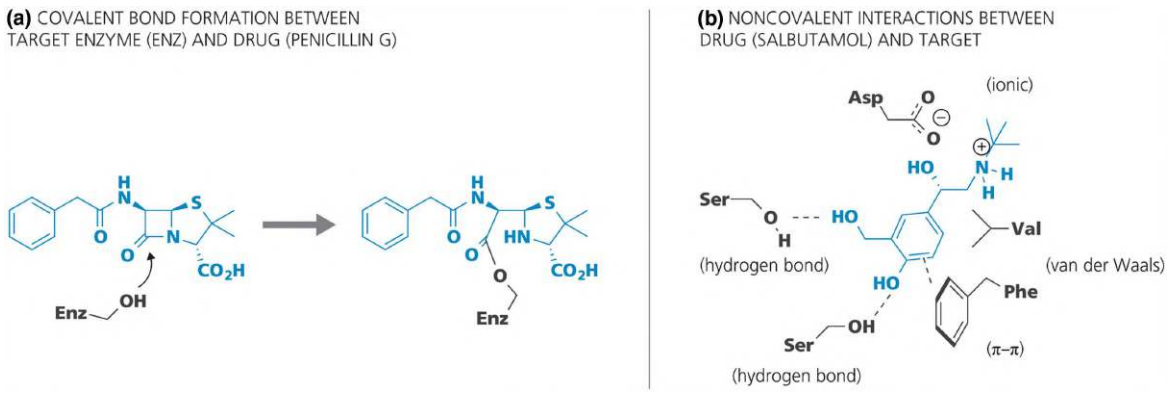


Figure 3.3 Examples of drugs binding through covalent and noncovalent interactions. (a) A covalent bond forms between the carbonyl of penicillin G (blue) and a serine of the target enzyme (ENZ, black). (b) Salbutamol (blue) binds through H-bonding, π–π, ionic, and van der Waals interactions to its target (black), the beta-2 adrenergic receptor.

have been observed, including halogen bonding, orthogonal multipolar interactions, sigma hole interactions, and interactions between the sulfur of a methionine and carbonyls or aryl rings.

Most interactions between drugs and their targets involve noncovalent interactions. However, there are several instances where **covalent binding** occurs and a formal bond is formed between the protein and the drug. This interaction is the strongest type of binding (50–150 kcal/mol). Irreversible enzyme inhibitors, such as penicillin G, bind through covalent interactions, usually between a nucleophilic amino acid side chain residue, such as serine, threonine, or cysteine, and an electrophilic site on the drug (**Figure 3.3a**). Another example of irreversible covalent binding occurs in the anti-tumor nitrogen mustards, which covalently modify nucleophilic sites on DNA. More recently, drugs binding through reversible covalent interactions have also been developed through many of the same interactions usually between a nucleophile in the target and an electrophile in the drug.

Most drugs rely on a combination of various binding interactions to contribute to their potency and efficacy. These attributes are a direct result of **biomolecular recognition** or specific binding interactions between the drug and the target macromolecule. The asthma drug salbutamol, a member of the family of β-agonists that bind at β-adrenoceptors, is an illustrative example (see **Figure 3.3b**).

Binding interactions of salbutamol include an ionic interaction between the protonated amine and the carboxylate of an aspartic acid (Asp), a π–π interaction between the aromatic ring and a phenylalanine (Phe), van der Waals (hydrophobic) interaction between the aromatic ring and a valine isopropyl group (Val), and hydrogen-bonding interactions between the hydroxyls and two serine hydroxyls (Ser).

Stereochemistry also plays an important role in drug–receptor binding interactions. In a set of enantiomers, each contains identical functional groups that are capable of forming the same interaction with a specific residue of a protein target; however, their three-dimensional orientation is different. Therefore, their overall binding will differ. This effect was recognized in the 1930s when different effects of enantiomeric phenethylamines on adrenergic activity was observed.

Enantiomers of a drug often vary in potency. However, there may also be differences in rates of metabolism due to differences in affinity for metabolic enzymes, as well as differences in toxicity that each displays. The more active enantiomer is called the **eutomer**, and the less active enantiomer is the **distomer**. The ratio of activity of eutomer to distomer is the **eudismic ratio**, which gives an indication of stereoselectivity of the drug–target interaction. For example, citalopram, a selective serotonin reuptake inhibitor, has a eudismic ratio of 167, with the *S*-enantiomer being the active isomer. Citalopram is marketed as a racemic mixture (Celexa), and (*S*)-citalopram is marketed as escitalopram (Lexapro); both are used to treat anxiety and depression (**Figure 3.4**). Methylphenidate, a widely used medication for attention deficit hyperactivity disorder (ADHD), is a racemic mixture of *R*,*R*- and *S*,*S*-enantiomers and also has a high eudismic ratio. The more active *R*,*R*-enantiomer is also marketed.

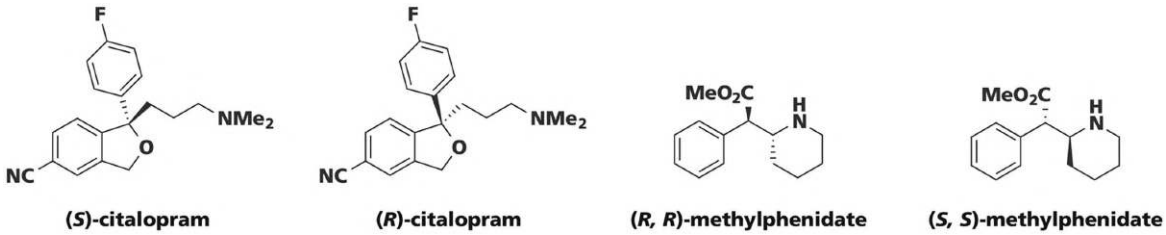


Figure 3.4 Enantiomers of citalopram and methylphenidate. (*S*)-citalopram is marketed as escitalopram and is 167 times more active than the *R*-enantiomer. Methylphenidate also exhibits a high eudismic ratio and is marketed as both the racemate and the *R*,*R*-isomer.

3.3 TOOLS USED TO DEVELOP STRUCTURE–ACTIVITY RELATIONSHIPS

As described in the previous section, the ability of a drug to bind to its target is dependent on its various structural features. Medicinal chemists employ multiple tools to understand how molecules bind to a target, as well as to design novel compounds based on the data generated in the iterative cycle in **Figure 3.1**. A key element of the design process is the development of **structure–activity relationships** (SAR), that is, the effect of a change in the structure of the molecule on its biological activity. This process is an important underpinning of medicinal chemistry. Tools that incorporate SAR include pharmacophore modeling, commonly used in ligand-based drug design, and structure-based drug design, which relies on knowledge of the structure of the macromolecular target bound to ligands. Artificial intelligence and machine learning algorithms also rely on the availability of SAR data.

The relationship between structure and activity was recognized even before our current understanding of protein structures and was addressed in the mid-1900s by the investigation of structure–activity relationships of **congeners**, compounds with similar core structures but varying functional group appendages. SAR studies are used throughout the lead optimization process. This entails modifying specific functional groups in the molecule, either actually or virtually, and measuring the activity of the new analogs to determine whether the specific moiety is involved in binding to the target or, more broadly, what effect that change in structure has on activity. For example, if a lead molecule contains a hydrogen bond donor, a molecule is designed in which that hydrogen bond donor is removed while the other features of the molecule are maintained. When tested, if the molecule is considerably less active, the hydrogen bond donor is deemed to be a critical feature of the structure. If the molecule maintains its activity, that hydrogen bond donor is not an essential feature. In the specific case of salbutamol (see **Figure 3.3b**), the removal of the benzylic hydroxyl group removes that particular hydrogen bond and also decreases its activity as a bronchodilator.

While this concept seems simple, replacing one moiety without modifying another aspect of the molecule can be challenging. A hypothetical example is shown in **Figure 3.5**. Replacing the hydroxyl group in a potential hydrogen bond donor, such as an alcohol, with a methoxy group will remove the H-bonding donor capacity, but it also changes the size and polarity of that particular feature. If the new molecule is less active, one conclusion is that the H-bond is important; however, the possibility that the new methyl group is too large to fit into the binding pocket cannot be ruled out. If it is similarly or more active, one conclusion might be that the H-bond is not important for binding, but the oxygen in the lead and the congener acts as an H-bond acceptor. Replacing the alcohol with a secondary amine would maintain H-bond donor and acceptor features. If either was important for binding, the activity of the secondary amine should be maintained or perhaps improved. To further corroborate the SAR data, the secondary amine could be changed to a tertiary amine, which removes the H-bond-donating capability. Inactivity of this analog would confirm the importance of the H-bond-donating feature. Equivalent or improved activity would indicate the importance of the H-bond-accepting capabilities. In either case, further optimization may include the

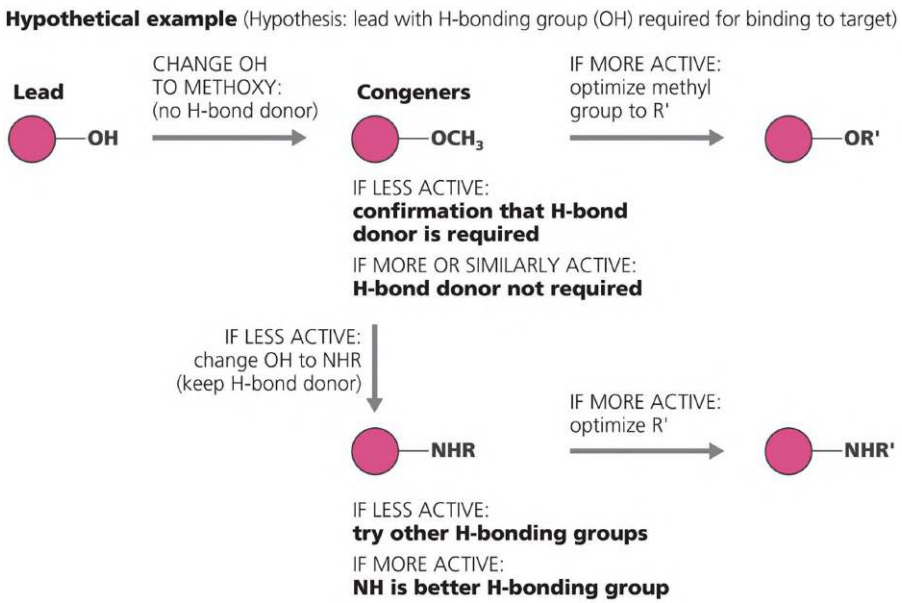


Figure 3.5 Hypothetical example showing lead optimization by synthesizing various congeners of a lead compound. This example shows how the logical process of systematically changing the structure of a specific functional group (OH group) leads to conclusions about essential binding interactions.

design and synthesis of more amine analogs with various alkyl or acyl (R') groups to probe the size of the binding pocket and perhaps optimize van der Waals interactions.

While the hypothetical example in Figure 3.5 shows a simple modification of one group, typically several different functional groups of the lead are modified to optimize each part of the compound. Furthermore, corroborating data need to be accumulated to support the conclusions based on each cycle of design, synthesis, testing, and analysis.

3.3.1 Historical Methods for Quantifying Structure–Activity Relationships Were Based on Physicochemical Parameters and Presence or Absence of Defined Groups

To streamline the process of drug discovery, starting in the mid-1900s, methods were developed to quantify structure–activity relationships. The earliest **quantitative structure–activity relationship** (QSAR) models were two-dimensional and relied on either properties of fragments of the molecules or properties of the entire molecule. Today these methods are not used frequently; however, they are the basis of many of the modern computational algorithms applied today, and therefore, an understanding of this background is valuable. The goal of two-dimensional QSAR was to develop a mathematical relationship between the structures and the activities of a series of compounds. This model could then be used to guide the future design of new compounds in the series. In general, *biological activity* was defined as a function of the physicochemical properties of the drug, with certain specific **physicochemical parameters** correlating with biological activity. The most common parameters quantified were electronic characteristics, lipophilicity, and size of certain functional groups.

Historically, **electronic properties** have been quantified in a number of ways. One of the earliest used was the Hammett equation, which had been developed to study organic reactions and their mechanisms. Hammett had derived a series of constants (σ_x) that characterize the effects of electron-donating or electron-withdrawing substituents (X) on the dissociation of benzoic acid. The constants quantified the effect on an equilibrium constant (K_x) compared to that of benzoic acid itself (K_H).

The importance of considering **lipophilicity**, the affinity of a molecule for lipids or nonpolar media, in SAR models came from the appreciation that drugs need to cross cell membranes in order to pass between the gut, blood, and tissues containing the

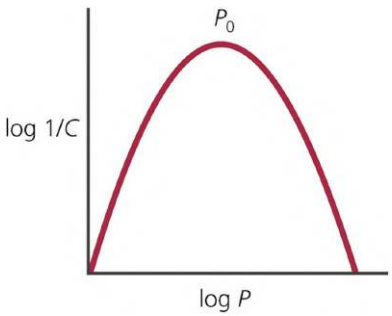


Figure 3.6 Relationship between drug concentration (log 1/*C*) and lipophilicity (log*P*) for a series of compounds where bioactivity correlates to lipophilicity. At low log*P* (nonlipophilic compound), a high concentration is required for activity; at the optimal log*P*, a low concentration of drug is required for activity; and at high log*P* (very lipophilic), a high concentration is required for activity.

macromolecular targets. If a molecule was not lipophilic enough, it would not cross the membrane, but if it were too lipophilic, it would get trapped in membranes and never reach the biological target. The pioneers of QSAR, Corwin Hansch and Toshio Fujita, proposed that a lipophilicity parameter should be incorporated into QSAR analyses. This property could be measured experimentally by partitioning the molecule between 1-octanol (which is lipid-like) and water. The partition coefficient, *P*, is usually defined as the concentration of the compound in 1-octanol divided by its concentration in water. If a compound is equally soluble in both, then log*P* = 0. If it is more soluble in water than in 1-octanol (hydrophilic), then log*P* < 0, and if it is more soluble in 1-octanol than in water (lipophilic), then log*P* > 0. The more lipophilic the drug, the higher the value of log*P*. According to Hansch, there is a parabolic relationship between the reciprocal of concentration of compound at which biological activity is observed and lipophilicity (**Figure 3.6**). For nonlipophilic compounds (those with low log*P*), a high concentration (*C*) would be required for activity; compounds that have an optimal log*P* require the lowest concentration of compound for activity, and compounds that are very lipophilic (a high log*P*) need a high concentration for activity.

Hansch also measured lipophilicity substituent constants, π, for various atoms and functional groups. The π constants are both additive and constitutive—meaning that π values for groups within a molecule could be added to give the overall lipophilicity or log*P* for similar compounds.

Finally, several methods were developed to measure the size of various substituents on a lead molecule. Robert Taft measured the rates of hydrolysis of esters with alkyl substituents of varying sizes, RCO_2R', compared to the corresponding acetate esters, CO_3CO_2R'. This value, the steric parameter E_S, was used as a surrogate measurement of size. A more widely used set of parameters are the Verloop parameters that are calculated by the program Sterimol. These values are based on bond angles, van der Waals radii, bond lengths, and possible conformers of various substituents to give five values: L = length of substituent along the bond axis from the parent and B_1–B_4 = width of substituent perpendicular to the bond axis in four directions. One other example of a parameter that reflects size is the molar refractivity (MR), which is based on molecular weight, density, and index of refraction. This parameter gives an indication of the molar volume of a substituent as well as its polarizability.

In order to correlate these parameters with biological activity, a number of mathematical equations were developed. The Hansch analysis, among the earliest, could be used for a congeneric series. This is a linear multiple regression analysis that relates bioactivity to physical properties. An example of the **Hansch equation** using electrostatic (σ), lipophilicity (π), and size (E_s) parameters is as follows (Equation 3.5):

$$\text{Hansch equation}: \log 1/C = -a\pi^2 + b\pi + \sigma + cE_s + d \qquad 3.5$$

To apply this, within a given series of compounds, a number of analogs are tested for biological activity, where *C* is the concentration at which some given specified activity is observed and the calculated physicochemical parameters π, σ, and E_s (described earlier) are taken from sources in the literature. Regression coefficients *a*, *b*, *c*, and *d* are derived from a training set of known compounds. The reciprocal of concentration reflects the fact that a lower dose correlates to an inherently more potent drug. Once the coefficients are known, the equation can be solved for hypothetical compounds to give the predicted biological activity. Therefore, the medicinal chemist can predict whether or not a proposed compound may have activity without going to the trouble of synthesizing and testing it. However, for any given series of compounds, a large number of analogs have to be synthesized and tested for activity in order to develop a highly predictive equation.

Another classical two-dimensional QSAR method was the **Free–Wilson method**, or **additivity method**, introduced by Spencer Free and James Wilson. This method

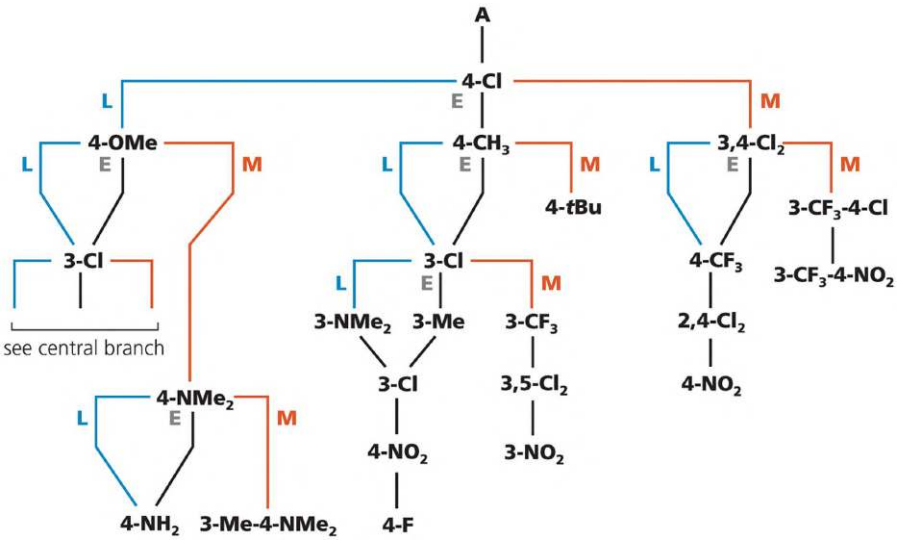


Figure 3.7 Original Topliss scheme for aromatic substituents. A substituted derivative (4-Cl) of an aromatic parent compound (A) is prepared and tested. The results of testing (L = less active; E = equally active; M = more active) determine the next analog to prepare. On this manner, an efficient, systematic process for evaluating SAR and lead optimization is pursued.

assesses the occurrence of additive substituent effects and estimates their magnitude. Unlike the Hansch method, this equation does not take into account any properties of the molecules or fragments, but it tries to directly relate the effects of changes in structure to changes in biological activity. The general form of the equation is as follows (Equation 3.6):

$$\text{Free} - \text{Wilson or additivity equation: BA} = \sum a_i X_i + \mu \qquad 3.6$$

In this case, *BA* is biological activity, a_i is the magnitude of the effect of a given substituent, and μ is average activity of the parent molecule. *X* is given a value of 1 if substituent *i* is present, and 0 if not. As in a Hansch analysis, a series of congeners needs to be prepared and tested for biological activity. These are then analyzed to determine values of *a* for various substituents, as well as a value for the average activity. Those values are then used to predict bioactivity for compounds in the same series that have not yet been synthesized and tested.

The development of either a Hansch equation or Free–Wilson analysis required the synthesis and testing of a fairly large number of compounds as a basis set. As an alternative to these methods, in the 1970s, John Topliss developed a flow diagram (**Topliss decision tree**) that began with the biological analysis of a single compound (**Figure 3.7**). The bioactivity of a compound containing a substituted benzene (active or inactive) determines which is the next analog to prepare. The scheme was designed on the basis of the electronic and lipophilic characteristics of the substituents and relied on Hansch principles.

For example, if a compound bearing an aromatic ring had biological activity, synthesis of the 4-chloro analog was undertaken, and that analog would be either less, equally, or more active than the parent compound. The specific data determine which analog to prepare next: If the 4-chloro analog is more active, then the 3,4-dichloro analog should be made; if the 4-chloro analog is less active, then the 4-methoxy analog is prepared next. This process was iterative, with various substituents and points of substitution modified depending on the data from the previous compound. The Topliss decision tree allowed medicinal chemists to make rational decisions about which analogs to make on the basis of smaller numbers of congeners compared to Hansch and Free–Wilson methods, and it avoided the need to develop complex equations that describe biological activity.

Today these historical methods *as described* are not used, but they have been modified and updated for the modern computational era. For example, Hansch analysis is the foundation for fragment-based QSAR, which uses molecular fingerprints,

or fragments, to carry out two-dimensional QSAR assessment based on a partial least squares analysis. Rather than experimental σ values for electronics, or Taft parameters for size, electrostatics (Coulomb) and sterics (Lennard–Jones) are calculated computationally for a molecule. Likewise, the concept of ligand efficiency, which will be described in Chapter 4, is essentially a progression of Free–Wilson analysis. Variations of the Topliss scheme are still used to decide on the synthesis of analogs with aromatic substituents; indeed, there is some concern about an overall bias toward *p*-chlorophenyl substituents in many libraries due to the historical widespread application of Topliss schemes.

3.3.2 Modern Computational Methods for QSAR

The classical two-dimensional QSAR method for drug design was a very time-consuming process. Many analogs had to be synthesized and tested for activity in a congeneric series in order to develop an appropriate equation. This could take years, depending on how many compounds were made. The initial process was essentially one of trial and error, with all imaginable changes being tried before a useful equation could be derived that accurately predicted active compounds prior to synthesis. While the roots of QSAR lie in these historical methods, analysis today is done using computer-assisted drug design (CADD), which, in more recent years, has progressed to AI/ML methods. The overall goal is still the same: developing a computational model that can predict bioactivity and streamline the drug discovery process.

All require input of molecular data and bioactivity to develop a model whose output is the predicted bioactivity of compounds. The molecular data that is "input" are descriptors that may range from 0D (MW, atom count, bond count) or 1D (fingerprints), and more commonly 2D (graphic representation) and 3D (molecular matrix), and finally, 4D (grid-based). In addition, the biological activity associated with the structures is also part of the input. These data can be in any form that measures biological data. In more sophisticated models, data from multiple assays (e.g., *in vitro*, *in vivo*, physical properties) are all part of the input. To start, a training set and a validation set are designed. Each is a series of known compounds plus their associated data. The training set is used to develop the model; once developed, the model is used to predict the activity of the known compounds in the validation set. If the model predicts the known compounds in the validation set with high accuracy, it is considered validated and then used to predict bioactivity of new compounds. If the model is inaccurate, refinements are made until it is validated. As data for new compounds become available, the model is continually refined.

Many algorithms have been developed for QSAR. These include multiple linear regression (MLR), such as was used in Hansch and Free–Wilson analyses, and random forest (RF) models, which are the average output of multiple decision trees. Currently, many models rely on machine learning methods, such as SVM (support vector machines), RF, and DNN (deep neural networks). These newer methods begin with the input of large amounts of data, but the machine learning aspect allows the prediction of bioactive compounds with features that were not in the original data input. The application of these techniques in hit identification is summarized in what follows:

- *SVM.* The support vector machine algorithm allows for the classification of data points. It identifies a hyperplane that gives the maximum distance between classes of data, with support vectors being those points closest to the hyperplane. In the context of hit identification, most often, the two sets of data points include active vs. inactive compounds.
- *RF.* A random forest algorithm builds a series of unique decision trees, each using different input. The results are then averaged to give the output.
- *DNN.* Deep neural networks consist of an input layer—then hidden layers of learning—that generates an output layer.

Rough visual representations of these techniques are shown in **Figure 3.8**; however, the specifics of the computational methods are beyond the scope of this text.

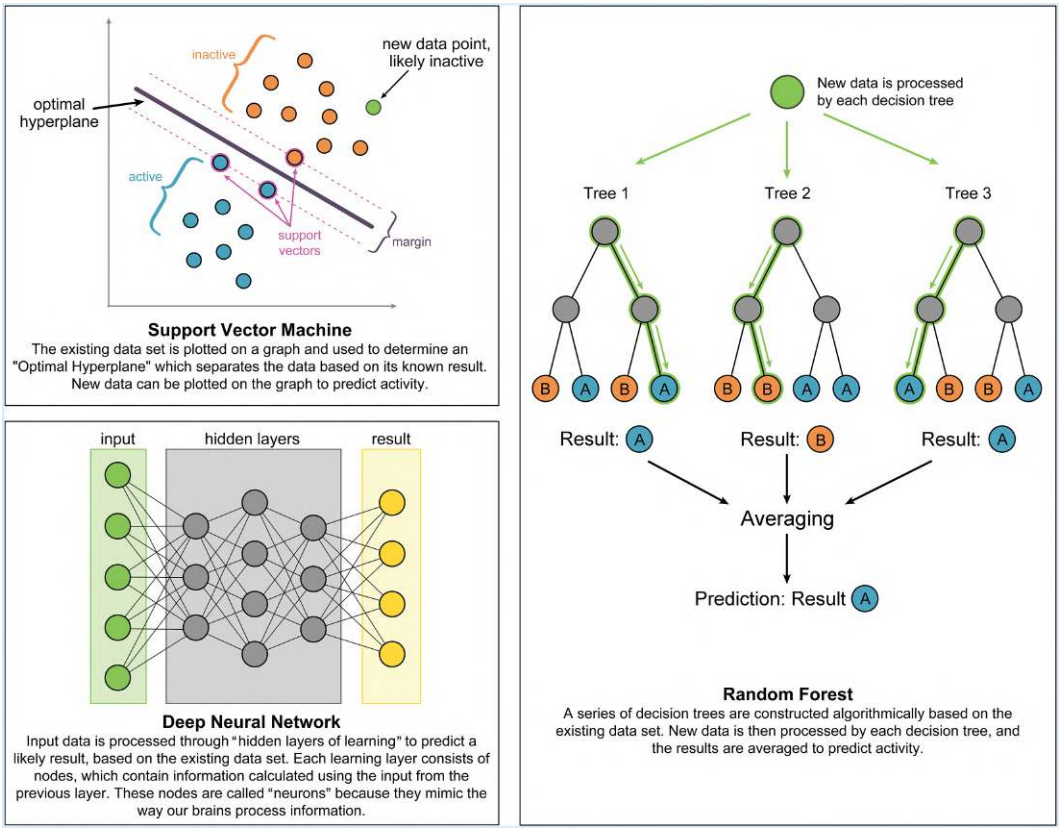


Figure 3.8 A summary of common AI/ML techniques used for hit and lead identification. These include support vector machines (SVM), random forest (RF), and deep neural networks (DNN).

3.3.3 Ligand- and Structure-Based Drug Design Rely on Knowledge of the Pharmacophore

Biomolecular recognition is a key component of drug design and relies on the interactions outlined in Section 3.2 and the predictions of the three-dimensional fit into a target's binding site. With advances in computational abilities, medicinal chemists have come to use the important tool of three-dimensional QSAR for rational drug design. Two general strategies are employed. **Ligand-based drug design** depends on the analysis of low-energy conformers of compounds that are known to bind to a specific target. **Structure-based drug design** relies on a known structure of the macromolecular drug target and is applied when a structure of the macromolecular target, or a closely related homolog, is available. All these methods rely on a linchpin concept in medicinal chemistry: the pharmacophore.

In any series of active compounds, there are key binding interactions with the target that are made by specific functional groups of the molecule (Section 3.2). The target has a defined three-dimensional structure, as do the important functional groups of the drug. The **pharmacophore** is defined by the International Union of Pure and Applied Chemistry (IUPAC) as "the ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target structure and to trigger (or to block) its biological response." Further notes specify that this is an abstract concept accounting for common interaction capacities of a group of molecules toward the target macromolecule. Much of the rest of the molecule is either scaffolding to hold the pharmacophoric elements in the appropriate three-dimensional orientation or fragments that modify physical and pharmaceutical properties. The goal during the lead optimization process is to develop and understand the pharmacophoric elements and to make changes that improve the specific properties of the molecule. Knowledge of the pharmacophore tells you which parts of the molecule can be optimized to retain or even improve binding and, conversely, what moieties are not part of the pharmacophore and, therefore, can be modified to improve physical and pharmaceutical properties while binding to the target is retained.

Typical pharmacophores involve at least three binding elements that can include hydrogen bond donors or acceptors, hydrophobic regions, aromatic regions, and positive or negative ionizable groups. Specific functional groups within a molecule will fill the role of these pharmacophore elements. For example, an H-bond donor element in a pharmacophore could be achieved by an alcohol (RO**H**), amine (RN$\mathbf{H}_2$), or amide (RCON$\mathbf{H}_2$). When a pharmacophore is described, the specific functional group (such as alcohol, amine, or amide) is less important than the element (H-bond donor) it presents.

3.3.4 Pharmacophore Models Are Widely Used in Ligand-Based Drug Design

Multiple methods to develop a pharmacophore model have evolved. One of the earliest appearances of the term *pharmacophore* was in a study by Lemont Kier on the use of molecular orbital calculations of the preferred conformations of three compounds (acetylcholine, muscarine, and muscarone) that bind potently to the muscarinic acetylcholine receptor. By identifying the compounds' low-energy conformations and comparing similarities of their structures and the placement of functional groups, a pharmacophore model was derived. As seen in **Figure 3.9**, the three pharmacophoric elements identified were a positive ionizable group, presented by the quaternary amine in all three structures (red); an H-bond acceptor, presented by the carbonyls in acetylcholine and muscarone and by the oxygen atom of the alcohol in muscarine (blue); and a second H-bond acceptor, represented by the ester oxygen in acetylcholine and by the ring oxygens in muscarine and muscarone (green).

Overlaying these three structures provided the distance constraints between the pharmacophoric elements. Based on the model, any compound that fit these constraints should bind to the target. The concept of using a three-dimensional model to describe how small molecules interact with a target is central to ligand-based drug design. In these cases, the model developed from compounds known to bind to the target (ligands) is used to predict whether or not novel analogs will bind. The structure of the binding pocket of the target may not be known.

The **active analog approach**, pioneered by Garland Marshall in the late 1970s, used the three-dimensional arrangement of the low-energy conformers of multiple active compounds to find the minimal recognition elements. To apply the active analog approach, a set of both active molecules and inactive analogs is accumulated. After the common pharmacophore elements are identified, conformers of each ligand are generated computationally. The conformers are then superimposed on each other so that the common pharmacophore elements are matched. Many times these will be superimposed on a rigid active ligand that will necessarily have those elements in a fixed three-dimensional conformation.

A three-dimensional pharmacophore model can then be proposed and used in the design of new compounds. Applying this approach to a target macromolecule led

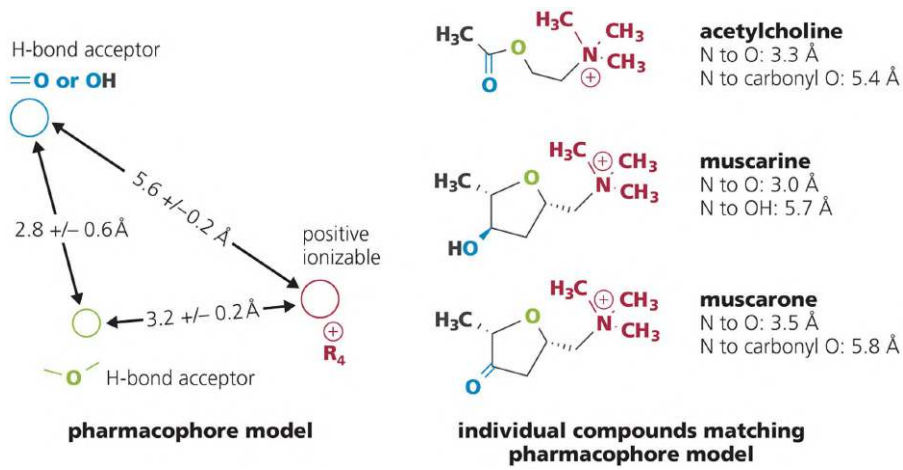


Figure 3.9 Early example of pharmacophore model based on compounds that bind to the acetylcholine receptor (acetylcholine, muscarine, and muscarone). The model on the left shows distances between pharmacophore elements in angstroms. Three groups are defined: blue and green = hydrogen bond acceptors; red = positive ionizable group. Three molecules on the right show specific functional groups that present the pharmacophoric elements highlighted and the distances between them.

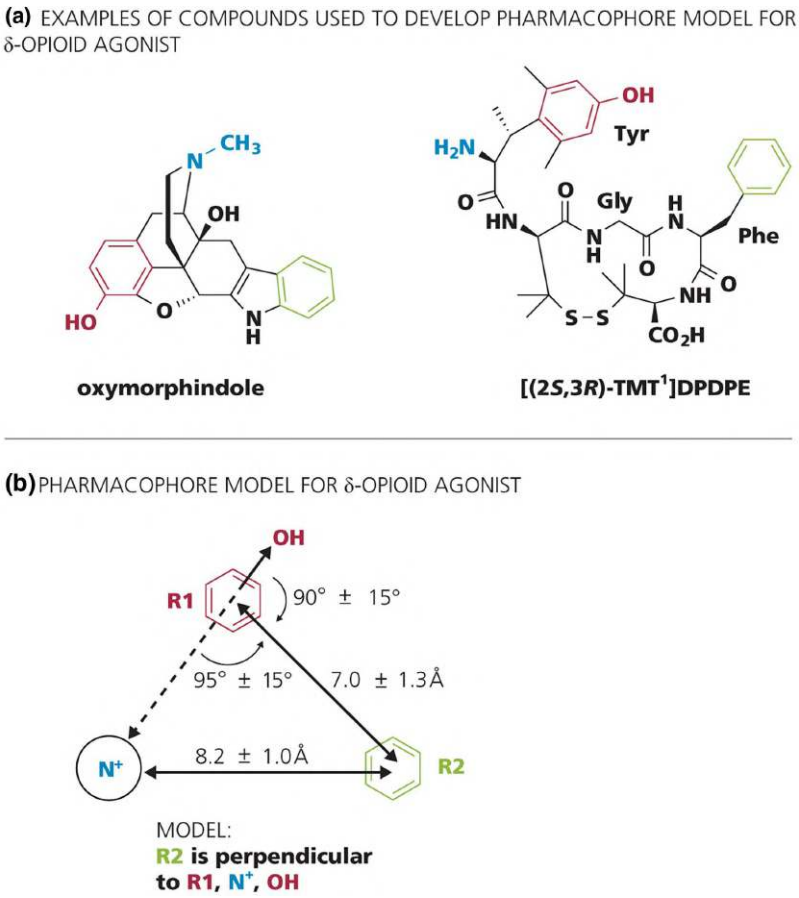


Figure 3.10 Hruby's δ-opioid agonist pharmacophore model: (a) Two of the compounds, oxymorphindole and [(2*S*,3*R*)-TMT[1]] DPDPE, used to develop the model. (b) Pharmacophore model contains four pharmacophore elements: red = aromatic group (R1) and OH, blue = positive ionizable group, green = aromatic group (R2). Distances in angstroms between groups are defined, with the R2 group perpendicular to both R1, OH, and the positive ionizable group.

to the concept of **receptor-excluded volume** (binding pocket area occupied by the drug) and **receptor-essential area** (binding pocket area occupied by the macromolecule target). The receptor-excluded volume is generated from the superimposition of all active compounds that contain the pharmacophore in their lowest energy conformation. A volume map is then generated from inactive compounds that also contain the pharmacophore. The assumption is that these inactive compounds, despite containing pharmacophoric elements, do not bind tightly because their volume extends into the area that is occupied by the target (receptor-essential area). This comparison between the volume of active analogs and the volume of inactive analogs provides a rough three-dimensional map of the size and shape of a binding pocket within a protein target and provides the boundaries beyond which active small-molecule binders cannot extend.

As an example, Victor Hruby and co-workers developed a pharmacophore model for selective δ-opioid agonists based on cyclic analogs of enkephalins, such as the cyclic peptide [(2*S*,3*R*)-TMT[1]]DPDPE (**Figure 3.10a**). The pharmacophore elements, in **Figure 3.10b**, were defined as an aromatic and an OH group (red, substituted tyrosine in [(2*S*,3*R*)TMT[1]]DPDPE), a positive ionizable group (blue, primary amine), and a second aromatic group (green, phenylalanine aryl ring). The model was developed by the use of rigid analogs with selectivity for the δ-opioid receptor, such as oxymorphindole, which contains the same pharmacophoric elements and allowed imposing distance constraints within the pharmacophore. The model developed (**Figure 3.10b**) contains four pharmacophoric elements and shows the spatial relationships between them.

Today, with the development of computational chemistry, pharmacophore modeling is often done by use of specific software programs. In these programs, the lowest energy conformations of a number of active and inactive ligands are determined and aligned to identify the common proposed points of interaction with a target. For example, scientists at GlaxoSmithKline used a commercial software program to

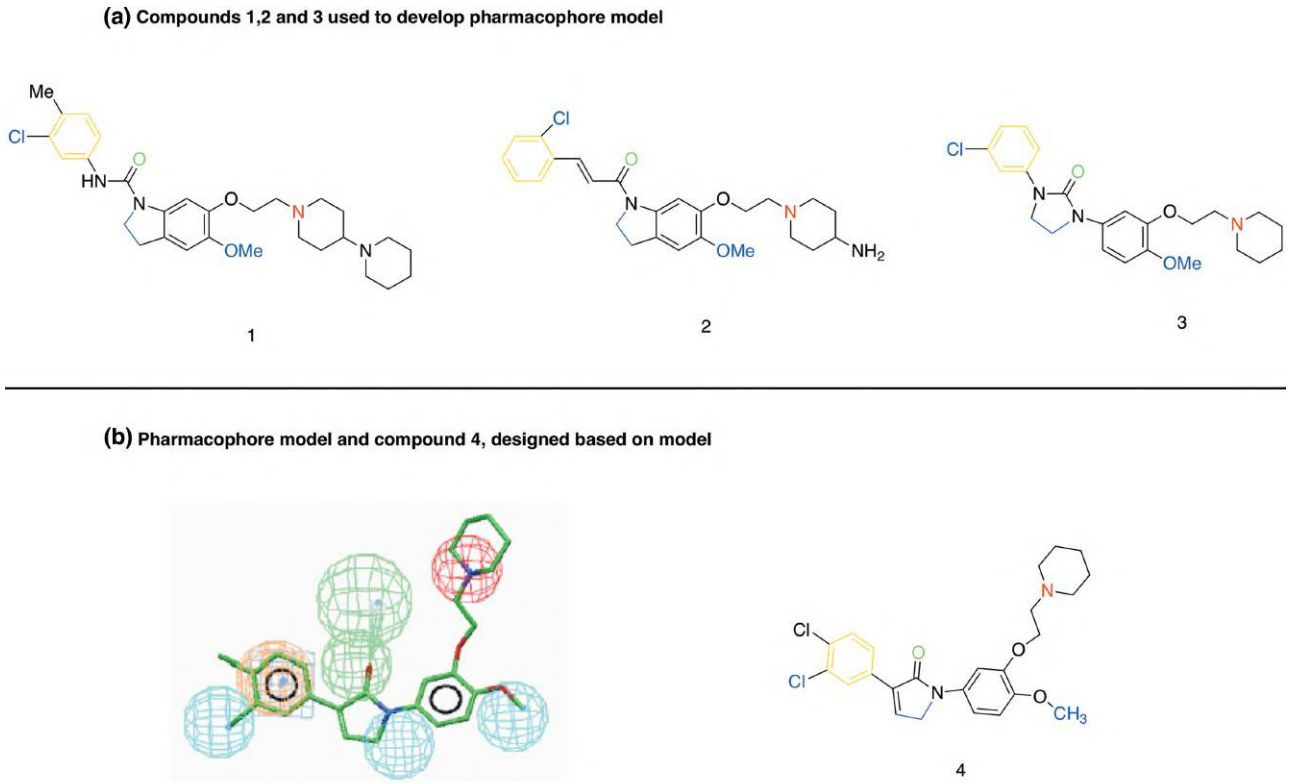


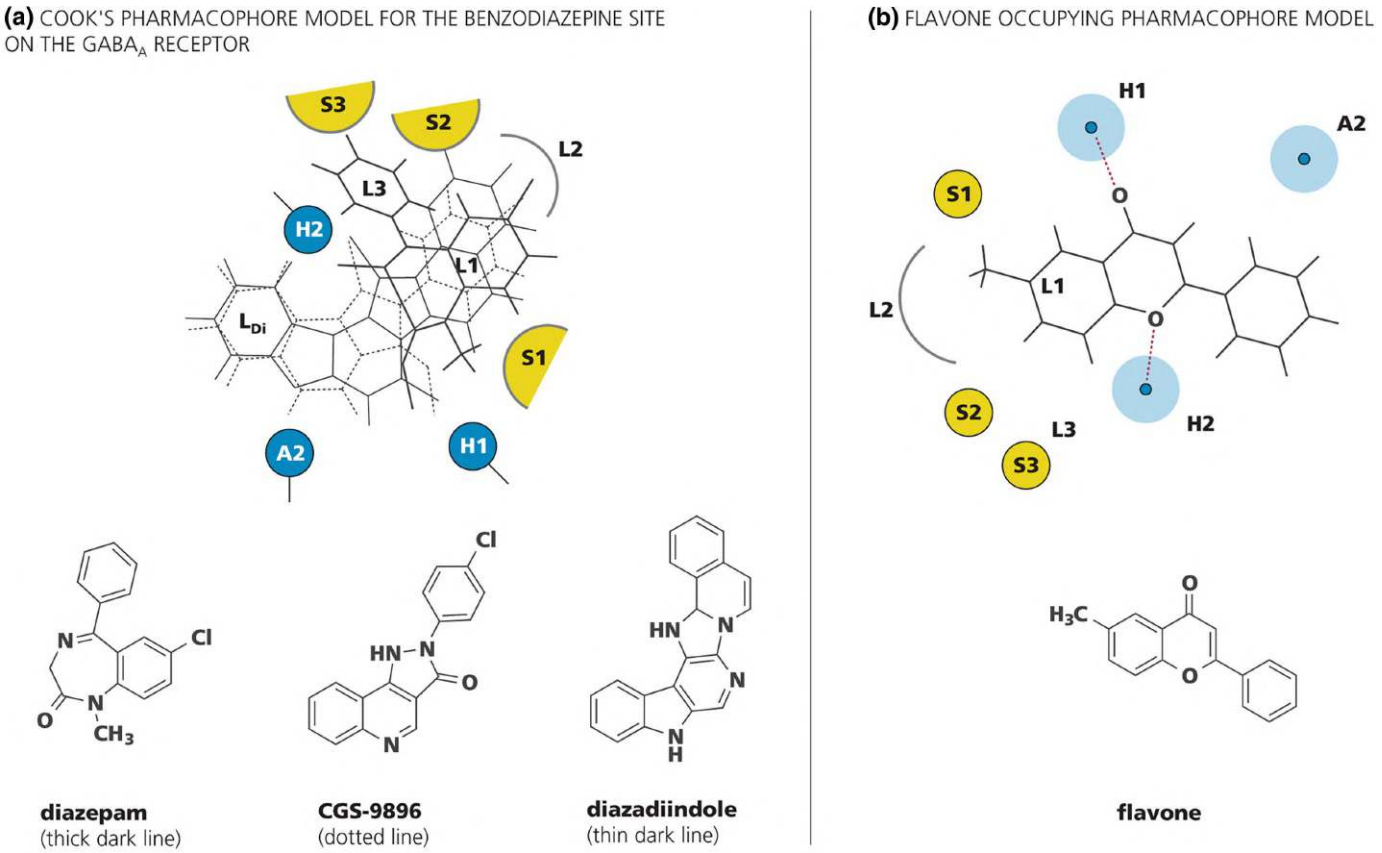
Figure 3.11 Three-dimensional pharmacophore model for ligands binding at the 5-HT$_{2c}$ receptor. (a) Analogs 1, 2, and 3 are examples of compounds that were used to develop a pharmacophore model; some key pharmacophoric elements are highlighted in colors defined later. (b) Pharmacophore model where the features are shown as spheres: green = H-bond acceptor; red = positive ionizable group; blue = hydrophobic groups; and yellow = aromatic ring. The volume and location of the sphere indicate the optimal position of the pharmacophoric element. The structure of the newly designed compound 4, developed on the basis of this model, is shown on the right. Common pharmacophore elements in all structures colored as pharmacophore model.

Source: From Micheli, F.; Pasquarello, A.; Tedesco, G; et al. (2006). *Bioorg Med Chem Lett* 16:3,906–3,912. With permission from Elsevier.

develop a pharmacophore model for ligands binding to one of the serotonin receptors (5-HT$_{2c}$). The pharmacophore model was based on compounds with high affinity and selectivity for the 5-HT$_{2c}$ receptor; the three structures in **Figure 3.11a** (**1–3**) are representatives of those compounds. These efforts resulted in a pharmacophore model (**Figure 3.11b**) in which the features are shown as spheres (green = H-bond acceptor; red = positive ionizable group; blue = hydrophobic group; and yellow = aromatic ring). The volume and location of the spheres define the optimal region in space where each of the elements must be located. A new compound (**4**) was designed on the basis of the model by ligand-based drug design; it is shown superimposed on the three-dimensional pharmacophore model.

Another example of the value of pharmacophore models is in virtual screening, as can be seen by the example of a pharmacophore model developed for a binding site on the GABA$_A$ receptor. Compounds such as benzodiazepines bind at this receptor and exhibit bioactivities, including anti-anxiety, muscle relaxation, and sedation. In the early 1990s, James Cook reported a model that was based on 136 benzodiazepines, β-carbolines, pyridodiindoles, pyrazoloquinolines, and several other scaffolds that bound to the benzodiazepine site on the GABA$_A$ receptor (**Figure 3.12a**). Later, as seen in **Figure 3.12b**, Tommy Liljefors and co-workers applied the model to include a different class of compounds, the flavones.

In these models, the features of both the small molecule and the target are described to characterize the pharmacophore: H1 and H2 represent sites on the receptor that contribute hydrogen bond donors (X-H from receptor), L1–L3 and L$_{Di}$



are lipophilic sites on the small molecule, S1–S3 are regions of negative steric interaction (corresponding to receptor-essential area), and A2 is a hydrogen bond acceptor site on the receptor.

In a new search for compounds binding at the $GABA_A$ receptor, Liljefors went further and used virtual screening based on this unified pharmacophore model and identified a 4-quinolone (K_i = 122 nM) as a potential lead for binding at the benzodiazepine receptor. This quinolone, which is a close analog of a flavone yet binds somewhat differently, fits the same pharmacophore model (**Figure 3.13a**). Lead optimization (**Figure 3.13b**) by extension of the ester side chain was based on the strategy to fill the L2 pocket, with a propyl group found to be the optimal length. Chain-branching (isopropyl) and ring-chain transformation at this position (cyclopentyl) led to analogs that were less active than the propyl analog, possibly due to steric interactions with the S1 and S2 sites. Optimization of the H1 binding interaction was accomplished by converting the ester to an amide, where the amide carbonyl oxygen is more electron-rich. By changing the R_2 group from trifluoromethyl to benzyl to fill the lipophilic interface region, an optimized lead with potent activity was identified.

Another example of a valuable computational tool for pharmacophore development, based on regression analysis and three-dimensional QSAR, is the use of CoMFA (comparative molecular field analysis) and CoMSIA (comparative molecular similarity index analysis) programs. In this type of analysis, introduced by Richard Cramer in the late 1980s, low-energy conformers for compounds with similar activity are generated and superimposed, matching pharmacophore elements. The computer establishes a three-dimensional grid around the compounds, and two fields, electrostatic and steric, are calculated for each molecule in every grid point by approaching the grid with atomic probes, such as an ion or a carbon atom. The calculations rely on Lennard–Jones and Coulomb potentials. A data table is generated, and a partial least squares analysis is done to determine the minimal set of grid points that explains the measured biological activity. Two different three-dimensional structures are generated:

Figure 3.12 Pharmacophore models for the benzodiazepine site on the $GABA_A$ receptor. (a) Cook's pharmacophore model (with diazepam, CGS-9896, and a diazadiindole overlaid), where H1 and H2 are H-bond donor sites and A2 is an H-bond acceptor site on the receptor; L1–L3 and L_{Di} are lipophilic sites on the small molecule, and S1–S3 are areas where steric interaction is unfavorable (receptor-essential area). (b) Cook's $GABA_A$ receptor pharmacophore (rotated 180° counterclockwise), with a flavone mapped onto the model. H2/A3 is either a donor or acceptor and corresponds to H2 in Cook's model.

Source: (a) From Clayton, T.; Chen, J. L.; Ernst, M.; et al. (2007). *Curr Med Chem* 14:2,755–2,775. With permission from Bentham Science Publishers. (b) From Kahnberg, P.; Lager, E.; Rosenberg, C.; et al. (2002). *J Med Chem* 45:4,188–4,201. With permission from the American Chemical Society.

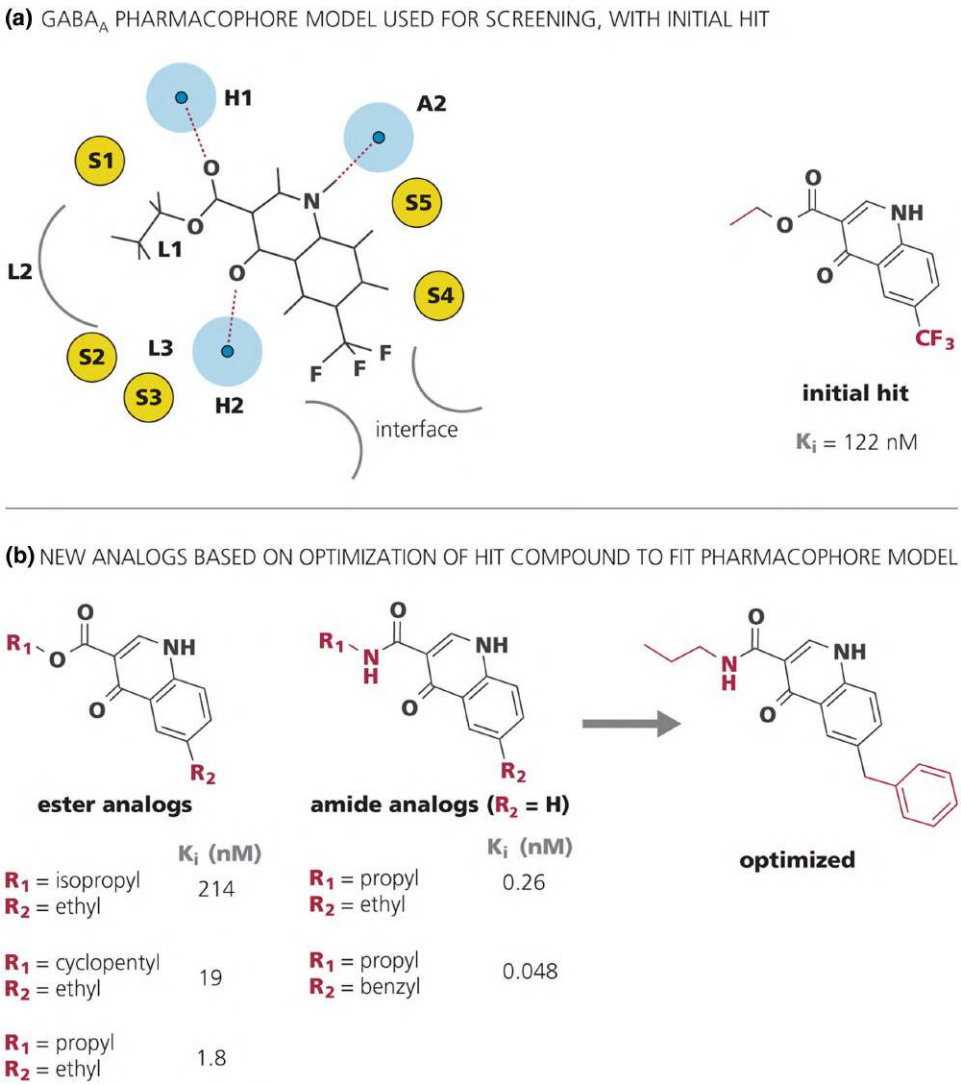


Figure 3.13 Use of a pharmacophore model in virtual screening and lead optimization. (a) The initial hit was discovered by virtual screening of a compound library to fit the pharmacophore model shown in Figure 3.12 with the addition of a lipophilic interface region. (b) Optimization of R1 and R2 groups and the ester linkage was done by fitting the model so that R1 (propyl) fills the L2 pocket and R2 occupies the lipophilic interface region. The corresponding amides improved the H1 interaction, and an optimized lead was generated by combining all changes. Lower *Ki* values indicate improved binding to the receptor.

Source: From Lager, E.; Andersson, P.; Nilsson, J.; et al. (2006). *J Med Chem* 49:2,526–2,533. With permission from the American Chemical Society.

one of the steric field, to show where it is desirable and undesirable to have steric bulk, and a second of the electronic field, to show where electron-rich and electron-poor areas are optimal. CoMSIA calculates similarity indices at the intersections of the surrounding grid using properties such as size, electrostatic potential, hydrophobic properties, and hydrogen bond donor and acceptor capabilities. This type of analysis is still incorporated in discovery programs in combination with ML techniques.

The availability of a pharmacophore model, in theory, allows a medicinal chemist to design new analogs, and if those analogs adhere to the pharmacophore model, theoretically they will be active. These models are the mainstay of medicinal chemistry and allow one to maintain and optimize potency and efficacy by understanding the structural features involved in binding to the macromolecular target. At the same time, optimizing physical and pharmaceutical properties while maintaining potency is possible by modifying the scaffold or other areas of the molecule that are independent of the pharmacophore.

3.3.5 Structure-Based Drug Design Requires Knowledge of the Structure of the Drug Target

The preceding section describes how pharmacophores are developed from knowledge of ligands and active and inactive molecules and how they are used in ligand-based drug design. However, with the advent of methods to determine the structures of macromolecules, such as X-ray crystallography, NMR, and cryo-electron

microscopy, structure-based drug design (SBDD) is now a common strategy. For example, a compound bound to its target can be visualized based on an X-ray crystal structure of the protein target that has been solved. Often, the structure is solved when the small molecule is bound to a macromolecular target, as this provides the most reliable data. However, sometimes only the macromolecule structure is determined and the small molecule is virtually docked into the site. In other cases, a structure of a related protein is known, and it can be used to make a model of the protein of interest (a **homology model**) for docking purposes. As mentioned in the previous chapter, AlphaFold has become an important tool for the prediction of three-dimensional protein structures. It allows the computation of a protein structure from the amino acid sequence and provides millions of structures that were previously unavailable for use in SBDD.

The availability of a three-dimensional view of the small molecule bound to the macromolecular target allows for the observation of specific binding interactions so that it is immediately apparent which functional groups are necessary for binding and identification of a pharmacophore is expedited. SBDD also provides an opportunity to incorporate additional elements into a pharmacophore model based on the structure of the target and to design improved compounds that may not have been apparent without the availability of the structure. SBDD is frequently used for the design of enzyme inhibitors or other proteins that are amenable to structure elucidation, and a number of such examples are described in later chapters.

An early example of the power of SBDD was the development of the human immunodeficiency virus (HIV) protease inhibitors in the 1990s. The antiviral drugs for acquired immunodeficiency syndrome (AIDS) and HIV infection that target HIV protease were all optimized by SBDD. X-ray structures of the protein itself and many co-crystal structures of inhibitors bound to the protein were solved. For example, the HIV protease inhibitor ritonavir was optimized from lead A-74704, shown in **Figure 3.14a**.

Compound A-74704 exhibited an IC_{50} of 3 nM against the protease enzyme, and an X-ray structure was solved with it bound to the active site. Two particular binding interactions observed were hydrogen bonds from two NH groups of the molecule to backbone carbonyls of glycines on the enzyme; however, this interaction did not have the optimal geometry for a hydrogen bond. Analogs, including A-77003 (**Figure 3.14b**), were designed that contained an additional central carbon

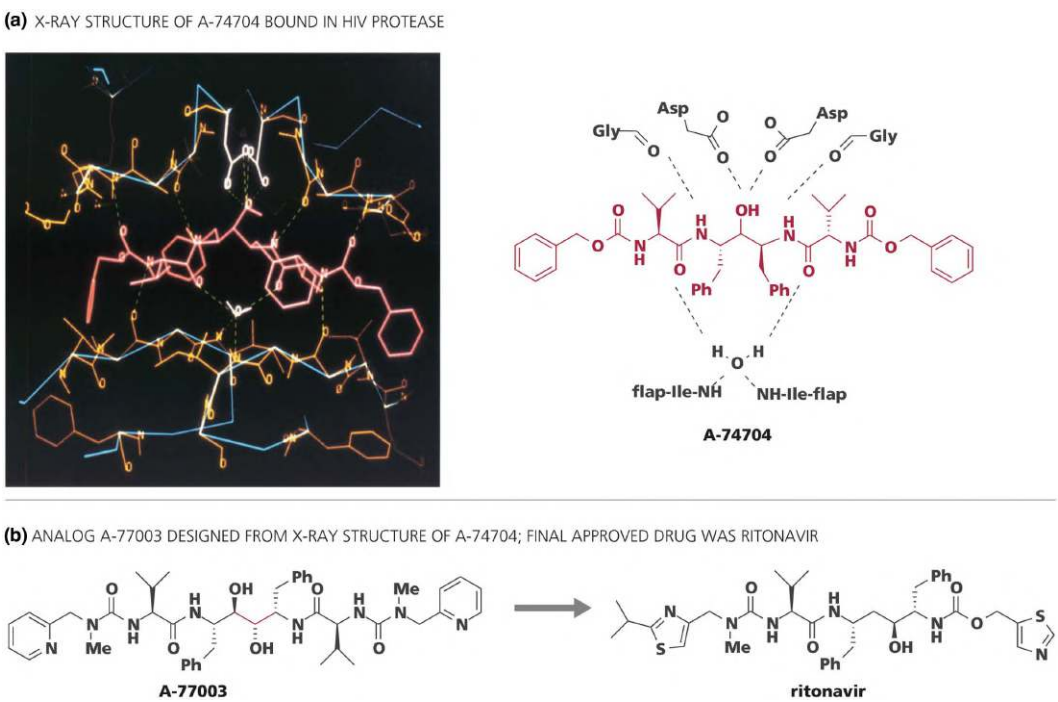


Figure 3.14 Structure-based drug design of HIV protease inhibitor ritonavir: (a) X-ray structure of A-74704 (pink) bound in the enzyme active site (side chains of enzyme in yellow, outline of backbone in blue, aspartic acid side chains and structural water in white), and depiction of A-74704 hydrogen bond interactions (dashed lines) with protein and structural water (shown in black). (b) Structures of A-77003 optimized for binding by inserting additional substituted carbon atom (in red) and ritonavir, the drug developed from these leads.

Source: From Greer, J.; Erickson, J. W.; Baldwin, J. J.; and Varney, M.D. (1994). *J Med Chem* 37:1,035–1,054. With permission from the American Chemical Society.

and allowed a more optimal geometry for H-bonding; the new analogs were 10–50 times more potent. The drug ultimately developed from this series was the AIDS/HIV drug ritonavir. The availability of the X-ray structure allowed for the visualization of the unoptimized H-bonds and the realization that by changing their geometry, potency could be improved.

The HIV protease inhibitor example is an early example of the power of structure-based drug design, which is now a common practice. This is especially true for the design of enzyme inhibitors, since these drug targets are more amenable to crystallization than membrane-bound receptors.

3.4 SUMMARY

While Chapter 2 described hit and lead identification strategies, this chapter highlighted some of the properties that are important during the lead optimization process. Biological, physical, and pharmaceutical properties are all evaluated and optimized in order to identify a viable drug candidate starting from very simple primary assays, advancing to more complex secondary assays, and eventually moving to *in vivo* assays. Biological properties, such as potency, efficacy, and selectivity of a small molecule, are optimized by taking into account its specific binding interactions with its protein target and by the development of a pharmacophore model.

Ligand- and structure-based designs are modern tools that take into account the three-dimensional conformation of the small molecule and are used to design and optimize improved analogs. Ligand-based design is used to develop models around compounds that bind to the same binding site of a target; the structure of the target does not need to be known. In contrast, structure-based drug design requires knowledge of the macromolecular target and, therefore, is not applicable to all drug discovery projects. Modern methods using artificial intelligence and machine learning are also being applied to develop pharmacophore models.

Knowledge of the pharmacophore is necessary for the lead optimization process. It provides insight into which elements are required for binding to the biological target and, therefore, which can be further optimized to improve potency and selectivity. At the same time, knowledge of the pharmacophore provides information about which aspects of the molecule can be modified in order to optimize physical and pharmaceutical properties without affecting target binding.

CASE STUDY 1: EXAMPLE OF THE USE OF A PHARMACOPHORE MODEL IN VIRTUAL SCREENING

Scientists at Schering-Plough reported a novel series of compounds that act to block the cannabinoid (CB1) receptor as potential anti-obesity drugs. One drug that was previously developed in this family is rimonabant, although it was removed from the market due to undesirable side effects. A combination of pharmacophore modeling and virtual screening was used to find lead compounds. The pharmacophore model was based on known ligands for the CB1 receptor and included eight compounds (**Figure 3.15**) with reported activity as antagonists (blockers) and inverse agonists (compounds that bind but with activity opposite to the natural ligand).

After a three-dimensional model of each compound was generated and low-energy conformers of each were calculated, pharmacophore models were generated by use of the Catalyst/HipHop commercial pharmacophore-generating program. Hydrogen bond acceptor (HBA), aromatic ring (AR), and hydrophobic (HPO) features were incorporated, and eight different pharmacophore models were generated. Each model was mapped with all eight compounds, and the one pharmacophore that best fit rimonabant was chosen (**Figure 3.16**).

Virtual screening of an in-house virtual library of approximately 500,000 compounds was carried out and provided ~22,000 hits that adhered to the pharmacophore model. Those hits were filtered for molecular weight and sample availability, as well as other criteria, and the remaining compounds met at least three of the following criteria: MW <500; calculated log*P* <5; number of HBA <10; number of HBD <3; number of rotatable bonds <10. This filtering step narrowed the pool of hits to 7,247 compounds. Further filtering was carried out by developing a model of probability of activity, leading to 2,100 compounds. Those compounds were then clustered on the

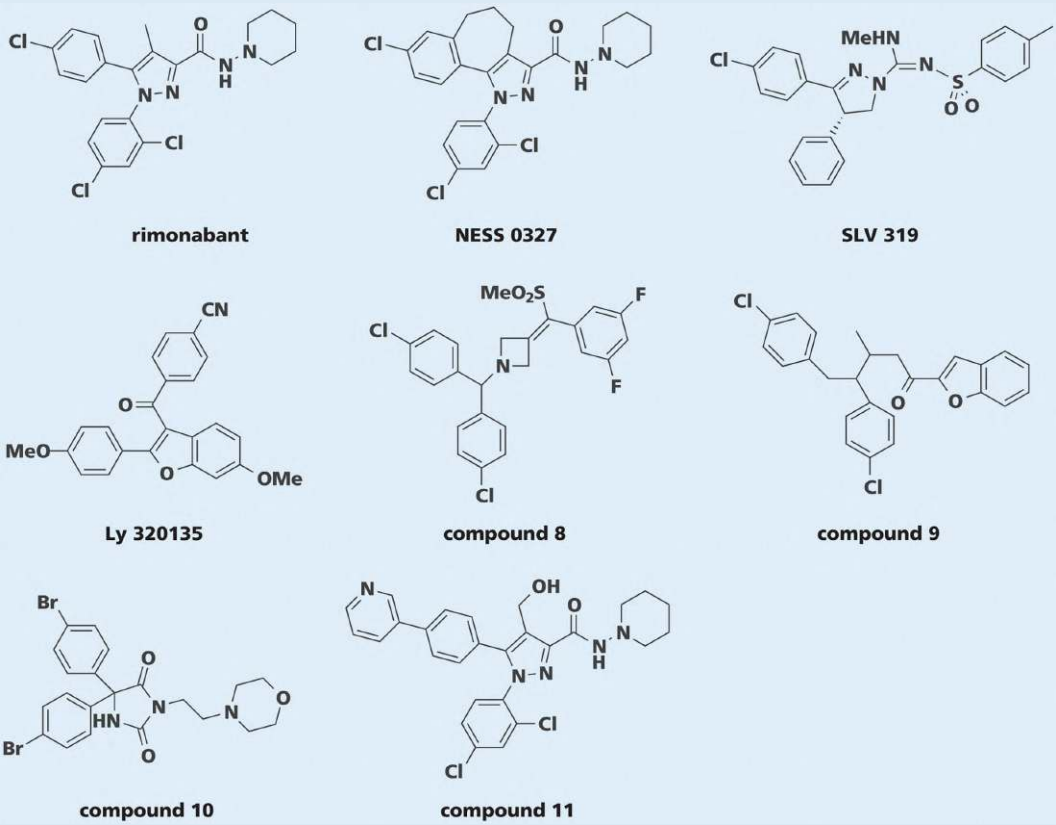


Figure 3.15 **Structures of compounds used to develop the pharmacophore model for CB1.**

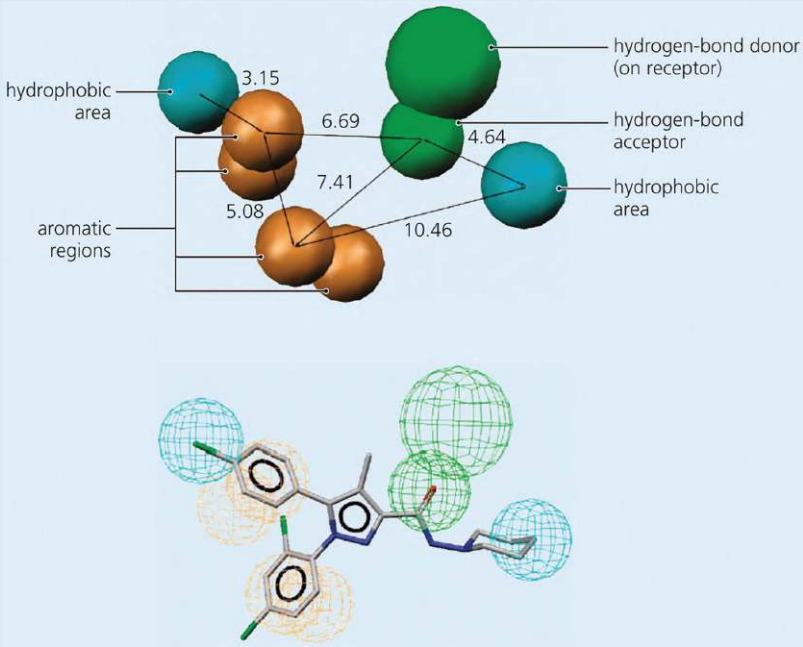


Figure 3.16 CB1 pharmacophore model. Brown spheres = aromatic ring (*AR*); cyan spheres = hydrophobic area (*HPO*); and small green sphere = hydrogen bond acceptor (*HBA*). Large green sphere = hydrogen bond donor (*HBD*) on receptor. Lower part of the figure shows rimonabant superimposed onto pharmacophore model.

Source: From Wang, H.; Duffy, R. A.; Boykow, G. C.; et al. (2008). *J Med Chem* 51:2,439–2,446. With permission from the American Chemical Society.

basis of similarity to afford 420 groups, and the compound with the highest predicted activity in each group was selected for testing in a CB1 binding assay. Five of the 420 compounds exhibited greater than 50% inhibition at 0.1 μM. The most promising was the azetidinone, compound E, which exhibited selectivity for CB1 over CB2 and was a novel scaffold for cannabinoid antagonists (**Figure 3.17**).

This example illustrates a number of techniques described in this and previous chapters, including the development of a pharmacophore model and virtual screening based on that model. Using these methods, the authors were able to screen a virtual library of 500,000 compounds to identify a novel scaffold as a lead for the development of a potential treatment for obesity.

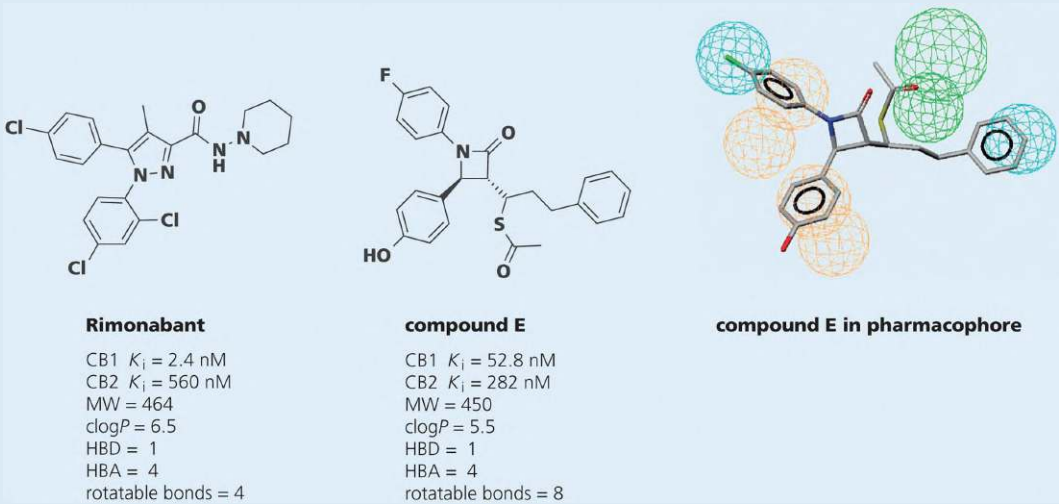


Figure 3.17 Structures of Rimonabant and novel CB1 antagonist compound E, with compound E superimposed on the pharmacophore model.

Source: From Wang, H.; Duffy, R. A.; Boykow, G. C.; et al. (2008). *J Med Chem* 51:2,439–2,446. With permission from the American Chemical Society.

CASE STUDY 2: EXAMPLE OF THE USE OF STRUCTURE-BASED DRUG DESIGN

Phosphodiesterases (PDEs) are a class of enzymes that hydrolyze the phosphodiester bonds of nucleotides, usually cGMP or cAMP. Numerous isoforms exist, and several are clinically significant, with PDE3 inhibitors having activity as cardiovascular drugs, and the erectile dysfunction drugs, such as sildenafil, acting at PDE5. The isoform PDE10A is expressed in the striatum of the brain, and inhibition is an approach for the treatment of schizophrenia, based on the goal of raising cGMP levels in the striatum. Authors in this study used SBDD to identify a clinical candidate with high selectivity for inhibition of PDE10A. The internal compound library at Pfizer was screened for PDE10A inhibitors, and triarylimidazole **10** was identified as a hit. The compound exhibited an IC_{50} of 35 nM and >100× selectivity over other PDEs. Co-crystallization of the lead with PDE10A suggested two reasons for the selectivity. First, all PDEs have a conserved glutamine Gln-726. In all previously solved co-structures, this side chain formed a hydrogen bond to inhibitors. However, in this case, rather than forming an H-bond, the amine of Gln-726 is positioned above the aromatic ring of the imidazole, forming a cation–π interaction. It is hypothesized that this difference in the specific binding interaction contributes to selectivity. Second, all PDEs have a hydrophobic cleft near the entrance to the active site, but the PDE10A cleft is deeper and has a specific lipophilic pocket known as the selectivity pocket. In this case, the thiophene of **10** occupies that unique selectivity pocket of PDE10A (Figure 3.18).

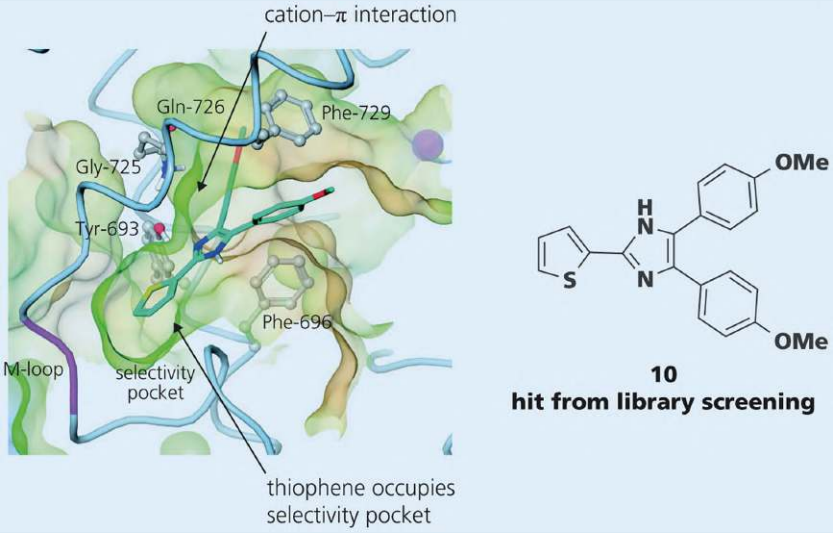


Figure 3.18 Library hit compound 10 co-crystallized with enzyme PDE10A. Important interactions with enzyme include Gln-726 to imidazole cation–π interaction and hydrophobic interactions from occupation of the selectivity pocket with the thiophene moiety.

Source: From Verhoest, P. R.; Chapin, D. S.; Corman, M.; et al. (2009). *J Med Chem* 52:5,188–5,196. With permission from the American Chemical Society.

The nature of the selectivity pocket is unique in PDE10A due to the presence of a small amino acid (glycine) at position 725, next to the conserved Gln-726. All other PDEs have a larger amino acid at that same position, which blocks this pocket. Consequently, this pocket is deeper in PDE10A than in other PDEs. Another observation in the X-ray structure highlighted additional differences: PDE10A contains a Tyr at position 693 in that pocket that may be able to form a hydrogen bond to an inhibitor, while other subtypes do not offer the same H-bonding opportunities.

On the basis of this structural analysis of the hit–enzyme co-crystal, the authors pursued a second round of screening based on compounds that could fit into that

lipophilic selectivity pocket and also hydrogen-bond to Tyr-693 in the pocket. Other factors for choosing compounds included the early incorporation of druglike physicochemical properties, such as MW <400, moderate lipophilicity, and minimal number of hydrogen bond donors. These criteria were applied in order to identify compounds likely to cross the lipophilic blood–brain barrier. On the basis of these factors, pyrazole **11** was identified as a hit, and analog **12** was the optimized lead. X-ray analysis of a co-structure confirmed that the quinoline fits into the selectivity pocket and a hydrogen bond between the quinoline nitrogen and Tyr-693 is observed. This compound has greater selectivity for the PDE10A isoform of the enzyme compared to the thiophene hit compound **10**, presumably due to the larger quinoline group that extends farther into the pocket than the original thiophene, as well as the ability to form a hydrogen bond there. Figure 3.19 shows the X-ray co-structure, with **11** in orange and **12** in green. Key binding interactions of **12** are the hydrogen bond to Tyr-693 and the cation–π interaction between Gln-726 and the aryl ring. The pyridyl nitrogen also forms a hydrogen bond to a water molecule (top of Figure 3.19) that binds to the backbone of the enzyme. Inhibition is improved from IC_{50} = 35 nM for the original compound **10** to IC_{50} = 0.42 nM, with >1,000× selectivity over other PDEs for **12**.

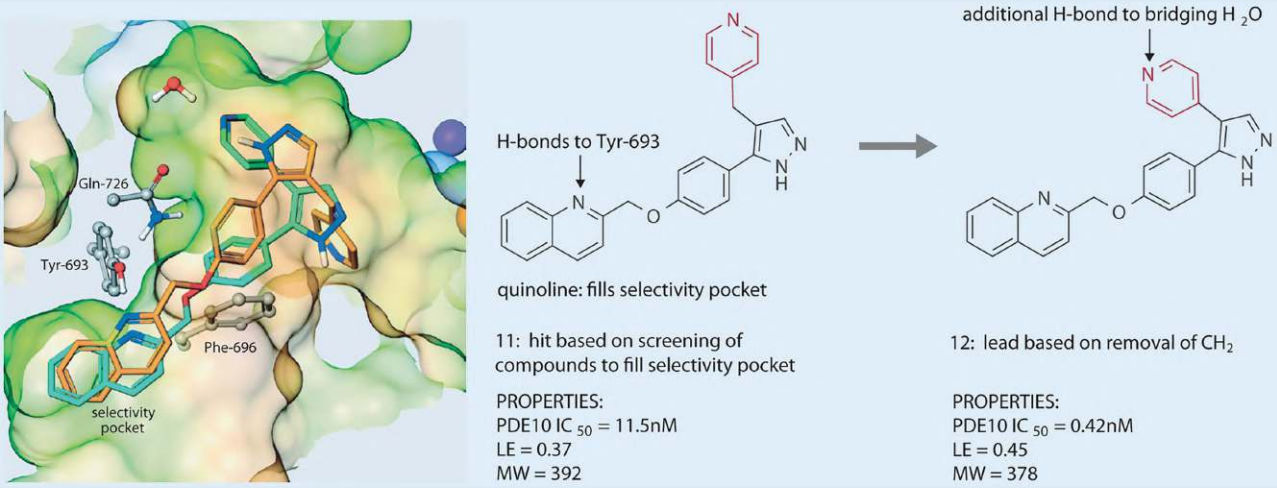


Figure 3.19 X-ray co-structure of 11 and 12 bound to PDE10 showing specific interactions. Structures of hit compound **11**, discovered by virtual screening for compounds that fill the selectivity pocket and make an H-bond to Tyr-693, and optimized lead compound **12**. X-ray structures of **11** (orange) and **12** (green) show occupation of the selectivity pocket by the quinoline, as well as an H-bond to Tyr-693. Also evident are cation–π interaction between Gln-726 and the aryl ring. In the case of **12**, an H-bond between the terminal pyridine and a structural water is also observed.

Source: From Verhoest, P. R.; Chapin, D. S.; Corman, M.; et al. (2009). *J Med Chem* 52:5,188–5,196. With permission from the American Chemical Society.

Although **12** was promising on several levels, such as selectivity and binding affinity, results *in vivo* were not adequate. One secondary assay measured an increase in levels of cyclic guanosine monophosphate (cGMP) as a downstream result of PDE10A inhibition. While **12** did raise cGMP levels in the striatum of mice by 350% at a dose of 10 mg/kg, it had only modest brain penetration, which meant that it was not reaching its target *in vivo*.

SAR analysis was combined with SBDD for further optimization. For example, replacing the nitrogen of the pyridine with carbon, as well as incorporating other aryl groups, led to compounds with reduced activity, confirming the importance of the H-bond to the bridging water molecule. Other heterocycles as substitutes for the pyrazole were investigated, including an isoxazole and a pyrazine, which both had decreased potency. Replacement of the quinoline weakened the important H-bond

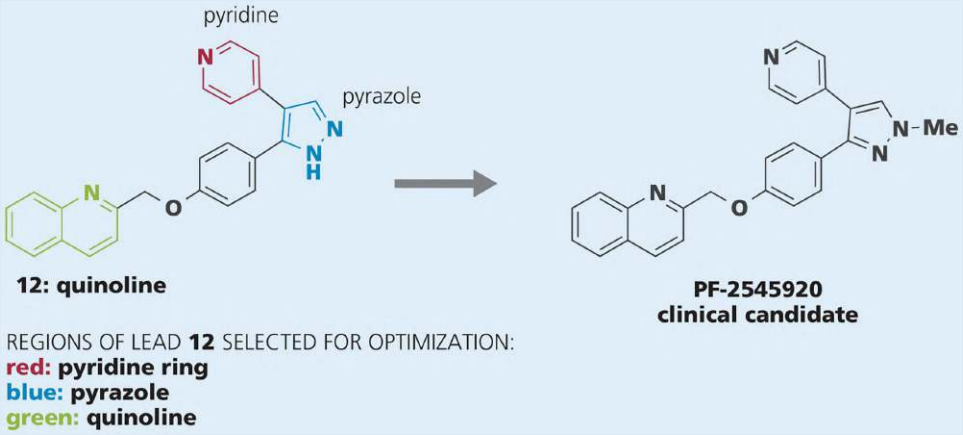


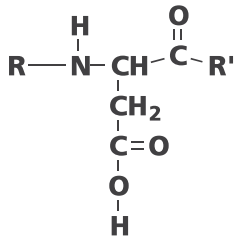
Figure 3.20 SAR analysis of lead compound 12, showing optimization to give the clinical candidate PF-2545920.

to Tyr-693 and led to diminished potency. One successful strategy replaced the NH of the pyrazole ring with a methyl group, thereby removing one H-bond donor and increasing lipophilicity; this led to the clinical candidate PF-2545920 (Figure 3.20).

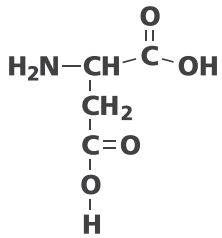
Overall, PF-2545920 (mardepodect) maintained the best *in vivo* properties and became the first PDE10A inhibitor to enter clinical trials as a potential treatment for schizophrenia, and it progressed through Phase II trials.

3.5 REVIEW QUESTIONS

1. Aspartic acid is an amino acid found in many proteins. It is drawn in what follows as a fragment of a generic protein. Identify five molecular interactions aspartic acid could participate in with a drug molecule. Identify the specific atom(s) that would participate in each.



2. If aspartic acid itself were a drug candidate, explain whether desolvation would be an energetically favorable or unfavorable process.



3. List some of the properties that were incorporated in early, historical QSAR models and some of the parameters used to reflect that property.

4. Explain the difference between receptor-excluded volume and receptor-essential area.

5. Define the concept of a pharmacophore.

6. What is required for structure-based drug design?

7. What properties are optimized during the lead optimization process?

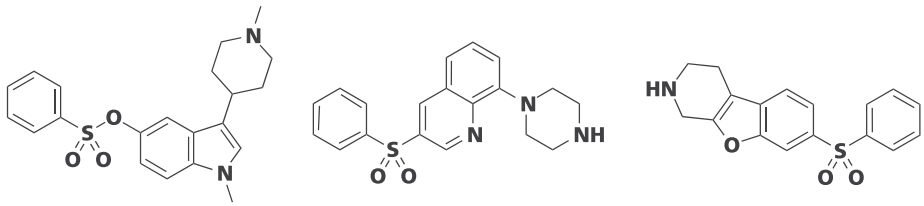
8. How does the equilibrium dissociation constant of a drug–target interaction relate to the Gibbs free energy of binding?

9. What is a eudismic ratio?

10. Explain why hydrophobic interactions, despite being weak interactions, are strong contributors to drug–protein binding.

3.6 APPLICATION QUESTIONS

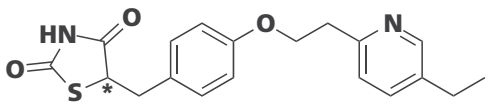
1. The three molecules shown here are all potent antagonists at the 5-HT_6 receptor.



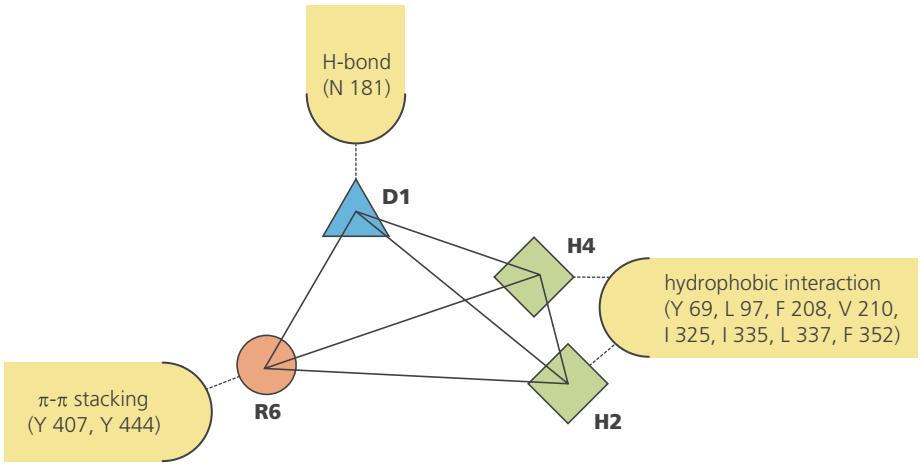
Design a pharmacophore for 5-HT_6 antagonism that is consistent with the three structures, and map the pharmacophore onto each compound (that is, identify which part of each molecule represents each pharmacophoric element).

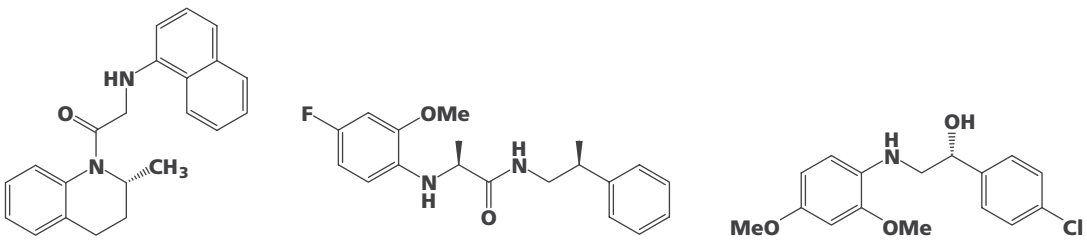
2. Using the three 5-HT_6 antagonists in Application Question 1, design three different molecules that test your pharmacophore hypothesis. Explain your design and what you will conclude if testing shows your new compound to (a) maintain or improve activity or (b) lose activity.

3. Pioglitazone is a drug used to treat diabetes. It contains an asymmetric center on the thiadiazinedione ring (indicated by *), yet the racemate is used clinically. Research why the racemic mixture of pioglitazone is used rather than an enantiomerically pure form.

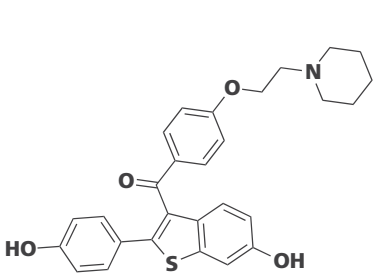


4. A pharmacophore model for monoamine oxidase A inhibitors is shown. It consists of an H-bond donor (D1, blue), two hydrophobic groups (H2, H4, green), and a π–π stacking interaction (R6, orange). The three compounds shown beneath adhere to this pharmacophore. Map the features of this pharmacophore onto each structure. (Figure adapted from Shelke, S. M.; Bhosale, S. H.; Dash, R. C.; et al. [2001]. *BioorgMed Chem Lett* 21:2,419–2,424. With permission from Elsevier.)

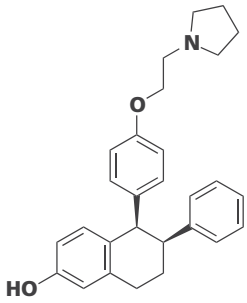




5. Structures of the selective estrogen receptor modulators raloxifene and lasofoxifene are shown. Draw a pharmacophore that is consistent with both structures. Identify the features of each molecule that correspond to each pharmacophoric element (map or label it).

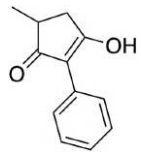


raloxifene

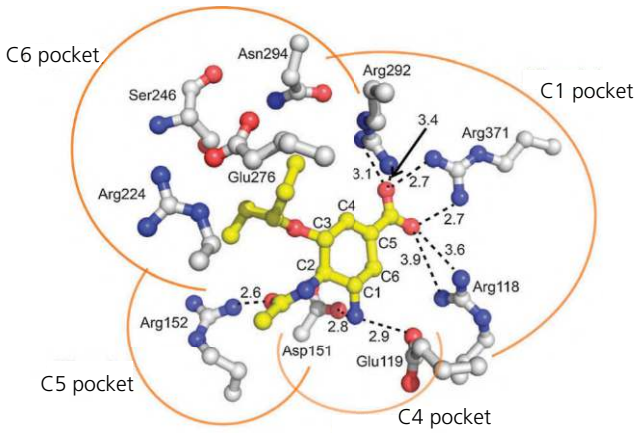


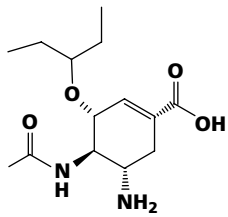
lasofoxifene

6. Identify five types of molecular interactions that the given fragment can make with a protein target. Identify which atoms or moieties are involved in each type of interactions.



7. Research the structure of biotin and the molecular interactions it makes with streptavidin or avidin. What is its binding constant (K_d)? What are biotin and streptavidin or avidin binding partners used for?
8. Research a drug that works by making a covalent bond to its target. What is the structure of the drug, and what bond is made?
9. Do a literature search to find an example of the use of machine learning in QSAR. Explain what type of ML was used, and identify at least two compounds with predicted activity.
10. Structure-based drug design was used in the design of the antiviral drug oseltamivir, which inhibits the enzyme neuraminidase (see Chapter 11). The drug structure (yellow) and an X-ray crystal structure of the drug bound in the active site of the enzyme are shown. (Figure from Naumov, P.; Yasuda, N.; Rabeh, W. M.; and Bernstein, J. [2013]. Chem Commun 49:1,948–1,950. With permission from Royal Society of Chemistry.)





oseltamivir

a. Identify a key ionic binding interaction.
b. Identify a hydrogen bonding interaction.
c. Based on the view of the active site, in what pocket is the alkyl ether? Looking at the X-ray structure, might it be wise to change the ether from 3-pentyl to a larger alkyl group? Why or why not?

FURTHER READING

Binding Interactions

Bissantz C, Kuhn B & Stahl M (2010) A medicinal chemist's guide to molecular interactions. *J Med Chem* 53:5061–5084.

Ferenczy GG & Keseru GM (2010) Thermodynamics guided lead discovery and optimization. *Drug Discov Today* 15:919–932.

Friere E (2008) Do enthalpy and entropy distinguish first in class from best in class? *Drug Discov Today* 13:869–874.

Smith AJT, Zhang X, Leach AG & Houk KN (2009) Beyond picomolar affinities: quantitative aspects of noncovalent and covalent binding of drugs to proteins. *J Med Chem* 52:225–233.

Quantitative Structure–Activity Relationship

Cherkasov A, Muratov EN, Fourches D et al. (2014) QSAR modeling: where have you been? Where are you going to? *J Med Chem* 47:4977–5010.

Dara S, Dhamercherla S, Jadav SS, Babu CHM & Ahsan MJ (2022) Machine learning in drug discovery: a review. *Artif Intell Rev* 55:1947–1999.

Free SM & Wilson JW (1964) A mathematical contribution to structure activity studies. *J Med Chem* 7:395–399.

Hansch C & Fujita T (1964) ρ-σ-π analysis: a method for the correlation of biological activity and chemical structure. *J Am Chem Soc* 86:1616–1626.

Topliss JG (1972) Utilization of operational schemes for analog synthesis in drug design. *J Med Chem* 15:1006–1111.

Pharmacophore

Guarav A, Yadav MR, Giridhar R et al. (2011) 3D-QSAR studies of 4-quinolone derivatives as high-affinity ligands at the benzodiazepine site of brain $GABA_A$ receptors. *Med Chem Res* 20:192–199.

Huang Q, He X, Ma C et al. (2000) Pharmacophore/receptor models for $GABA_A$/Bz subtypes (α1 β3γ2, α5β3γ2, and α6β3γ2) via a comprehensive ligand-mapping approach. *J Med Chem* 43:71–95.

Lager E, Andersson P, Nilsson J et al. (2006) 4-Quinolone derivatives: high-affinity ligands at the benzodiazepine site of brain GABAA receptors. Synthesis, pharmacology, and pharmacophore modeling. *J Med Chem* 49:2526–2533.

Leach AR, Gillet VJ, Lewis RA & Taylor R (2010) Three-dimensional pharmacophore methods in drug discovery. *J Med Chem* 53:539–558.

Marshall G, Barry CD, Bosshard HE et al. (1979) The conformation parameter in drug design: the active analog approach. *ACS Symp Ser: Comput-Assisted Drug Des* 112:205–226.

Tebib S, Bourguignon J-J & Wermuth C-J (1987) The active analog approach applied to the pharmacophore identification of benzodiazepine ligands. *J Comput-Aided Drug Des* 1:153–170.

Structure-Based Drug Discovery

Erickson J, Neidhart DJ & VanDrie J et al. (1990) Design, activity, and 2.8Å crystal structure of a C2 symmetric inhibitor complexed to HIV-1 protease. *Science* 249:527–533.

Lounnas V, Ritschel T, Kelder J et al. (2013) Current progress in structure-based rational drug design marks a new mindset in drug discovery. *Comput Struct Biotechnol J* 5:e20130211.

Visualizing Noncovalent Interactions

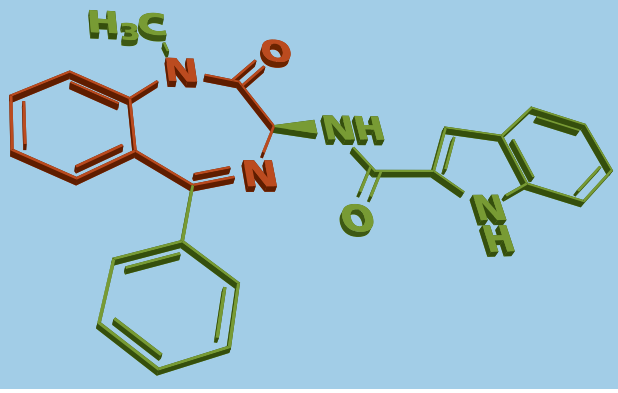
Contreras-Garcia J, Johnson ER, Keinan S et al. (2011) NCIPLOT: a program for plotting noncovalent interaction regions. *J Chem Theory Comput* 7:625–632.

Johnson ER, Keinan S, Mori-Sanchez P et al. (2010) Revealing noncovalent interactions. *J Am Chem Soc* 132:6498–6506.

Case Studies

Verhoest PR, Chapin DS, Corman M et al. (2009) Discovery of a novel class of phosphodiesterase 10A inhibitors and identification of clinical candidate 2-[4-(1-methyl-4-pyridin-4-yl-1 *H*-pyrazol-3-yl) phenoxymethyl]-quinoline (PF-2545920) for the treatment of schizophrenia. *J Med Chem* 52:5188–5196.

Wang H, Duffy RA, Boykow GC et al. (2008) Identification of novel cannabinoid CB1 receptor antagonists by using virtual screening with a pharmacophore model. *J Med Chem* 51:2439–2446.



Lead Optimization: Properties Optimized and Medicinal Chemistry Strategies

4

LEARNING OBJECTIVES

- Understand the specific properties that are targeted during lead optimization.
- Become familiar with lead optimization assays and their target values.
- Understand the various "rules" that have been developed to guide lead optimization.
- Learn multiple strategies that are applied during the lead optimization process.
- Recognize when specific medicinal chemistry strategies should be applied.

4.1 INTRODUCTION

The process of lead optimization to obtain a clinical candidate requires multiparameter optimization (MPO) of potency, selectivity, and physical and pharmaceutical properties, as well as safety. As described in Chapter 3, understanding the pharmacophore and optimization of binding is one key step in this process. It allows the medicinal chemist to maintain or improve potency and selectivity, and therefore safety, by focusing on the pharmacophoric elements that are essential for biological activity. At the same time, physical and pharmaceutical properties are optimized by maintaining the pharmacophoric features and modifying the portions of the molecule that are not essential to binding. These **druglike properties** are the necessary features for a compound to become a drug and include chemical and metabolic stability, solubility, lipophilicity, and safety, among others. The characteristics of a drug, including binding affinity, biological activity, and the aforementioned druglike properties, are interrelated, and modification of one often affects the others. The challenge for the medicinal chemist is to optimize all the properties despite their interdependency.

4.2 PROPERTIES EVALUATED DURING LEAD OPTIMIZATION

This section describes some of the important biological, physical, and pharmaceutical properties that are evaluated and refined during the lead optimization process. On the basis of target goals for each of these parameters, medicinal chemists employ a number of strategies during the iterative optimization process to eventually generate a clinical candidate. These properties and the target values are described in the

DOI: 10.1201/ 9781003465607-5

following sections, and some specific medicinal chemistry strategies follow. While some of these are categorized according to the property improved, in many cases, more than one property is affected, and therefore this division should not be considered absolute. Finally, these target values should not be considered inviolate, as there are always examples of approved drugs that do not meet specific criteria.

4.2.1 Biological Activity Is Optimized as a Component of the Lead Optimization Process

Chapter 3 described the molecular interactions that mediate binding between a small molecule and its biological target and how they contribute to its observed activity. However, the biological activity of a small molecule is based not only on its binding to the target but also on other attributes. The data generated in a biological assay reflect the aggregated effects of all these traits. The lead optimization process uses data from increasingly stringent biological assays to inform the iterative rounds of design and synthesis, with the understanding that target binding, combined with a number of other properties, contributes to the observed effects.

In a typical drug discovery program, primary assays that measure potency, and sometimes selectivity, toward a biological target are first utilized. These *in vitro* assays are relatively simple; they may be the same assays used during the original hit identification process, and many times (but not always) they measure the activity of the molecule toward a specific biological target. As described in Chapter 2, they can range from assays that quantify the activity of a compound on an isolated protein to a phenotypic screen that measures the effect of a small molecule on a particular cellular phenomenon. Secondary assays are then incorporated to measure potency and efficacy of the molecule in a more complex testing system that often reflects a downstream effect of interference with the primary target. For example, if a primary assay measures the ability of a compound to inhibit an enzyme, a secondary assay could measure the amount of the same enzyme's substrate and product in cells. An effective enzyme inhibitor in such a secondary assay would result in accumulated substrate and reduced product levels compared to controls. These assays are more rigorous than primary assays since they require that the small molecule permeate the cell, be stable, reach the specific target, and interact with it, as does the natural substrate. Furthermore, the molecule should not interfere with other biochemical processes that might also be ongoing in the cell. Finally, *in vivo* assays are utilized to evaluate the effects of the compound in animals. *In vivo* assays are considerably more rigorous and relevant compared to primary and secondary assays. The use of a whole animal requires that the compound be absorbed, not be rapidly metabolized and eliminated, reach the site of action at appropriate and sustained concentrations, and interact with the specific target to produce a measurable effect that is relevant to the disease, without causing significant side effects. Throughout the progression from primary and secondary *in vitro* to *in vivo* assays, the criteria for advancement become more stringent, and the compounds become further optimized. The three common components evaluated in optimization of biological activity are potency, efficacy, and selectivity.

Potency is the concentration of a compound that is required to achieve the desired effect. In contrast, **affinity** of a molecule, usually measured as K_d or K_i, reflects its ability to bind to its target. As was defined in Chapter 3, K_d is the dissociation constant of a drug from a drug–target complex. The inhibition constant, K_i, for an enzyme (E) is the equilibrium dissociation constant of an enzyme inhibitor complex, where $K_i = [E][I]/[E - I]$. Potency of hits and leads (usually measured as IC_{50} or EC_{50}) at the beginning of the lead optimization process can be in the 1–10 μM range in primary assays, and it is not unusual for hits and early leads to be inactive in secondary and *in vivo* assays. Values targeted during the lead optimization stage and common for clinical candidates are <10 nM in an isolated biochemical assay (primary assay), <100 nM in a cellular assay (secondary assay), and ≤50 mg/kg in an animal model of disease (*in vivo* assay). One common value calculated when a compound is active in *in vivo* models is the **minimum effective dose** (MED), which is the lowest dose that produces a statistically significant desirable effect. $\mathbf{ED_{50}}$ refers to the dose that produces 50% of the maximal response.

Efficacy refers to the degree of effect of the compound and is usually measured as a percentage compared to a standard (100%) and technically refers to receptor agonists. However, it is often applied to other systems. For example, complete inhibition of the activity of an enzyme (no substrate is turned over) is considered 100% efficacious. The target efficacy value is determined by the specific disease: In cancer drug discovery, 100% inhibition of tumor growth is required, while in some neuropsychiatric disorders, complete inhibition of neurotransmitter signaling may not be desirable, and target efficacies may be only 50%.

Selectivity for a specific target, with minimal activity against other biological targets, is an important criterion for the safety of the drug candidate. Usually, compounds will be tested for selectivity against related targets or those that are essential to avoid in a primary assay. Primary, secondary, and *in vivo* assays are also used to detect off-target effects, and ratios of >10- to 100-fold in favor of the desired target are expected.

4.2.2 Physical and Pharmaceutical Properties Are Optimized in Parallel with Biological Properties

The biological activity of a small molecule is based not only on its binding to the target but also on its physical properties. One barrier to drugs reaching their target is the cell membrane: either cellular membranes they must traverse to reach intracellular targets or those that the compounds must pass through in the gastrointestinal (GI) tract to enter circulation. Therefore, appropriate physical properties based on lipophilicity and aqueous solubility are required for membrane permeability. Furthermore, if a molecule is highly polar, it will prefer to be in an aqueous solution and, therefore, will never reach nonpolar sites of biological targets, since the desolvation penalty will be high. These properties, in addition to biological activity, must also be optimized during the lead optimization process and are often specifically measured in primary and/or secondary assays. The method by which drugs are eventually administered to humans (oral, intravenous, transdermal, etc.) may require specific physical properties and also must be considered during this stage.

Lipophilicity refers to the affinity of a compound for a lipophilic environment. It relates to a compound's ability to move from an aqueous medium, when first administered, through lipophilic membranes, to be absorbed into the aqueous environment of blood circulation, and then through additional membranes, in order to reach its protein target. While an optimized lead needs to have sufficient lipophilicity, compounds that are highly lipophilic also tend to have higher levels of target promiscuity, leading to safety issues. Compounds that are too hydrophilic (low lipophilicity) will not cross the biological membranes.

Lipophilicity can be measured experimentally by a number of methods, but one of the most common is by partitioning a compound between 1-octanol and water and measuring the concentration of the drug in each layer. As described in Chapter 3, the partition coefficient, P, is the concentration of compound in 1-octanol divided by the concentration in water, and log$P$ is the most commonly reported value. A more useful value for acidic or basic compounds is log$D$, the logarithm of the apparent partition coefficient at a specified pH, which uses a buffered system as the aqueous layer. This important variation eliminates the effects of the test compound's acidity or basicity on the aqueous solution and therefore takes into account the ionization state of the molecule being measured. Log$D$ values always include the pH at which they were measured: For example, log $D_{7.4}$ indicates that the measurement was taken at pH 7.4. As an alternative to experimental methods, calculated log$P$ values (referred to according to the specific program used, such as cLog$P$, aLog$P$, or xLog$P$) are available in chemical drawing programs and in cheminformatics and computational chemistry software.

In order to optimize lipophilicity, a careful analysis of the functional groups in a hit or lead molecule is required. For compounds that contain a large number of polar functionalities (low log$P$), the elimination, replacement, or incorporation of additional lipophilic groups (such as halogens or alkyl groups) is considered. For those compounds whose log$P$ value is too high, the introduction of heteroatoms is a frequent strategy.

Lipophilicity is a particularly important characteristic of drugs that passively cross the blood–brain barrier (BBB). Rather than the typical log*P* value of 2–4 for orally available drugs, molecules that cross the BBB fall in a slightly higher range of 3–5.

Solubility is another important characteristic of drugs, and it is influenced by the lipophilicity of the molecule. Drugs will be in an aqueous environment in the body and should have acceptable aqueous solubility. This characteristic, like lipophilicity, is related to a compound's ability to be absorbed into the bloodstream and cross biological membranes. A compound must be in solution in order to be permeable. In addition, the aqueous solubility of a compound will also allow it to be easily formulated for delivery to animals and humans such that consistent levels are absorbed. Solubility depends on the specific functional groups within the molecule as well as the physical characteristics of the solid. Features that increase aqueous solubility include the presence of hydrogen bonding and ionizable groups. Both kinetic and thermodynamic solubilities are usually measured at physiological pH, and typical values are in the range of 5–100 μM.

Permeability is a property that relies on both the solubility and the lipophilicity of a compound. In order to reach the desired target organ, compounds need to be lipophilic enough to permeate a cell membrane but hydrophilic enough to be soluble. The permeability of a lead compound can be measured in primary assays, such as a parallel artificial membrane permeability assay (PAMPA) that measures the rate of passage through an artificial membrane. A common secondary assay is the Caco-2 assay, which measures the rate of transit through a monolayer of epithelial cells and takes into account not only passive but also active transport. MDCK2 permeability assays can detect not only permeability through a cell but also the potential for compounds to be actively effluxed (resulting in overall poor permeability). Permeability is measured as a rate and is characterized by comparison to reference standards. High or medium permeability is preferred, and having highly permeable compounds with low or moderate solubility can sometimes still be acceptable. For instance, all other things being equal, to achieve a projected dose of 1 mg/kg, a compound with medium permeability will require aqueous solubility of about 50 μM, whereas a compound with high permeability might require solubility of only about 10 μM.

Pharmaceutical properties are evaluated by absorption, distribution, metabolism, and elimination (ADME) studies, which are discussed in more detail in Chapter 5. While these evaluations become increasingly detailed as a lead is optimized to a clinical candidate, early indications of some of these properties are an important component of the lead optimization phase of drug discovery. For example, a prediction of *in vivo* metabolic stability and clearance may come from an *in vitro* assay based on incubation of a compound with liver microsomes that contain active metabolic enzymes. Although a compound may exhibit suitable chemical stability in pH 7.4 buffer, it may be rapidly degraded in the presence of metabolic enzymes in liver microsomes and cleared, thereby resulting in very low levels in circulation.

Another frequently desirable pharmaceutical property is oral bioavailability, which is measured in early *in vivo* models. **Oral bioavailability** is the concentration of drug in the blood after oral dosing, as compared to the concentration after intravenous dosing. High oral bioavailability allows a low dose to be administered, thereby reducing the risk of off-target effects that might occur if high doses of compound were required. A low dose also translates to other benefits, such as simpler formulation development, easier dosing regimen for the patient, and ultimately, lower cost of the drug substance. A highly permeable, soluble, and metabolically stable compound will most likely have high oral bioavailability.

These biological, physical, and pharmaceutical properties are interrelated. For example, a compound that is highly potent in secondary cell–based biological assays will, by necessity, be permeable and soluble. Conversely, if a compound is highly potent in primary biochemical assays but not in secondary cellular biological assays, poor physical and pharmaceutical properties may be the culprit. While each specific project or disease therapy has specific requirements for biological, physical, and pharmaceutical properties, **Table 4.1** summarizes the characteristics and values that are generally targeted. Most of these properties are measured routinely throughout the

Table 4.1 Summary of Some Properties and Measurements Evaluated during Lead Optimization

Biological Properties	
Affinity	• Tendency of a small molecule to associate with its target measured in primary binding assays
	• Target values: K_d or K_i <10 nM
	• Impacts potency, safety, and eventual dose of final drug substance
Potency	• Concentration or dose at which a compound elicits a specific effect
	• Measured in primary, secondary, and *in vivo* assays
	• Target values for primary and secondary assays: IC_{50} and EC_{50} <10 nM in isolated system; <1 μM in cell-based system[a]
	• Target values for *in vivo* assays: ED_{50} or EC_{50} <50 mg/kg
	• Minimum effective dose (MED) <50 mg/kg (lowest dose possible desirable)
	• Related to eventual dose and safety of compounds in humans
Efficacy	• Degree of a given biological response (technically refers to receptor agonists)
	• Measured in primary, secondary, and *in vivo* assays reported as a percentage, usually compared to a control
	• Target value determined by specific indication and protein target
	• Reflects the physiological effect of the compound
Selectivity	• Ratio of magnitude of desired effect compared to undesired effect
	• Usually calculated as a ratio of potency and/or affinity in primary, secondary, and *in vivo* assays
	• Target values usually >10 but may be as low as 2 for some life-threatening diseases
	• Reflects the safety of the drug
Physical and Pharmaceutical Properties	
Aqueous Solubility	• Amount of compound that dissolves in an aqueous solution, under either kinetic or thermodynamic conditions
	• Measured in primary (kinetic) and secondary (thermodynamic) assays; values of >10 μM targeted
	• Related to a compound's ability to cross biological membranes, to be absorbed after administration, and to be formulated for delivery to animals and humans
Lipophilicity	• Affinity of a compound for lipophilic environment
	• Measured in primary and secondary assays by various measurements, such as partition coefficients
	• Target values are 2–4 for log*P* or log*D*
	• Reflects the ability of a compound to cross biological membranes, its safety, and its ability to be formulated
Permeability	• Rate at which a compound transits across a membrane
	• Measured in various primary or secondary assays, such as PAMPA (primary) or Caco-2 and MDCK (secondary)
	• High or medium permeability, with low efflux usually preferred
	• Reflects ability of the drug to reach the bloodstream or the target organ
Metabolic Stability	• Ability of a compound to withstand metabolic transformation by liver enzymes
	• Measured in primary or secondary assays, such as liver microsomes, to detect half-life of compound
	• $T_{1/2}$ >60 min ideal, although very long $T_{1/2}$ may be unsafe; often converted to CLint (intrinsic clearance), which measures intrinsic ability of the liver to metabolize the small molecule
	• Contributes to drug achieving concentrations at the site of action that are essential for efficacy at the lowest dose possible

(*Continued*)

Table 4.1 Summary of Some Properties and Measurements Evaluated during Lead Optimization (*Continued*)

Oral Bioavailability	• Concentration of drug in the blood via oral route, compared to concentration achieved with intravenous administration
	• Measured in animal models; values of >50% preferred
	• When preferred human dosing regimen via oral route, high oral bioavailability used to ensure consistent drug concentrations achieved at lowest possible doses
State (solid, liquid)	• Physical state of compound
	• For solids, measured via melting point or differential scanning calorimetry (DSC)
	• Crystalline, non-hygroscopic, stable solid samples preferred
	• Will affect solubility, stability, and reproducibility of drug sample

[a]IC_{50} = inhibitory concentration, concentration that reduces response by 50%; EC_{50} = effective concentration, concentration that causes 50% of maximal response; ED_{50} = effective dose, median dose that causes 50% of maximal response.

iterative design, synthesis, and assay process, with *in vivo* assays being implemented on fewer compounds at the latter stages.

4.3 OPTIMIZATION OF LEAD PROPERTIES

Modifying the lead structure optimizes the specific properties outlined in **Table 4.1**. The data generated in the assays described earlier are analyzed, and new compounds are designed that are predicted to have improved properties. In the following sections, tools and strategies that have been developed to aid the medicinal chemist are described.

4.3.1 Rules and Metrics Have Been Developed to Aid the Lead Optimization Process

A number of metrics have been developed and are used by medicinal chemists as a guide during the optimization process. By applying these metrics and rules, specific properties (biological, physical, and pharmaceutical) are identified that need to be improved during the iterative design, synthesis, and testing cycle, and newly designed analogs are designed based on them. These metrics and rules are not steadfast, but they provide a tool that helps simplify a complex, multidimensional set of interdependent properties.

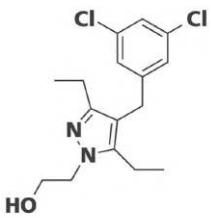
Several different metrics are used in order to assess the value of specific atoms or fragments of a compound. These include ligand efficiency and lipophilic ligand efficiency, which evaluate atoms' or a fragment's contribution to affinity and potency and to lipophilicity, respectively. **Ligand efficiency** (LE) assesses and compares compounds on the basis of how efficiently they bind (affinity). This metric is a ratio of the molecular weight of a molecule compared to its affinity and is defined as the free energy of binding of a ligand averaged for each non-hydrogen or heavy atom (*HAC* = heavy atom count). The term used for the binding parameter is ideally K_d or K_i, but EC_{50} or IC_{50} values can also be substituted. If it is assumed that an oral drug candidate should have MW <500 and K_d or K_i (or EC_{50} or IC_{50}) <10 nM, then an LE of ≥0.3 is considered a desirable number.

$$\mathrm{LE} = -\Delta G / \mathrm{HAC} = -RT \ln K_d \left(\text{or } \mathrm{IC_{50}} \text{or } \mathrm{EC_{50}}\right) / \mathrm{HAC} \qquad 4.1$$

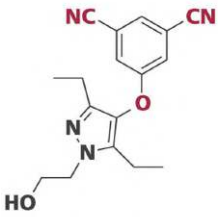
For $T = 300$ K, this equation can be simplified to:

$$\mathrm{LE} = -1.37 / \mathrm{HAC} \times \log K_d \left(\text{or } \mathrm{pIC_{50}}\right) \left(\text{where } \mathrm{pIC_{50}} = -\log \mathrm{IC_{50}}\right)$$

This concept allows for the evaluation and comparison of hits and leads on the basis of their efficiency of binding, and it sets a goal (≥0.3) or direction (greater LE) for



1
IC_{50} = 0.66 μM
clog*P* = 4.3
LLE = 1.9



2 (lersivirine)
IC_{50} = 0.12 μM
clog*P* = 2.1
LLE = 4.9

Figure 4.1 Optimization of LLE in a series of reverse transcriptase inhibitors. Compound 1 exhibited an LLE = 1.9; to improve the LLE, the linkage to the aryl ring was replaced with the more polar oxygen atom, and the substituents on aryl ring were replaced with nitriles to give compound 2 (changes are shown in red). These changes decreased clog*P*, increased potency, and improved the LLE to 4.8.

medicinal chemistry efforts. *Ligand efficiency* refers to the whole molecule, whereas the idea of *group efficiency* (GE) refers to specific functional groups within the molecule. Within a series of similar analogs, the value of individual groups can be assessed by comparing changes in free energy of binding that occur when a functional group is changed. In this case, HAC refers to the heavy atom count of the individual group.

As an example of how to use ligand efficiency metrics, consider two compounds, A and B, that both have IC_{50} values of 10 nM. Compound A contains 30 non-hydrogen atoms, but compound B has 40 non-hydrogen atoms. Using the equations LE = $-RT$ $\ln IC_{50}$/HAC and $\ln IC_{50}$ = −18.4, the LE values (at 298 K) for compounds A and B would be calculated as follows:

Compound A

$$\mathrm{LE} = -(0.00198\mathrm{kcal/deg\cdot mol})(298\mathrm{K})(-18.4)/30\mathrm{atoms} = 0.36$$

Compound B

$$\mathrm{LE} = -(0.00198\mathrm{kcal/deg\cdot mol})(298\mathrm{K})(-18.4)/40\mathrm{atoms} = 0.27$$

In this example, even though both compounds have identical potency, the efficiency with which they achieve that potency differs. Compound A requires fewer atoms (and has a higher LE, 0.36) than compound B to exhibit the same potency. During the optimization process, efficiency is considered carefully and monitored, and modifications that improve potency and maintain or even improve efficiency are preferred.

Ligand lipophilic efficiency (LLE), or lipophilic ligand efficiency (LipE), is a gauge of how the lipophilicity of a compound contributes to its potency. It is defined as log of potency minus log*P* (either experimental or calculated, clog*P*) and can be used as an assessment of the molecule's affinity for its target vs. its general affinity for a lipophilic environment.

$$\mathrm{LLE} = \mathrm{LipE} = \mathrm{p}IC_{50} - \log P \qquad 4.2$$

For compounds that will be administered *via* oral dosing, an LLE of 5–7 is targeted. An example of the use of LLE in the optimization of a series of HIV nonnucleoside reverse transcriptase inhibitors is shown in Figure 4.1. An early compound (**1**) exhibited IC_{50} of 0.66 μM (pIC_{50} = 6.2) and had a clog*P* of 4.3, and therefore an LLE of 1.9. Substituting cyano groups for the chlorines and an oxygen for the methylene led to **2** (lersivirine), with a lower clog*P* of 2.1 and greater potency (IC_{50} = 0.12 μM; pIC_{50} = 6.9), and therefore an improved LLE of 4.8. Maintaining a high LE and optimal LLE often helps guide the optimization process.

In addition to metrics that focus on optimizing each fragment for contributions to potency and lipophilicity, rules have been developed for the optimization of the pharmaceutical property of oral availability. In order for a drug to be orally available, it needs to survive the acidic environment of the gastrointestinal tract, be able to penetrate membranes (either through passive or active transport), and survive the metabolic machinery of the liver. Some of the general factors that can decrease oral bioavailability include export by transporter enzymes and first-pass metabolism.

In 1997, Christopher Lipinski and co-workers at Pfizer proposed a set of rules that predicted the likelihood of a compound being orally bioavailable. The **Lipinski rules**, also known as the **rule of five** (Ro5), are based on the common properties of compounds that had progressed to Phase II clinical trials and were dosed orally. A modified version, proposed by Miles Congreve and co-workers at Astex, applies to the smaller-molecular-weight hits identified in fragment-based drug discovery and is called the **rule of three**. Another set of rules was proposed in 2002 by Daniel Veber and co-workers at GlaxoSmithKline to predict oral availability. The **Veber rules** incorporate molecular flexibility and polar surface area and take into account the unexpected oral bioavailability of certain larger-molecular-weight drugs. Both sets of rules are

Table 4.2 Rules for Optimization of Oral Bioavailability

Lipinski Rules: Rule of Five[a]
• A compound must have no more than five hydrogen bond donors (counted by the number of NH and OH groups).
• There must be no more than ten hydrogen bond acceptors (counted by the number of N and O atoms).
• The molecular weight must be under 500 Da (daltons).
• The log*P* (partition coefficient; lipophilicity) must be no more than 5.
Variation of Lipinski Rules: Rule of Three[b]
• A compound should have no more than three hydrogen bond donors. There should be no more than three hydrogen bond acceptors.
• The molecular weight should be under 300 Da.
• The log*P* should be no more than 3.
Veber rules
• A compound should have no more than ten rotatable bonds.
• The hydrogen bond count (donors and acceptors) should be no more than 12, or polar surface area (PSA) should be no more than 140 $(Å)^2$.

[a]All the components of the rules are multiples of five.
[b]Used most often in fragment-based screening and hit triage.

summarized in **Table 4.2**. While these correlations are termed *rules*, there are a significant number of drugs that deviate from them, such as Lipitor, at one time the world's best-selling drug. More recently, the validity of these rules has come into question as more and more orally bioavailable drugs that fall outside (and some, very far outside) these parameters are developed. For example, a more recent analysis (1998–2019) of orally bioavailable drugs revealed statistically significant deviations from the Ro5 and Veber parameters in all metrics except hydrogen bond donor count. Practically, these rules should be considered guidelines and not hard-and-fast requirements.

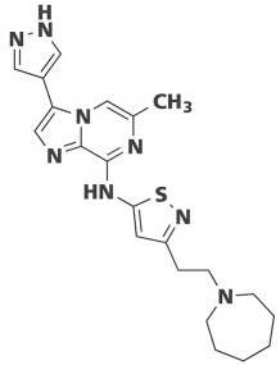
Drug discovery scientists continue to develop an understanding of what physical properties of drugs are most important to consider during the optimization process and which accurately predict oral bioavailability and drug development success.

4.4 SPECIFIC STRATEGIES TO ADDRESS DEFICIENCIES IN LEAD STRUCTURES

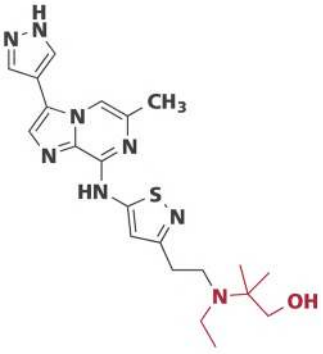
The overall goal of lead optimization is to take a promising lead and progress it to a druglike clinical candidate. While the rules and metrics described earlier can be used as target values, the following sections describe specific medicinal chemical strategies that are applied to improve physical and pharmaceutical properties of drug candidates. The improvements in properties typically also result in improvements to potency, and therefore, these strategies cannot be considered in isolation.

4.4.1 Lipophilicity Is Optimized so That a Drug Can Reach Its Target

Lipophilicity is one of the most important physical characteristics of a drug candidate. If a drug is too lipophilic, it will not be soluble, or it may compartmentalize into lipid membranes, never reaching its target tissue or organ. Lipophilic compounds are known to bind indiscriminately to proteins, another mechanism that prevents them from reaching their targets. Furthermore, highly lipophilic compounds are often more rapidly metabolized than less lipophilic compounds, decreasing their bioavailability. Conversely, compounds that are not sufficiently lipophilic may not cross cell membranes and may be so highly solvated that binding to the target protein is not a favorable process. Therefore, the appropriate balance between hydrophobicity and hydrophilicity is essential.



Aurora kinase inhibitor



SCH 1473759

Figure 4.2 Optimization of solubility by incorporating polar functional groups. To improve the aqueous solubility of an aurora kinase inhibitor, the azepine was replaced with an acyclic alcohol moiety (changes are shown in red), resulting in SCH 1473759 that displayed 500× greater solubility than the original aurora kinase inhibitor.

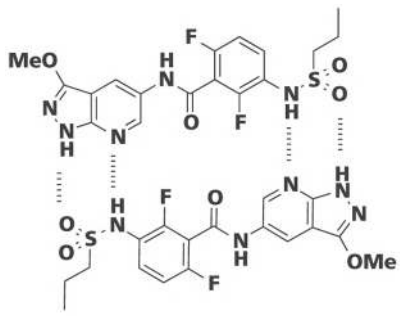
4.4.2 Leads Need to Be Optimized for Solubility in Aqueous Media

Small-molecule binding sites on a protein target are often lipophilic, and therefore, potent hits and leads may likewise be lipophilic, and they often suffer from poor or low aqueous solubility. However, high solubility is a prerequisite for both *in vitro* and *in vivo* activity, so improving solubility may be a significant effort during the optimization stage. The two most important factors affecting aqueous solubility in the context of medicinal chemistry are the polarity of the molecule and the crystal lattice energy of the solid crystal form; thus, the most commonly applied strategies to improve solubility are the introduction of polar solubilizing groups and the disruption of crystal packing.

Adding aqueous solubilizing groups, a common method to improve solubility, typically involves adding polar functional groups, such as alcohols, amines, or acids, to sites of the molecule that are not required for binding to the target. In addition to adding hydrophilicity, some of these functional groups also offer the opportunity to make salts of the molecule that may further increase solubility. In the example shown in **Figure 4.2**, replacing the azepine ring (seven-membered ring) in an aurora kinase inhibitor with the more polar acyclic alcohol (shown in red) improved solubility by almost 500-fold. An X-ray co-crystal structure of SCH 1473759 bound to aurora A kinase showed that the alcohol group extended out of the binding pocket and into a solvent-accessible site.

A second strategy to improve solubility is disruption of crystal packing. This approach is most often applied when there is evidence of high crystal lattice energy, such as a high melting point and/or small-molecule X-ray crystal structure showing numerous intermolecular interactions. By the incorporation of functional groups that disrupt those interactions, lower melting points and higher water solubility are obtained.

This strategy was successfully applied to a series of inhibitors of B-Raf^{V600E} kinase that showed potential as anti-cancer agents. The small-molecule X-ray structure of the original compound showed a head-to-tail arrangement and four hydrogen bonds between two molecules (**Figure 4.3**). In addition, π stacking and hydrogen bonds with molecules above and below the plane of the dimers were also observed. These interactions combined to afford high crystal packing lattice energy. Replacement of one aryl fluorine with a larger chlorine atom interfered with the dimer interface, and substitution of the methoxy group with a cyclopropane disrupted the π stacking between planes of the dimer (analog B; changes in red). By designing a molecule that disrupted some of these interactions, a >10-fold improvement in aqueous solubility was observed, along with >50° decrease in melting point.



B-Raf^{V600E} kinase inhibitor A: strong crystal lattice
AQUEOUS SOLUBILITY:
9 μg/mL at pH 6.5, 7.4



analog B: disrupted crystal lattice
AQUEOUS SOLUBILITY:
127 μg/mL at pH 6.5, 7.4

Figure 4.3 Optimization of solubility by disruption of crystal lattice. Inhibitor A self-associates through hydrogen bonds (shown as dashed lines) and π stacking between planes. The incorporation of the larger chlorine atom and a cyclopropyl group (shown in red) disrupts hydrogen bonding and π stacking and leads to higher aqueous solubility in analog B.

4.4.3 Optimization of Metabolic Stability Is a Necessary Component of Lead Optimization

One of the most frequent causes of poor bioavailability is rapid metabolism of the lead molecule, typically through oxidative processes. The process of metabolism will be described fully in Chapter 5; however, a common strategy used early in the optimization process is

to identify the metabolically labile positions and minimize their potential to be oxidized by metabolic enzymes. Standard strategies include reducing lipophilicity (more lipophilic compounds are more likely to be metabolized), inverting stereochemistry (metabolic processes may be stereoselective), modifying the steric environment of the metabolic site, manipulating electronic characteristics, and introducing conformational constraints. For example, the introduction of fluorine is a common method to reduce metabolism, since its small size may not drastically change the conformation of the active molecule. In some examples, direct replacement of a metabolically labile hydrogen atom or methyl group with a fluorine atom or trifluoromethyl group, respectively, physically blocks metabolism. Alternatively, the introduction of a fluorine atom or another electron-withdrawing group on an aromatic ring deactivates the entire ring toward oxidative metabolism and will often improve metabolic stability.

The introduction of deuterium in order to block metabolism is a similar strategy. In these cases, due to the deuterium isotope effect, cleavage of a C-D bond may be ten times slower than cleavage of the corresponding C-H bond, slowing drug metabolism and increasing metabolic stability. Recently, the first deuterium-containing drug was approved. Deutetrabenazine replaces six hydrogen atoms in a previously known drug, tetrabenazine, with deuterium atoms. This change results in significantly longer half-life of the drug and its active metabolites.

Increasing steric hindrance at or near the site of metabolism is another frequently applied approach. In the example shown in **Figure 4.4**, the *N*-methylpiperazine (red) of a potential allergy medication was found to be rapidly metabolized, particularly through demethylation. The sterically encumbered and conformationally constrained bicyclic piperazines are typically more stable to metabolism, and when these moieties were incorporated into the parent molecule, stability toward metabolism was significantly improved. Within the same series, reducing the lipophilicity by replacement of the piperazine with a smaller aminopyrrolidine or aminoazetidine also produced a stabilizing effect, as measured by the percent of compound remaining after incubation with rat liver microsomes (RLM).

4.4.4 Targeting Strategies and the Use of Prodrugs to Improve Drug Properties

When physical properties of a drug are non-optimal, a prodrug strategy can sometimes be applied to solve the problem. A **prodrug** is a molecule, typically inactive, that when metabolized within the body affords the active drug molecule. This approach is

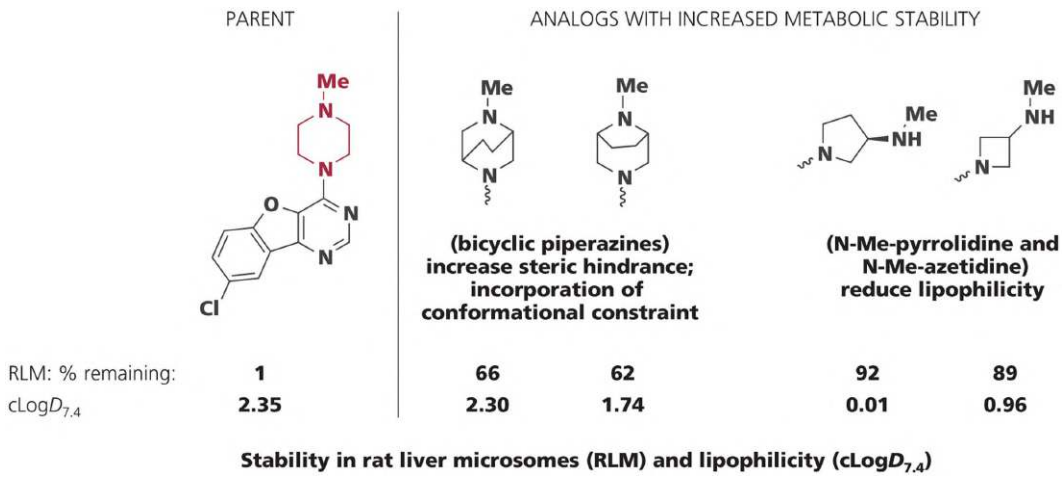


Figure 4.4 Optimization of metabolic stability by increasing steric bulk, introducing conformational constraint, and modifying lipophilicity. The parent compound (left) is rapidly metabolized at the *N*-methylpiperazine group (shown in red), with only 1% remaining after incubation with rat liver microsomes (RLM). Replacement of the piperazine with conformationally constrained and more sterically hindered bicyclic piperazines, or decreasing ring size (to five- or four-membered rings) to reduce lipophilicity (right), improved metabolic stability, as evidenced by the increased % remaining.

used to solve issues of poor permeability and/or solubility and thereby improve oral bioavailability. However, it can also be applied to affect the distribution of the drug and improve safety, among other characteristics. It has been estimated that approximately 15% of all approved drugs are actually prodrugs. In general, however, prodrug strategies are applied when other strategies have been unsuccessful. The reasons for this are based on several factors, including a relatively poor understanding and predictability of the biological processes and phenomena that can be exploited for prodrug design and the additional complications during the drug development process that prodrugs can encounter due to the fact that two molecules, the prodrug and the active drug, need to be studied and characterized. A schematic of the use of a prodrug with a masked functional group is shown in **Figure 4.5**.

One of the functional groups most frequently masked in a prodrug is a carboxylic acid. The charged, polar nature of the acid can severely impede its permeability, thereby making these drug candidates ineffective, since they never reach their intended biological target. In these instances, the acid can be masked as a hydrolyzable group, such as an ester, amide, carbonate, or other acid derivative that is less polar and more permeable, and usually inactive. Once absorbed, the derivative is cleaved by physiological enzymes to reveal the active acid moiety and the by-product, which is often a biologically inert molecule.

The influenza drug oseltamivir is one example of a prodrug (**Figure 4.6a**). It is administered as an ethyl ester and, once absorbed, is cleaved by endogenous esterases to generate the carboxylic acid, the active molecule. By use of an ethyl ester prodrug rather than the acid, the oral bioavailability of the active molecule improves from 5% to over 80%. A second example is the anti-cancer drug irinotecan. The active drug SN-38 (**Figure 4.6b**) is highly insoluble and, therefore, is difficult to formulate. By the incorporation of a water-solubilizing prodrug, the dipiperadinocarbamate, the solubility and formulation problems were solved.

In the first two examples in **Figure 4.6**, the prodrug was designed for cleavage by nonspecific esterases or other physiological enzymes. However, an alternate strategy incorporates a prodrug moiety that is specifically recognized by transporters or enzymes in or near certain cells or organs in the body, thereby accumulating and/or releasing the active molecule only at the site of action. In this manner, the drug can be targeted and may be more efficacious and less toxic. This strategy has been most often applied to central nervous system (CNS) and cancer targets due to the unique features of the blood–brain barrier and the tumor environment, respectively. The drug levodopa, used for Parkinson's disease, is one such example (**Figure 4.6c**). It is a substrate for a brain-specific transporter called large neutral amino acid transporter (LAT1) and therefore accumulates in the brain. Once there, it is decarboxylated to dopamine, the active species. A more recent example of a prodrug is remdesivir, originally developed as an antiviral drug to treat hepatitis C, respiratory syncytial virus, and Ebola but also found to be effective against SARS CoV-2. Its development is described in Case Study 2 at the end of this chapter.

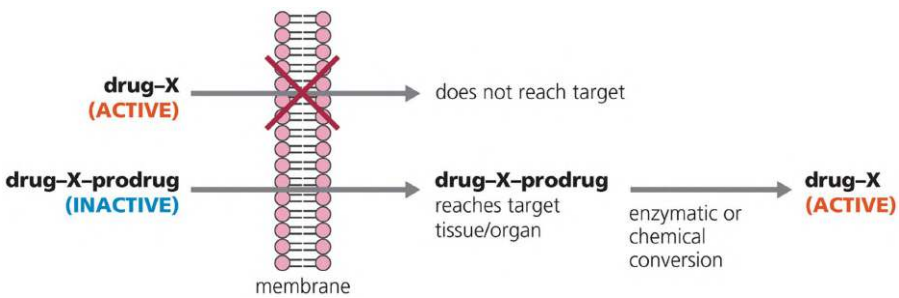


Figure 4.5 Schematic of prodrug crossing biological barrier that the active drug cannot. The active agent (drug-X, where *X* indicates a group required for activity but that also contributes to poor permeability) cannot cross the membrane, so no drug reaches its target. The prodrug, where *X* is masked by a readily removable functional group, can cross the membrane by virtue of improved properties. Biotransformation inside the barrier affords the active drug.

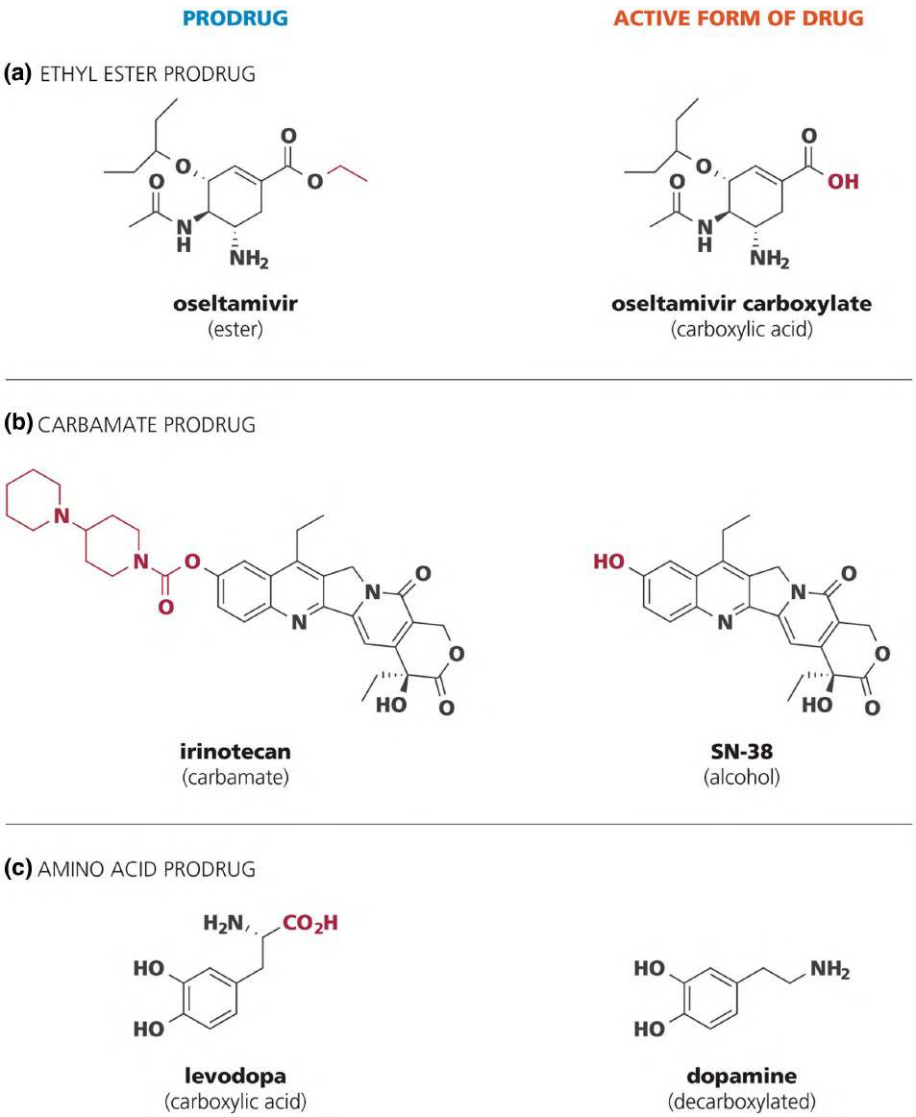


Figure 4.6 Prodrug examples: (a) The ethyl ester in oseltamivir is cleaved to generate the active carboxylic acid and provides a significant increase in bioavailability compared to the acid. (b) The prodrug carbamate of irinotecan shows improved solubility. (c) Levodopa is recognized by a CNS-specific transporter and accumulates in the brain, where it is decarboxylated to generate the active dopamine. Prodrug functional groups in red.

The anti-cancer drug capecitabine is a prodrug example that highlights how complex the route to the active moiety can be. Capecitabine undergoes three different biotransformations to generate 5-fluorouracil, an established anti-cancer agent. First, enzymes in the liver (CES1, CES2) cleave the carbamate ester moiety to generate the carbamic acid (not shown), which is rapidly decarboxylated to form 5'-deoxy-5-fluorocytidine (5'-dFCyd). Enzymes called cytidine deaminases (CDA) that are located primarily in the liver and tumor tissues perform the next bioconversion, which generates 5'-deoxy-5-fluorouridine (5'-dFUrd). Finally, an enzyme that is highly expressed in tumor cells, deoxythymidine phosphorylase (dThdPase), converts 5'-dFUrd to the active substance 5-fluorouracil (5-FU), as shown in Figure 4.7. In this way, the

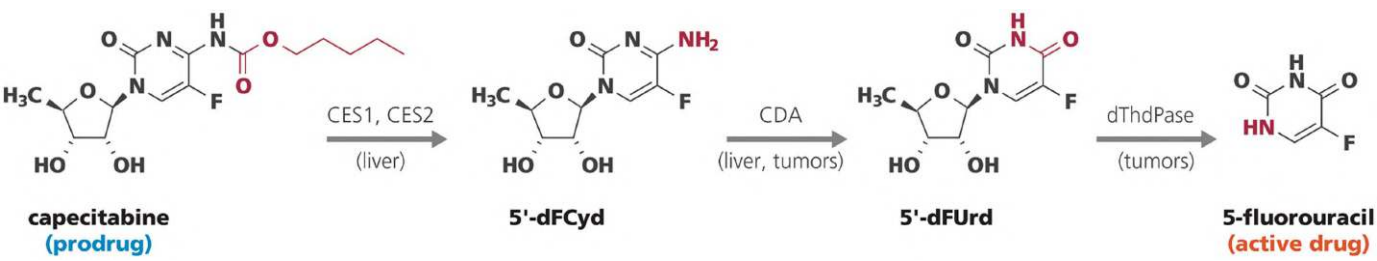


Figure 4.7 Transformation of capecitabine to the active form 5-fluorouracil. Hydrolysis of the carbamate by enzymes CES1 and CES2 and decarboxylation gives the amine 5'-dFCyd, oxidation by CDA gives 5'-dFUrd, and then cleavage of the base from the carbohydrate ring gives the active drug 5-fluorouracil.

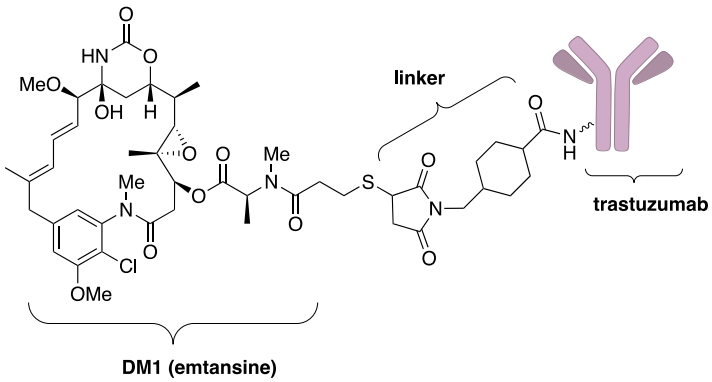


Figure 4.8 Structure of trastuzumab-emtansine, an antibody–drug conjugate used to treat cancer. A thioethyl analog of the natural product maytansine, DM1 (emtansine), is connected through a linker to the antibody (trastuzumab), which selectively delivers the drug to cancer cells.

cytotoxic 5-FU is released selectively in tumor cells, thereby significantly reducing its exposure to normal cells and improving the therapeutic window of this drug. In addition, the oral bioavailability of the molecule is improved and reaches almost 100%.

Another approach to targeting takes advantage of the exquisite specificity of antibodies to recognize certain molecules on the surface of certain cell types and has been successfully applied to cancer therapies. It allows extremely toxic molecules to be targeted to tumor cells while reducing exposure of normal cells, thereby improving the safety profile of the drug. As an example, maytansine is a natural product that is highly potent in anti-cancer assays. Unfortunately, the drug also exhibited unacceptable toxicity in clinical trials and was never advanced. To alleviate the toxicity, a thioethyl analog (DM1, or emtansine) was linked through the sulfur to an antibody called trastuzumab, which recognizes specific proteins on certain metastatic breast cancer cells and is an effective drug on its own (**Figure 4.8**). The antibody directs the cytotoxic agent to the HER2 receptor on metastatic breast cancer cells, where the conjugate is absorbed by the cancer cell. Proteolytic degradation of the antibody in the lysosome releases DM1, as an active metabolite conjugated to lysine. The conjugate, called trastuzumab-emtansine (Kadcyla), was approved in 2013 for the treatment of certain types of breast cancer. Many other **a**ntibody–**d**rug **c**onjugates (ADCs) are either approved or in clinical trials.

4.5 MEDICINAL CHEMISTRY STRATEGIES APPLIED DURING LEAD OPTIMIZATION

Optimizing the affinity of a small molecule against a biological target requires the optimization of the size, orientation, and polarity of the functional groups that interact at the binding site. As described in Chapter 3, a majority of the molecular interactions (for example, H-bonding, dipole–dipole, π–cation, and electrostatic interactions) that contribute to binding energy have a specific geometry that provides an optimal interaction. In addition, hydrophobic interactions are optimized by adjusting the size and the shape of the substituent. However, a drug candidate requires not only high affinity but also high potency and excellent physical and pharmaceutical properties. The lead optimization process requires all these parameters to be optimized in parallel, but improving one often results in worsening of another. A series of medicinal chemistry strategies have been developed that can be applied to address various liabilities of an early lead during the optimization process. They aim to maintain or improve potency while addressing weaknesses in specific physical or pharmaceutical properties.

4.5.1 Increasing Steric Bulk Affects Binding, Lipophilicity, and Metabolic Stability

The strategy of increasing steric bulk can be used to optimize affinity and potency, modify lipophilicity, and even improve metabolic stability. In the case of optimizing affinity and potency, these methods are most often used to take full advantage

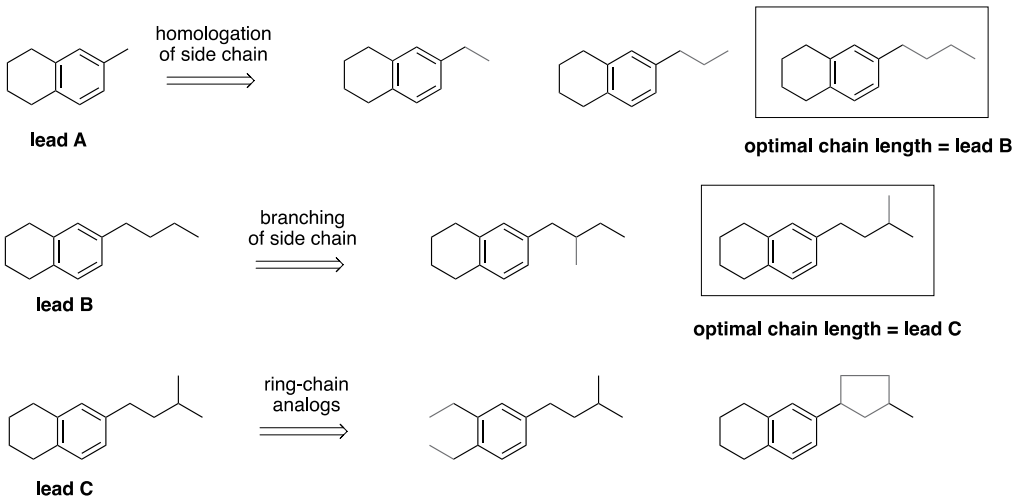


Figure 4.9 Theoretical example of lead modification, showing homologation, chain branching, and ring-chain transformations. Lead A is homologated from a methyl to ethyl, propyl, and butyl analogs (in red). If the homologated butyl analog (lead B, boxed) exhibits improved activity, in a second round of optimization, branch points are added to the butyl chain (red). If the branched analog lead C exhibits improved potency, further optimization may include either cleaving a ring (red, left) or forming a ring (red, right).

of hydrophobic sites on the protein target. Methods include homologation, chain branching, and ring-chain transformations, which are illustrated in Figure 4.9.

Homologation is the sequential addition or subtraction of methylene (CH_2) units to an alkyl chain. This change may affect binding to the target and will always have an effect on lipophilicity. In the hypothetical example shown in Figure 4.9, a methyl group in lead A is changed to an ethyl, propyl, or butyl group. If the butyl analog is the most active, it may be further modified. Increasing steric bulk can affect metabolic stability by hindering access to the metabolic site. The addition of alkyl branching points to an alkyl chain is termed **chain branching**. As with homologation, this will affect lipophilicity and may affect target binding. In the example shown starting from the newly optimized lead B, branching at different positions may provide a new compound with optimized activity (lead C). The third type of direct change to alkyl groups involves **ring-chain transformations**. In this type of transformation, the size of a ring may be increased or decreased, or a chain may be converted to a ring and vice versa. As shown in the hypothetical example in Figure 4.9, a ring in lead C may be broken open, or a chain may be closed to give a new ring. These modifications have an effect on the entropy of the system that will impact affinity, in addition to an effect on metabolic stability.

4.5.2 Bioisosteres Are Used as Substitutions for Specific Functional Groups

Bioisosteres (or **isosteres**) were defined by Alfred Burger in 1991 as "compounds or groups that possess near-equal molecular shapes and volumes, approximately the same distribution of electrons, and which exhibit similar physical properties." The strategy of replacing a functional group with a bioisostere is frequently applied when the goal is to maintain the affinity of a molecule but change its pharmaceutical properties. Over time, their importance in medicinal chemistry has grown, and their characterization has evolved. Today, the definition of a *bioisostere* is even broader, and replacing one functional group with a bioisostere may result in changes in size, electronics, lipophilicity, and other physical and pharmaceutical properties, as well as addressing intellectual property rights. However, the overall goal is to maintain the optimal properties of the original molecule but address its liabilities.

Classical isosteres were originally described in the early 1900s. According to Grimm's hydride displacement law, they are defined as "atoms that bond to one to four hydrogens to reach similarity to an inert gas." By this definition, O, NH, and CH_2 would be isosteric with each other. This view was later expanded by Erlenmeyer to include atoms, ions, and molecules with identical peripheral layers of electrons, rather than just number of bonds. In this case, Cl, SH, and PH_2 would be isosteric, since all have seven peripheral electrons. Classical bioisosteres have been categorized as monovalent, divalent, trivalent, tetravalent, or ring equivalents, as shown in Table 4.3.

Table 4.3 Examples of Classical Bioisosteres, Where Each Entry in a Row Is a Distinct Isostere of the Parent Group or Atom

	group/atom	bioisostere example
monovalent	—H —OH —CH_3	—F —D —NH_2 —SH —F —Cl —Br —OH —NH_2
divalent	—CH_2— C=C	—O— —S— —N(H)— C=O C=N C=S
trivalent	—C= —P=	—N= —As=
tetravalent		
ring equivalents		

One of the most powerful strategies using a classical bioisostere is the incorporation of fluorine to modify the properties of a lead molecule. Despite its small size, the introduction of a fluorine atom into a molecule can drastically change its potency, physical properties, and pharmaceutical properties. Recent advances in synthetic methods have made fluorine-containing drug candidates more accessible, and there has been a corresponding increase in the percentage of drug molecules containing fluorine. Currently, approximately 25% of drugs contain a fluorine atom. Fluorine is a relatively small atom, with a van der Waals radius of 1.47 Å. It is larger than hydrogen atom (1.2 Å) and smaller than oxygen and nitrogen (~1.5 Å) or carbon (1.7 Å) atoms. A trifluoromethyl group is estimated to be between the size of an isopropyl group and a *tert*-butyl group. The other defining characteristic of fluorine is its strong electronegativity. For example, the introduction of one fluorine atom into acetic acid reduces the pK_a from 4.7 (CH_3CO_2H) to 2.6 for CH_2FCO_2H; the pK_a of CF_3CO_2H is 0.23. In addition, fluorine can act as a hydrogen bond acceptor and, in many contexts, renders a compound more lipophilic than the parent. These characteristics can influence the properties of lead molecules in a variety of ways. The electronic effects of fluorine on neighboring acid or base moieties can tune and optimize their acidity or basicity. Due to its lipophilicity and ability to form hydrogen bonds and participate in dipole–dipole interactions, the introduction of a fluorine atom can improve affinity by making additional molecular interactions with the target protein, without significantly changing the steric environment. For example, in a series of thrombin inhibitors (**Figure 4.10**), the addition of a fluorine atom at the 4-position of a benzyl group (X) in the parent compound resulted in a greater than 5-fold increase in potency. An X-ray structure of the molecule bound to thrombin revealed that the fluorine atom was in close proximity to a backbone carbonyl unit, suggesting a new hydrogen bond between the small molecule and acidic H alpha to a carbonyl of the target (F···H-CαCO) and/or a new

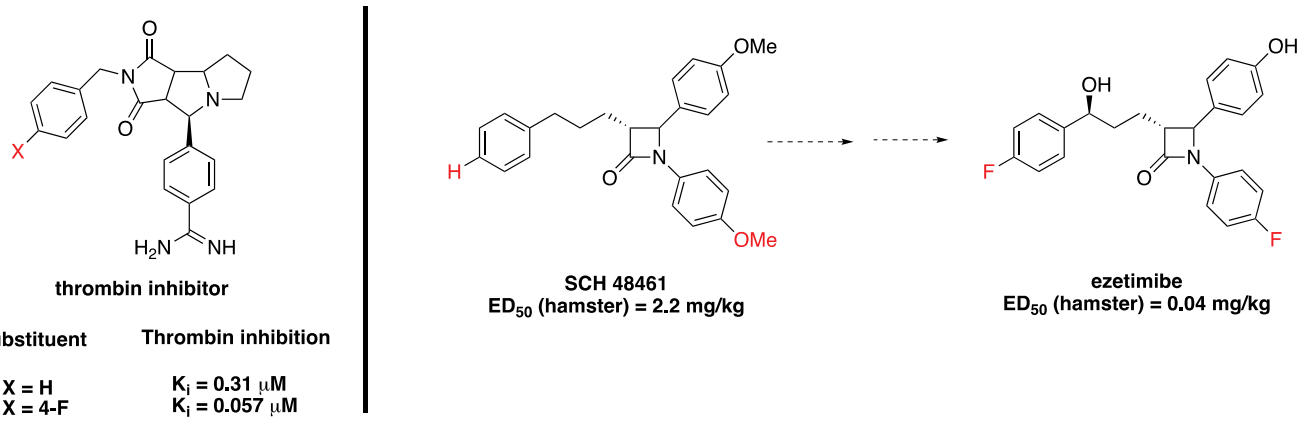


Figure 4.10 Improved binding (*left*) and metabolic stability (*right*) due to fluorine substitution. *Left*, substitution of a para fluorine substituent in a thrombin inhibitor improved binding to the enzyme by 5-fold (change shown in red). *Right*, an early lead, SCH 48461, was optimized in part by the addition of two fluorine atoms (shown in red) and led to ezetimibe, marketed as a cholesterol-lowering drug. The ED_{50} in a hamster model of disease improved from 2.2 to 0.04 mg/kg and reflected the improved metabolic stability imparted by the fluorine atoms.

dipole interaction between the fluorine atom and the backbone carbonyl carbon (F···C = O).

The incorporation of a fluorine atom can also have a profound effect on the physical and pharmaceutical properties of a lead molecule. Typically, lipophilicity increases with the addition of a fluorine atom, which often improves permeability and brain penetration. *Fluorination* is a strategy frequently employed to improve metabolic stability. By replacing a metabolically labile hydrogen (either aromatic or aliphatic) with a similarly sized fluorine atom, metabolism is blocked. Additionally, fluorination of an aromatic ring deactivates the ring toward oxidative metabolism. This effect is nicely demonstrated in the example of ezetimibe (Zetia), a cholesterol absorption inhibitor, as seen in **Figure 4.10**. The parent compound, SCH 48461, exhibited moderate oral bioavailability, as evidenced by the effective dose in a hamster model of disease (ED_{50}) of 2.2 mg/kg. The eventual drug, ezetimibe, contains two fluorine atoms that were incorporated to minimize metabolism and has an effective dose of 0.04 mg/kg, a greater than 50-fold improvement.

Another benefit of the incorporation of fluorine is that molecules containing a fluorine atom can be rapidly converted into imaging or diagnostic agents for positron emission tomography (PET) by incorporation of the radioactive isotope of fluorine (^{18}F).

Several other examples of the use of classical bioisosteres are shown in **Figure 4.11**. The substitution of deuterium (H^2) for hydrogen (protium H^1) has received much recent attention. The advantages include slightly decreased lipophilicity and larger molar volume. Owing to its 2-fold larger mass, the C-D bond is stronger than a C-H bond, resulting in reduced metabolic liability at that site, a result of the kinetic isotope effect (KIE). The first approved drug (2017) based on a "deuterium switch" strategy was deutetrabenazine, based on the known drug tetrabenazine, used to treat Huntington's chorea. Since its approval, a number of novel drugs have advanced to clinical trials or have been recently approved. Deucravacitinib (approved in 2022) is a novel TYK2 inhibitor approved for treatment of plaque psoriasis. TKY2 (tyrosine kinase 2) is a member of the family of Janus kinases, with numerous roles in inflammation. Deucravacitinib is highly selective for TYK2 over other kinases, in part due to the deuterated methyl amide. Due to the KIE, deuteration slows metabolism that generates the primary amide, an active metabolite which is less selective for TYK2. Other examples in **Figure 4.11** are the substitution of NH_2 for OH, as seen in aminopterin, an early anti-cancer drug, as well as the divalent substitution of -CH_2- for -O-, as seen in iloprost, an analog of prostacyclin, with improved stability.

Nonclassical bioisosteres do not necessarily have the same number of valence electrons or the same number of atoms, but they still result in similar biological activity as the original compound. Due to the broad definition, the success of this approach is often context-dependent, and a successful incorporation of a nonclassical bioisostere in one case does not always ensure success in another application. These substitutions may have effects on multiple parameters, such as size, hydrogen bonding ability, lipophilicity, and stability. For the sake of this discussion, nonclassical bioisosteres have been divided into two broad categories: cyclic/acyclic and functional group replacements.

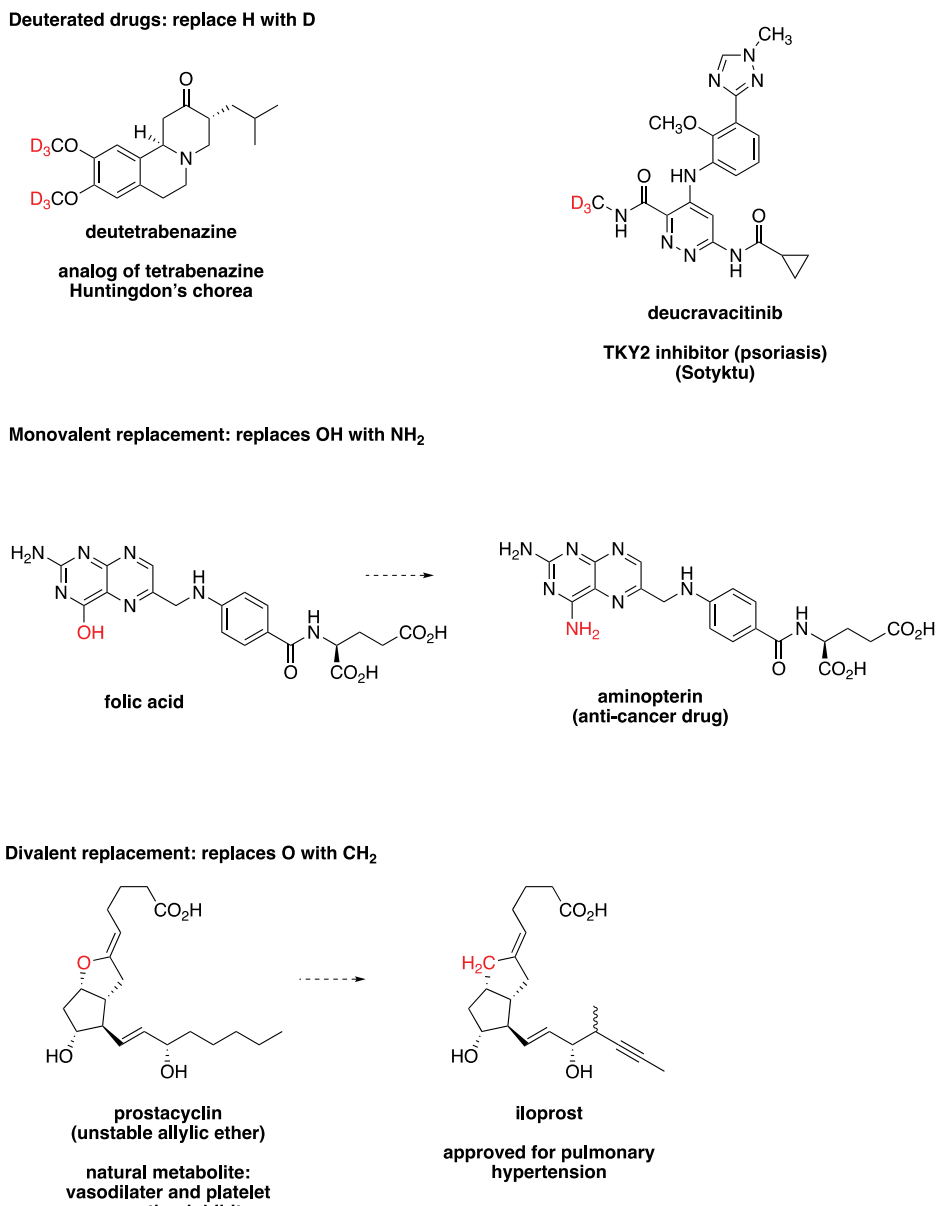


Figure 4.11 Examples of deuterated drugs and monovalent and divalent bioisosteres. Deutetrabenazine was the first deuterated drug approved and has similar activity to the parent tetrabenazine. Deucravacitinab, another deuterium-containing drug, is a selective TYK2 inhibitor approved to treat psoriasis. Aminopterin was an early anti-cancer drug and is an analog of folic acid, with an OH-to-NH_2 substitution. Substitution of CH_2 for oxygen in the natural ligand prostacyclin gave the more stable drug iloprost.

4.5.2.1 Nonclassical Bioisosteres: Cyclic/Acyclic

As it sounds, cyclic vs. acyclic bioisosteres are replacements of a ring with an acyclic analog, or *vice versa*. A classic example is that of estradiol and diethylstilbestrol, as seen in **Figure 4.12**. The rigid alkene is a bioisostere for the central fused cyclohexane rings of the steroid, as it orients the key pharmacophoric elements in the same position and maintains bioactivity as an estrogen receptor modulator.

4.5.2.2 Nonclassical Bioisosteres: Functional Group Replacements

Phenyl ring. Phenyl rings are found in nearly half of all drugs and are an important scaffold and binding element, due to the potential for π–system and hydrophobic interactions. However, phenyl groups, particularly unsubstituted ones, are prone to metabolic liabilities due to oxidation. The most common bioisosteres used to replace a phenyl group are heteroaromatics, both six- and five-membered rings. These can maintain π stacking interactions and may also provide additional hydrogen bond acceptors and have the advantage of reduced lipophilicity and improved solubility. Aliphatic bioisosteres, which maintain hydrophobic interactions, have seen more use

recently and have also been found to be effective and, in some cases, can improve properties such as solubility. The anti-diabetic drug carmegliptin is one example; it contains a pyrrolidinone ring as replacement for a phenyl group found in the lead compound (**Figure 4.12**). The aliphatic bioisosteres have an added advantage of "escaping from flatland," a recent trend based on the fact that increased numbers of sp^3 carbons give added complexity and natural product-likeness. Some specific examples of phenyl bioisosteres, as well as other functional group bioisosteres, are shown in **Figure 4.13**.

Carbonyl-containing derivatives. Bioisosteres of carbonyl groups, such as ketones, esters, and carboxylic acids, are varied, and most maintain the electronic properties of the carbonyl carbon and/or the oxygen lone pairs.

Ketones/aldehydes are traditionally replaced with thiocarbonyls, sulfoxides, and sulfones and also substituted ethylenes, such as dicyanoethylene. In some of these

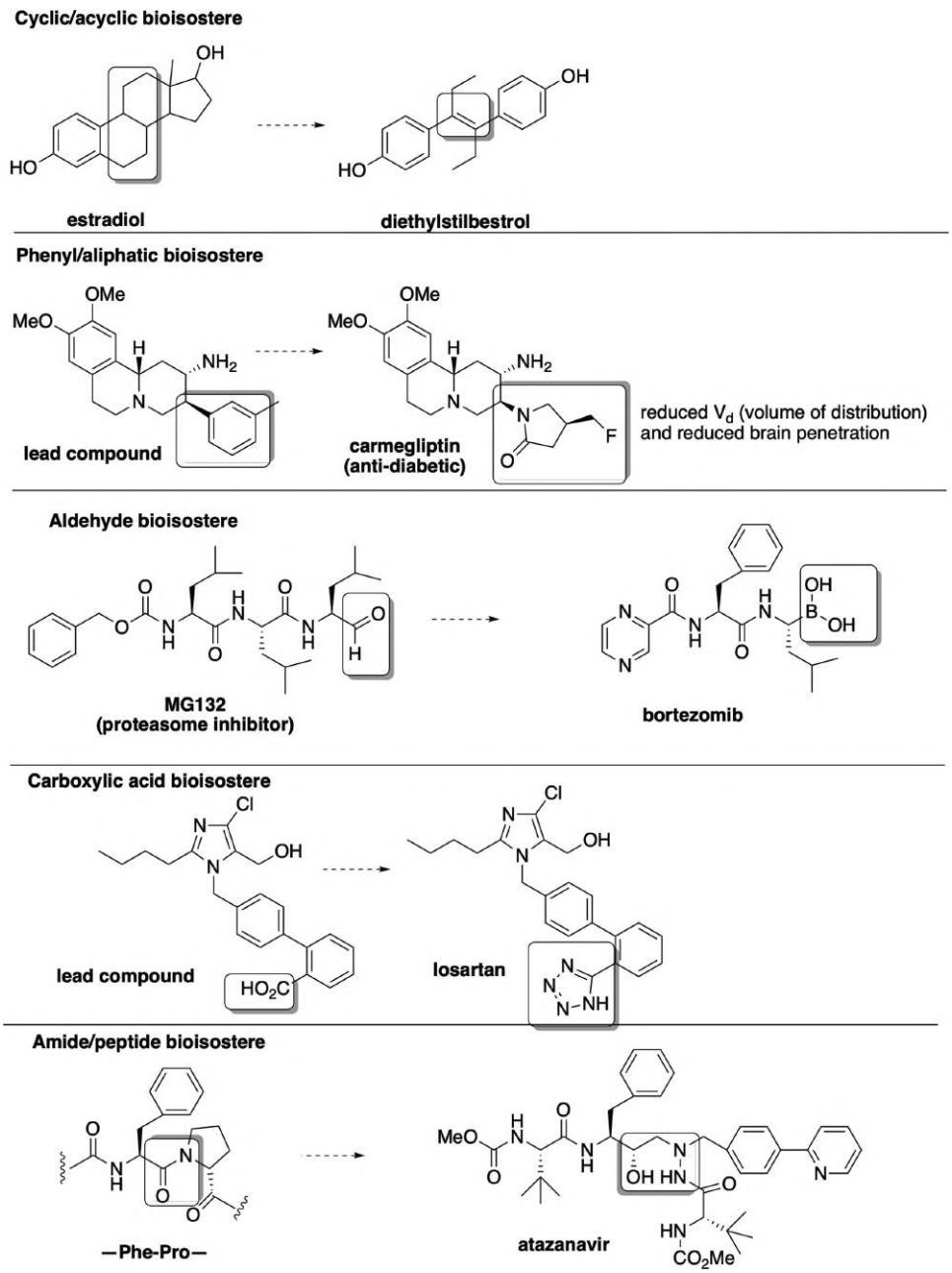


Figure 4.12 **Examples of drugs based on non-classical bioisosteric replacements.**

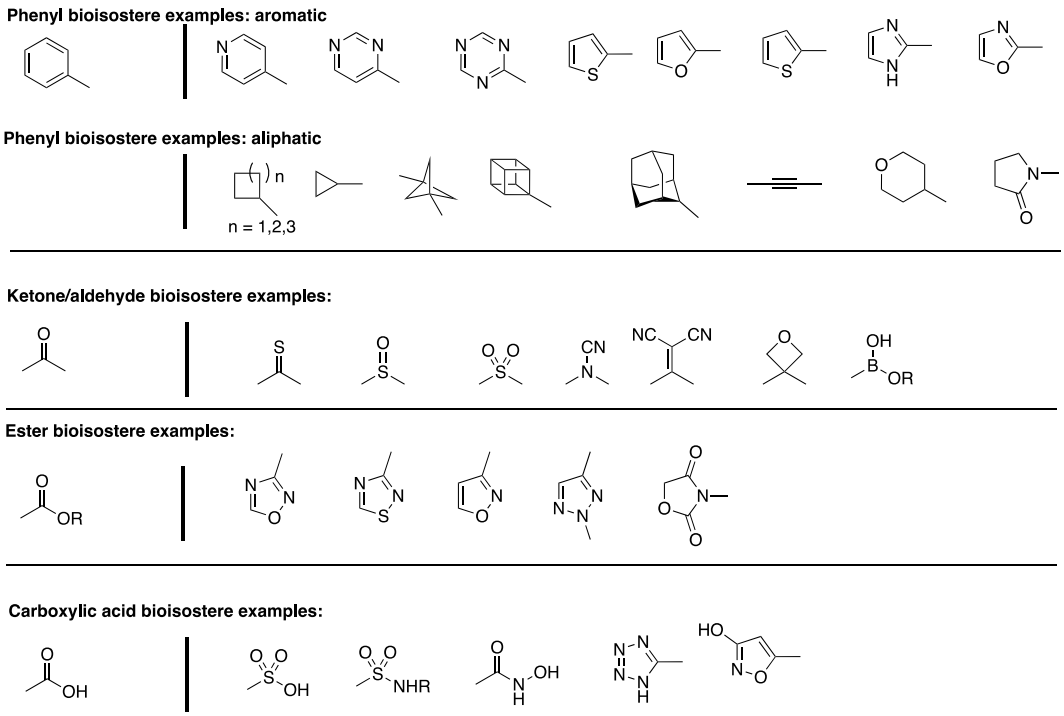


Figure 4.13 **Examples of nonclassical bioisosteres for selected functional groups.**

examples, it is relatively easy to see the similarity; for example, a carbonyl group and a sulfonyl group are similar in both size and electronic structure. However, the use of dicyanomethylene or N-cyano as a replacement for oxygen is less obvious, and one needs to consider the nitriles mimicking the oxygen lone pairs. More recent examples of ketone/aldehyde bioisosteres include cyclic oxetanes, providing three-dimensionality but maintaining the oxygen lone pairs. A number of boron-containing isosteres mimic carbonyls, particularly aldehydes, with the electrophilic boron able to form reversible covalent bonds with a nucleophile in a target and the oxygen maintaining hydrogen bonds. Several boron-containing drugs have been approved, such as bortezomib (**Figure 4.12**) and ixazomib, used to treat multiple myeloma.

Carboxylic ester mimics are mainly heterocyclic replacements, such as oxadiazoles, thiadiazoles, and oxazolidinediones. These heterocycles remove the liability of hydrolysis by esterases while maintaining binding properties.

Carboxylic acids are frequently important binding elements in leads, and there are many examples of bioisosteric replacements that contain ionizable hydrogens (pK_a = 1–8) or H bonding donor or acceptor capabilities. By the incorporation of a carboxylic acid isostere, the same binding interactions to the target are usually maintained, but the physical and pharmaceutical properties of the molecule, such as lipophilicity, permeability, or metabolic stability, can be modified. Substitutions for the carboxylic acid may be cyclic or acyclic but typically maintain an acidic hydrogen. While carboxylic acid bioisosteres may not be identical in size compared to the acid, they often have similar pK_a values for the acidic hydrogens. For example, the widely used carboxylic acid bioisostere 1*H*-tetrazole has a pK_a of ~4.6, and the corresponding acetic acid pK_a is ~5.0. This replacement is seen in the blood pressure drug losartan (**Figure 4.12**). Bioisosteric replacement of carboxylic acids is typically aimed at maintaining potency but improving permeability and/or metabolic stability.

Amide/peptide mimics can be considered a special class of functional group bioisosteres. For the purposes of this discussion, we will treat them similarly even though a peptide bond is specifically formed between two α-amino acids, whereas many drugs contain amide bonds that are not part of peptide linkages. While peptides or segments of proteins exert a biological effect as ligands for receptors or as substrates for enzymes, they are often not suitable as drugs due to their physical properties

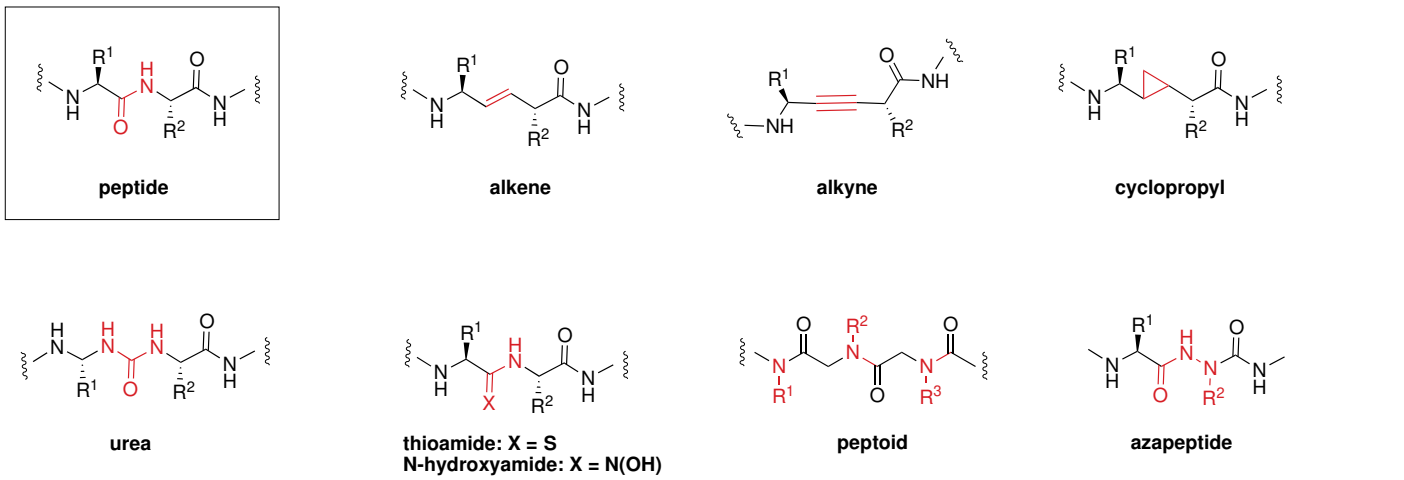


Figure 4.14 Examples of peptide mimetics. Replacement of an amide bond (red) in a peptide (boxed) with various isosteres can result in similar potency and improved properties, particularly metabolic stability. Alkene and alkyne isosteres mimic the planar features of a peptide bond; a cyclopropane isostere orients the side chains (R^1, R^2; purple) in an orientation similar to the amide bond; ureas, thioamides, N-hydroxy amidines, and peptoids provide similar H bonding capacity. Azapeptides (semicarbazides) replace an alpha carbon with a nitrogen and have improved metabolic stability. In each case, the amide replacement is shown in red.

and susceptibility to hydrolysis. To address these limitations, peptide mimetics have evolved. These include functional groups that are designed to mimic the peptide bond's polarity, H bonding capabilities, and other features but exhibit improved stability. Examples of peptide mimetics, some of which are shown in Figure 4.14, include alkenes and alkynes that mimic the planar features of the peptide bond, cyclopropanes that correctly orient the amino acid side chains, and ureas, thioamides, N-hydroxyamidine, azapeptides, and peptoids that maintain some hydrogen bonding features. While there are only a few examples of these bioisosteres in currently approved drugs, they continue to be important tools and are an area of active research. One example (Figure 4.12) is the HIV protease inhibitor atazanavir, which is an azapeptide mimic of the amide of the Phe-Pro portion of the enzyme substrate.

The specific classification of isosteres has become less important as this strategy is finding broad use and new examples of bioisosteric replacements are being reported. One online database of bioisosteric replacements is the widely used tool found at http://www.swissbioisostere.ch.

4.5.3 Scaffold Hopping Is a Subset of Bioisosterism

While bioisosteres are typically replacements for pharmacophoric elements, **scaffold hopping** involves the replacement of the non-pharmacophoric scaffold of a molecule. The goal of this tactic, first described by Gisbert Schnieder at Roche in 1999, is to identify novel bioactive compounds by replacement of the scaffold or core. Scaffold hopping is most often used to improve the physical or pharmaceutical properties of the lead structure, to expand intellectual property around a series, or to simplify the structure and therefore the methods for synthesis. In the simplest case, a new scaffold places the pharmacophoric elements in the identical orientation as the original scaffold. In some cases, the new scaffold itself may engage in additional binding interactions, providing improvements in affinity or potency. Computational chemistry approaches (described in what follows) are very frequently used to identify novel scaffolds.

Scaffold hopping can be considered a subset of bioisosterism and can involve traditional strategies, such as isosteric replacements within rings and chains. As described earlier, typical changes to a lead scaffold may include replacing one heterocycle with another or the incorporation of rings or substituents to reduce conformational flexibility. Some examples of drugs with different scaffolds but similar orientation of pharmacophoric elements and similar bioactivity are shown in Figure 4.15. An older scaffold hopping pair are the antihistamines pheniramine and the more rigid cyproheptadine. In the overlay, it can be seen that the pharmacophoric elements, the two aryl rings and the amine, are held in the same position (scaffold differences highlighted in red). In the case of cyproheptadine, the new scaffold imposes conformational constraint. In a second example, the cyclooxygenase (COX) inhibitors celecoxib and rofecoxib have different heterocyclic scaffolds (red) but very similar functional

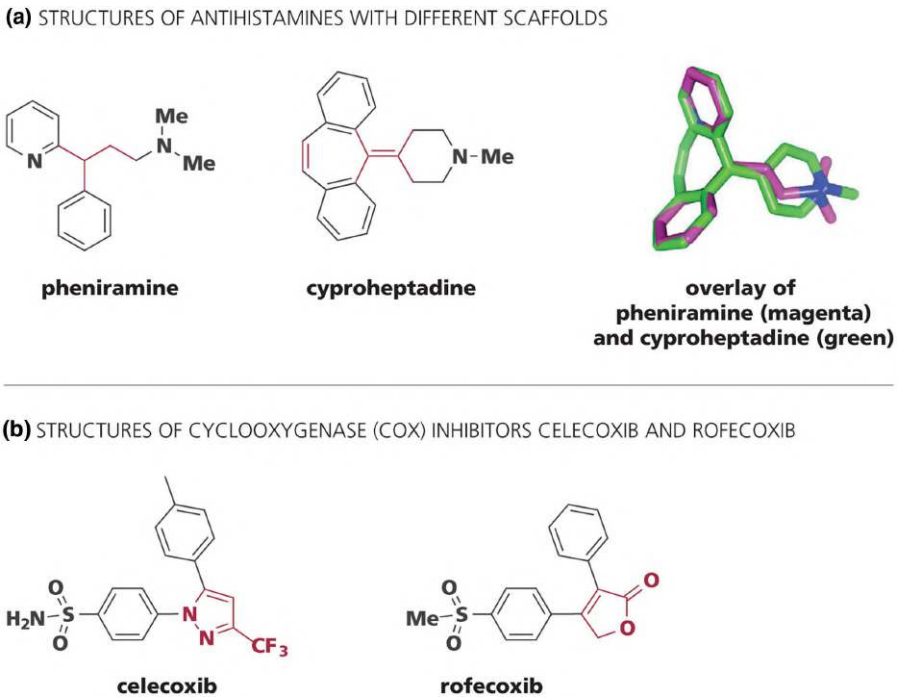


Figure 4.15 Examples of pairs of compounds with the same bioactivity but different scaffolds: (a) Two antihistamines, pheniramine and cyproheptadine, contain the same pharmacophoric features (two aryl groups and a positive ionizable group), on distinct scaffolds, with key scaffold differences highlighted in red. The overlay (*right*) shows that both (pheniramine, magenta; cyproheptadine, green) orient the terminal groups in an identical orientation. (b) Two approved COX inhibitors, celecoxib and rofecoxib, each containing nearly identical aryl sulfonamide and phenyl groups, but different heterocyclic scaffolds (in red).

Source: Overlay of structures from Sun, H.; Tawa, G.; and Wallqvist, A. (2010). *Drug Discovery Today* 17:310–324. With permission from Elsevier.

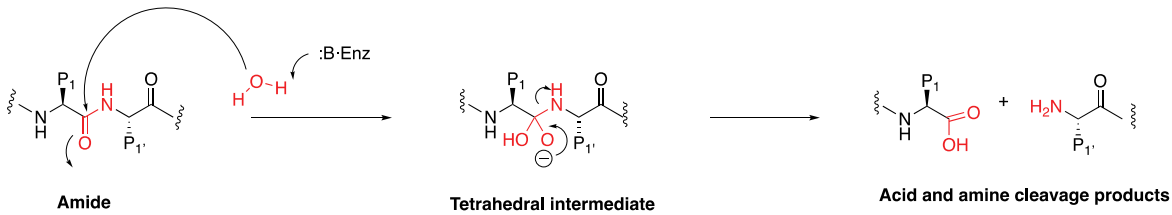
groups on their periphery. While the two drugs have similar potencies, their pharmacokinetic profiles are different.

Computational chemistry techniques are frequently used to identify new scaffolds that maintain the pharmacophore but are dissimilar enough to provide novel structures. These techniques traditionally rely on pre-defined databases and may include 3D shape similarity and/or fingerprint searching, pharmacophore matching, or fragment replacement. A common set of defined scaffolds are the BM scaffolds, introduced by Bemis and Murcko in the late 1990s. Some newer approaches involve generative AI models that produce *de novo* design of novel structures.

4.5.4 Transition State Mimetics Are Common Features of Enzyme Inhibitors

One strategy that has been particularly valuable in the development of enzyme inhibitors is the use of **transition state mimetics**. Since an enzyme stabilizes the transition state of the reaction it catalyzes, effective inhibitors can be developed by mimicking the transition state with a stabilized structure. Inhibitors of proteases, enzymes that catalyze the hydrolysis of a peptide bond, are one such example where this approach has been particularly successful. During the hydrolysis of the substrate peptide, an sp^2 carbonyl is attacked by a nucleophile (usually derived from either serine, cysteine, or threonine, or a water molecule) and generates an sp^3-hybridized tetrahedral intermediate, which then collapses to eventually generate the hydrolysis products. This process, for an aspartyl protease when the nucleophile is water, is outlined in abbreviated form in Figure 4.16.

Figure 4.16 Abbreviated mechanism for aspartyl protease–catalyzed amide bond hydrolysis, showing tetrahedral intermediate and products. An attack of a deprotonated water molecule on the carbonyl of the amide bond generates a tetrahedral intermediate, which then collapses to generate the acid and amine fragments.



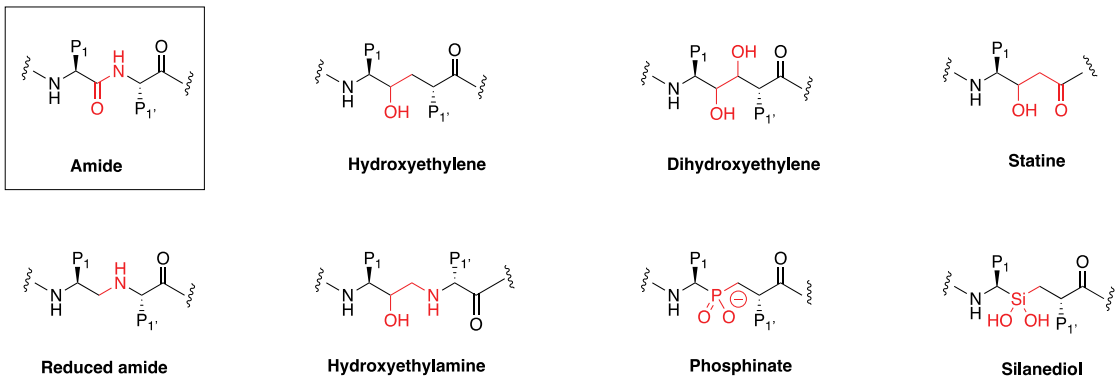


Figure 4.17 Amide bond that is cleaved by proteases, and examples of transition state analogs. P_1 and $P_{1'}$ are the amino acid side chains on either side of the cleaved (scissile) bond (red, boxed), and the specific transition state mimetic is shown in red. Hydroxyethylenes, dihydroxyethylenes, statines, hydroxyethylamines, phosphinates, and silanediols contain a hydroxyl group at the sp^3 center that may mimic the alkoxide present in the transition state. Reduced amides maintain all the features of the transition state, except the hydroxyl group.

On the basis of this hydrolysis mechanism, a stable sp^3-hybridized mimic of the tetrahedral intermediate is predicted to bind tightly to and inhibit aspartyl proteases. A number of examples of transition state mimetics of proteases have been developed and are shown in Figure 4.17. Hydroxyethylenes, dihydroxyethylenes, statines, hydroxyethylamines, phosphinates, and silanediols not only contain the tetrahedral carbon mimetic but also display a hydroxyl group that may mimic the alkoxide that is present in the transition state. Other examples, such as reduced amides, maintain all the features of the transition state, except the hydroxyl group.

These peptide-like transition state mimetics resemble the extended conformation of the substrate peptide and are incorporated into a larger scaffold that is based on the natural substrate. Hydroxyethylene and hydroxyethylamine transition state mimetics were incorporated into the HIV protease inhibitors indinavir and amprenavir, respectively (Figure 4.18), and are examples of the successful application of this strategy.

Another example of transition state mimetics is found in the influenza drugs zanamivir (Relenza) and oseltamivir (Tamiflu). The viral surface glycoprotein neuraminidase (labeled as HB-Enz) is a glycohydrolase enzyme that relies on acid–base catalysis to cleave terminal sialic acid residues from sialic acid glycoconjugates on the virus and host cell surfaces, allowing for the release of viral particles (Figure 4.19). The reaction proceeds through a planar intermediate prior to the addition of water, which generates sialic acid. The unsaturated rings in zanamivir and oseltamivir mimic the sialosyl cation, in which the anomeric carbon transitions from an sp^3-hybridized carbon to an sp^2 carbon.

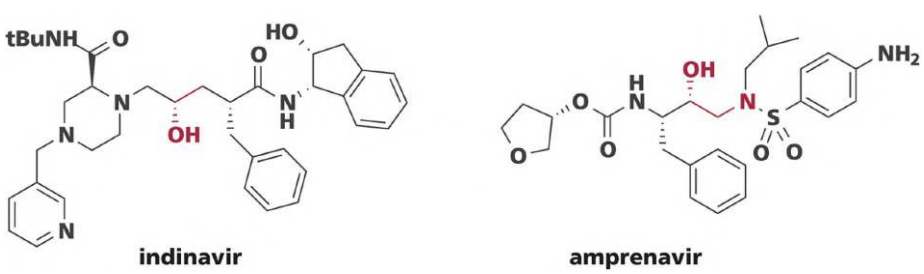


Figure 4.18 Examples of HIV protease inhibitors that contain transition state mimetics (shown in red). Indinavir contains a hydroxyethylene moiety in place of the amide bond, and amprenavir contains a hydroxyethyl amine group.

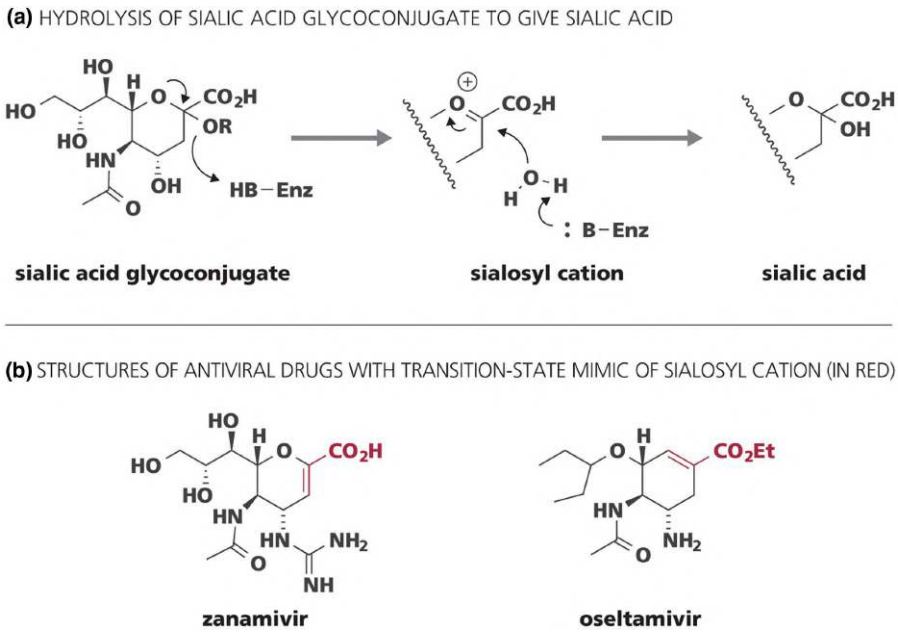


Figure 4.19 Glycohydrolase mechanism and examples of transition state mimetics. (a) Sialic acid glycoconjugate is hydrolyzed by the enzyme (HB-Enz) using acid–base chemistry, to generate a sialosyl cation in the transition state, which eventually generates sialic acid. (b) Structures of antiviral agents zanamivir and oseltamivir. Transition state mimetic moieties (here mimicking the planar transition state intermediate) are shown in red.

The use of transition state mimetics is a powerful strategy for the design of enzyme inhibitors. In theory, if the transition state of an enzymatic reaction is understood, inhibitors can be designed. The current challenge in this area is the design of transition state mimetics of multicomponent enzymatic reactions and those that involve high-valency intermediates, for example, the transition state in phosphate transfer reactions catalyzed by kinases that are not easily mimicked by traditional organic functional groups. Chapter 7 provides additional details on transition state mimetic strategies as applied to drugs and drug design.

4.5.5 Conformational Constraints Can Be Used to Improve Binding to the Drug Target

The goal of introducing conformational constraints into an active molecule is to maintain the enthalpic effects contributed by the binding interactions but reduce the entropic effects. The overall result is an increase in binding energy, since the molecule does not require energy to rotate into an active conformation. This strategy is frequently applied by incorporating rings into acyclic structures. For example, the natural ligand for the metabotropic glutamate receptors (mGluR) is the excitatory amino acid glutamic acid. A number of cyclic analogs of this flexible molecule have been developed that bind tightly to the receptor. LY354740 (eglumegad), which progressed to clinical trials as a potential anti-anxiety drug, uses a rigid bicyclic scaffold to constrain the amine and carboxylic acid binding elements into specific orientations, as shown in **Figure 4.20a**. It is likely that this constraint mimicked the conformation of glutamic acid that was recognized by the targeted receptor and not other glutamate receptors, thereby also enhancing selectivity and safety. In another example, a series of macrocyclic analogs of a hexapeptide inhibitor of the hepatitis C viral protease HCV NS3 were shown to be 26–50 times more potent than the acyclic analog. The macrocyclic ring (red) introduces a conformational constraint compared to the flexible linear peptide.

The final example in **Figure 4.20a** shows a kinase inhibitor that was identified by high-throughput screening as a potential cancer drug. Its common aniline-pyrimidine scaffold core was reported in multiple patent applications covering kinase inhibitors. In this case, macrocyclic analogs were designed (e.g., SB1317/TG02; red highlights the constraining atoms) that would maintain binding interactions as well as

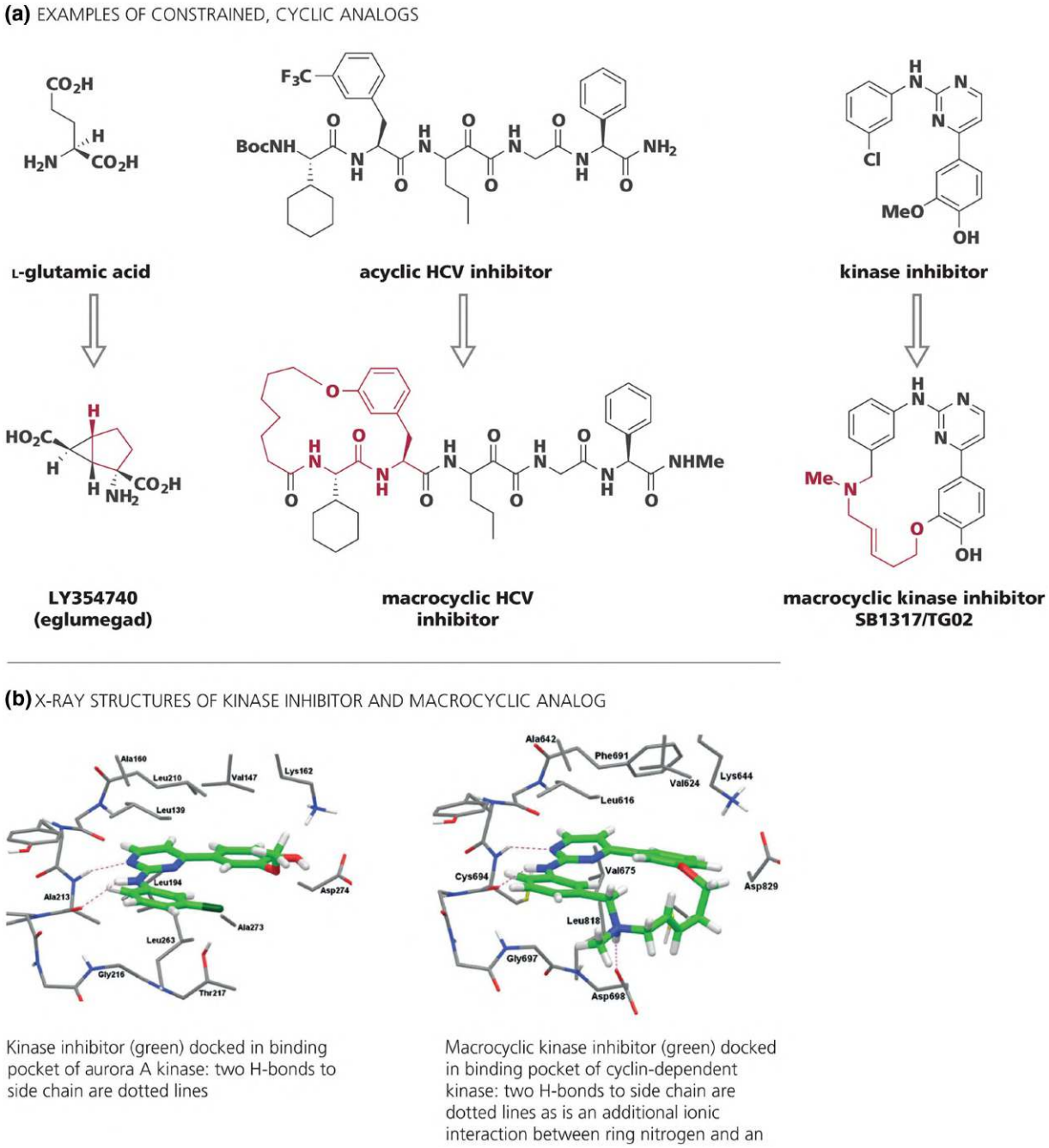


Figure 4.20 Examples of using conformational constraint as a medicinal chemistry strategy. (a) Examples of constrained cyclic analogs and original linear analogs: Eglumegad contains a rigid bicyclic core and mimics L-glutamic acid; the linear peptide HCV protease inhibitor was constrained by the incorporation of a macrocycle, which improved potency, and a macrocyclic kinase inhibitor, based on a common kinase inhibitor core afforded SB1317/TG02 that had improved patentability (conformational constraints are shown in red). (b) Structures of kinase acyclic inhibitor (*left*) docked into the ATP pocket in an X-ray of aurora kinase, and macrocyclic analog (*right*) docked into the ATP pocket in an X-ray of a related cyclin-dependent kinase 2 (CDK2).

Source: Docked X-ray structures from Poulsen, A.; William, A.; Blanchard, S.; et al. (2013). *J Mol Model* 19:119–130. With permission from Springer.

improve patentability. Docking of the acylic and cyclic inhibitor into a kinase model shows they adopt nearly identical conformations (Figure 4.20b) and detects hydrogen bonds to the backbone of the binding pocket, with an additional ionic interaction between the macrocycle nitrogen and an aspartic acid. A structural analog of SB1317 called pacritinib is approved as an orphan drug to treat a rare bone marrow disorder.



Figure 4.21 Diagram showing the general structure of stapled peptides. The incorporation of alkene-containing amino acids (red balls) into a peptide strand (pink alpha helix) is a strategy to form stapled peptides. Ring-closing metathesis of the alkenes affords a stapled peptide, with an alkene-containing macrocycle linking the two amino acids.

In addition to incorporating a ring into a linear structure to constrain its conformation, other strategies include incorporating unsaturation (alkenes and alkynes) into an alkyl chain, incorporating steric bulk to restrict rotation and/or favor one conformation, and designing intramolecular H-bond pairs into molecules. By pre-organizing the small molecule into the active conformation, considerable increases in affinity, potency, selectivity, and permeability can be realized.

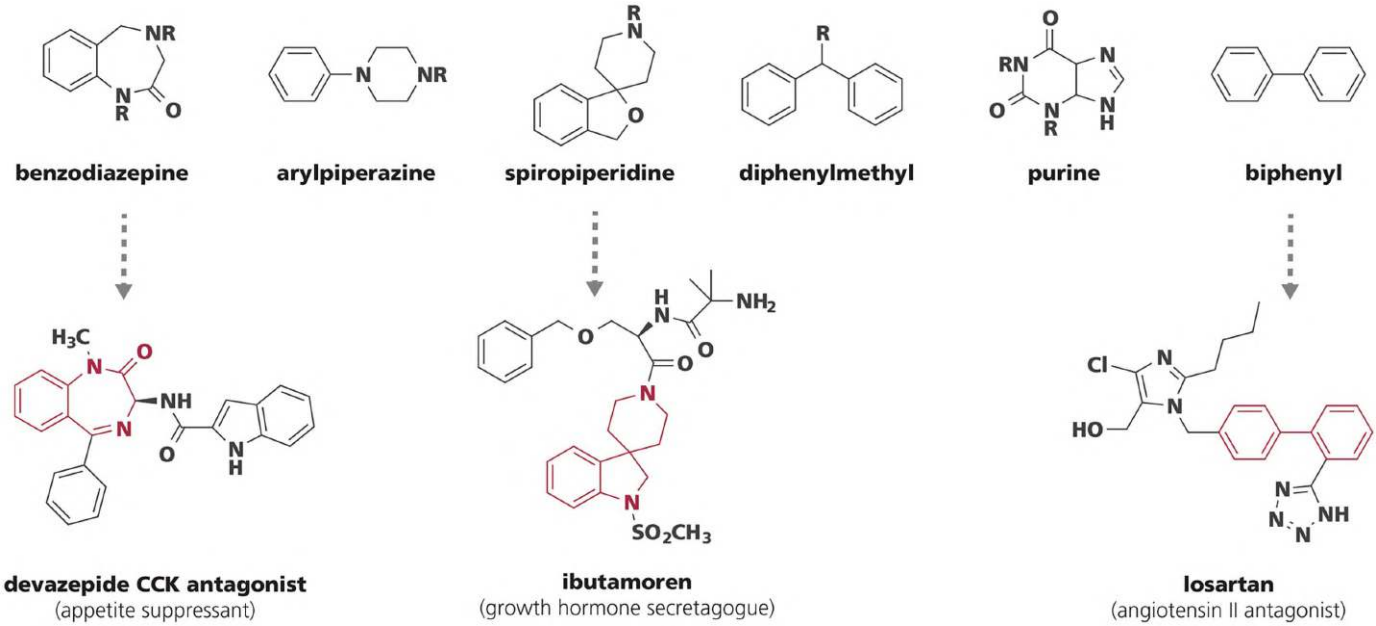
The strategy of constraining conformation has also been applied to larger structural motifs, for example, secondary structures of peptides, such as α-helices and β-turns. In these cases, a linear peptide fragment that is known to exert some desirable biological activity is constrained so that it mimics the conformation of the binding fragment within the larger, natural peptide. Considerable effort has been devoted to designing scaffolds that mimic a β-turn motif. **Stapled peptides**, where a macrocyclic carbon-based ring is formed between the residues of a small peptide, are another example, this time mimicking an α-helix (Figure 4.21). Frequently, peptides containing olefins are incorporated into specific points in the linear peptide, and a ring-closing metathesis reaction is used to favor an α-helical structure that mimics the active conformation and also increases the stability of the peptide. The use of this strategy is discussed further in Chapter 8.

4.5.6 Privileged Structures Are Structural Templates Found in Multiple Drugs

Privileged structures are structural motifs that are capable of binding to a variety of biological targets. The concept of privileged structure was defined in 1988 by Ben Evans at Merck and first applied to the benzodiazepine core, which binds to a number of different G protein–coupled receptors (GPCRs), as well as ion channels and enzymes. Subsequently, it was noted that certain high-throughput screening libraries based on privileged structures had higher hit rates than others. Examples of privileged structures (Figure 4.22) include benzodiazepines, arylpiperazines, spiropiperidines, diphenylmethyls, purines, and biphenyls.

The benzodiazepine structure is found in compounds such as devazepide that bind to the cholecystokinin (CCK) receptor (Figure 4.22), in anti-anxiety drugs such as Valium that bind to ion channels, and in compounds that bind to other G protein–coupled receptors and enzymes. The spiropiperidine core is found in growth hormone secretagogues, such as ibutamoren (Figure 4.22), as well as in compounds that bind to the melanocortin 4 (MC4) receptor. Buspirone contains the same core and binds to the serotonin receptor. The biphenyl core is found in the anti-hypertensive drug losartan, as well as in a number of other drugs.

Why privileged structures are privileged is not well understood. Some suggest that the structures have an optimized balance of conformational bias to adopt various geometries and excellent drug and pharmaceutical properties. Medicinal chemists often prioritize hits and early leads that contain these functional groups over other hits or look to incorporate privileged scaffolds into new structures during lead optimization programs.



4.6 SUMMARY

The process of lead optimization is an arduous process, where many changes to the original lead may be necessary in order to prepare a compound with all the properties appropriate for clinical development. Chemists have developed multiple strategies that are commonly used in this process in order to optimize biological, physical, and pharmaceutical properties. These include bioisosteric replacement, transition state mimetics, conformational constraint, and use of privileged structures. Optimal binding to the target must be achieved, while other properties, such as lipophilicity, metabolic stability, and solubility, are optimized. In some instances, it may be necessary to use a prodrug or targeting agent that, while not active itself at the target, will be converted in a living system to the active drug.

Figure 4.22 Examples of privileged structures and drugs containing those substructures (shown in red). Some common examples of privileged structures are given on the top row in the figure. Specific examples of drugs that contain those privileged structures (in red) include devazepide, with a benzodiazepine; ibutamoren, with a spiropiperidine; and losartan, with a biphenyl structure.

CASE STUDY 1: LEAD OPTIMIZATION OF LOSARTAN

An example of the optimization of a lead using a number of the strategies outlined earlier can be found in the development of the anti-hypertensive drug losartan. The initial lead was reported in the patent literature as an angiotensin receptor II (AII) antagonist (blocker). Chemists at DuPont Merck used a combination of molecular modeling and medicinal chemistry strategies to make hundreds of analogs of the initial lead, ultimately resulting in the first AII receptor antagonist to be approved, losartan. As outlined in **Figure 4.23**, the initial leads, which were the only reported non-peptide AII antagonists at that time, were modified on the basis of molecular modeling done by aligning the scaffold with the natural peptide substrate. It was predicted that the introduction of an acidic group in the para position of the benzene ring (shown in red) could mimic a tyrosine hydroxyl of the substrate. The designed compound exhibited improved binding over the original compound by 100-fold (IC_{50} = 1.2 μM vs. 150 μM). Further optimization of the position of the acid moiety, by inserting an additional aryl ring (shown in blue), constrained the acid into a specific position and led to a compound with potency in the submicromolar range (IC_{50} = 0.14 μM).

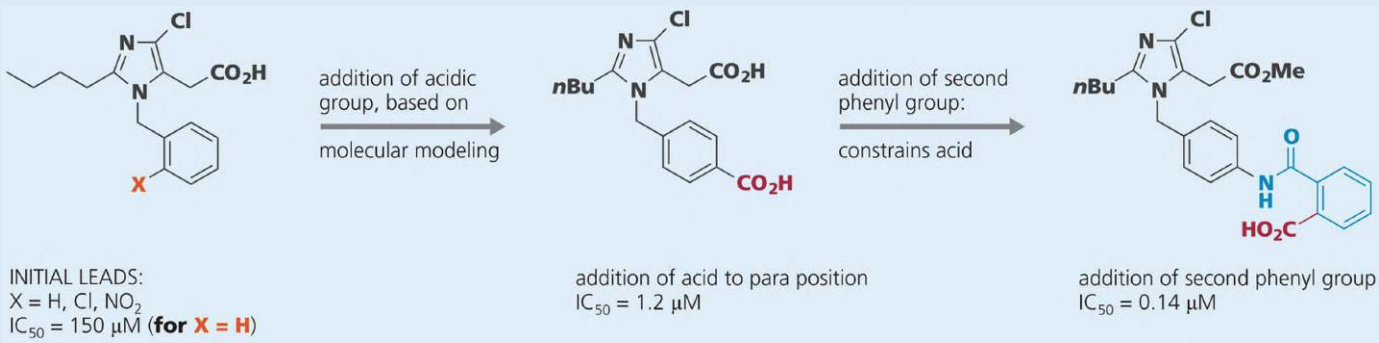


Figure 4.23 Losartan leads and modifications. Initial lead scaffold was optimized on the basis of molecular modeling by adding an acid group at the para position of the aryl ring (red), which provided a >100-fold improvement in potency. Further optimization by extending the position of the acid group via insertion of phenyl acetamide improved binding by constraining the acidic group into a specific position (red).

Optimization of the hydrophobic R^2 position in a simplified imidazole scaffold involved homologation, branching, and ring-chain transformations (**Figure 4.24** orange). For example, in the amide-linked series, the R^2 position was varied to H, methyl, ethyl, propyl, butyl, pentyl, hexyl, heptyl, isopropyl (branched), phenylethyl, benzyl, and cyclohexylmethyl (ring-chain). The conclusion was that the R^2 butyl group was optimal.

The amide linker (blue) was replaced with CO (ketone), O, S, OCH_2, *trans*-C=C, and NHCONH (carbamate), as well as eliminated (biphenyl). The amide, ketone, and O- and S-linked compounds, as well as the biphenyl analog, had similar potencies in a binding assay (not shown). In the series with the privileged biphenyl scaffold (**Figure 4.24**, *right*), when the R^2 butyl group was maintained, changes in the identity and size of the substituent at the R^5 position affected potency by about 10-fold. A summary of the SAR results is shown in **Figure 4.25a**.

The changes to this point had optimized receptor-binding activity (IC_{50}), but the analogs did not have acceptable oral availability. The most promising were those in the biphenyl series. The position of the carboxylic acid was found to affect oral bioavailability, with the 2-isomer (ortho) being optimal. A variety of bioisosteres that maintained the same pK_a (5–6) as the acid were prepared, including substituted triazoles and tetrazoles. The optimized tetrazole (losartan, **Figure 4.25b**) was almost ten times more potent than the acid (IC_{50} = 0.019 μM vs. 0.23 μM) and showed efficacy at doses of less than 1 mg/kg, while the acid required ~3 mg/kg to see the same effects. Losartan contains the bioisosteric tetrazole and the privileged biphenyl moiety.

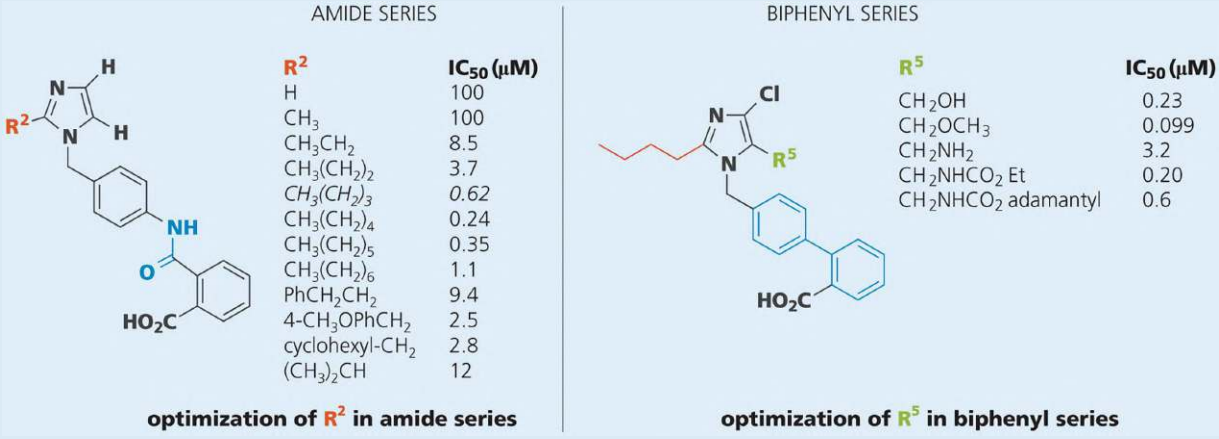


Figure 4.24 Examples of compounds prepared during the losartan optimization process in amide and biphenyl series. Strategies used for lead optimization in the amide series (*left*) included homologation, branching, and ring-chain transformations for the optimization of the R^2 group (orange). The biphenyl series (right) resulted from the incorporation of a privileged scaffold (biphenyl, blue) that eliminated the amide and the use of bioisosteres to investigate the R^5 group (green).

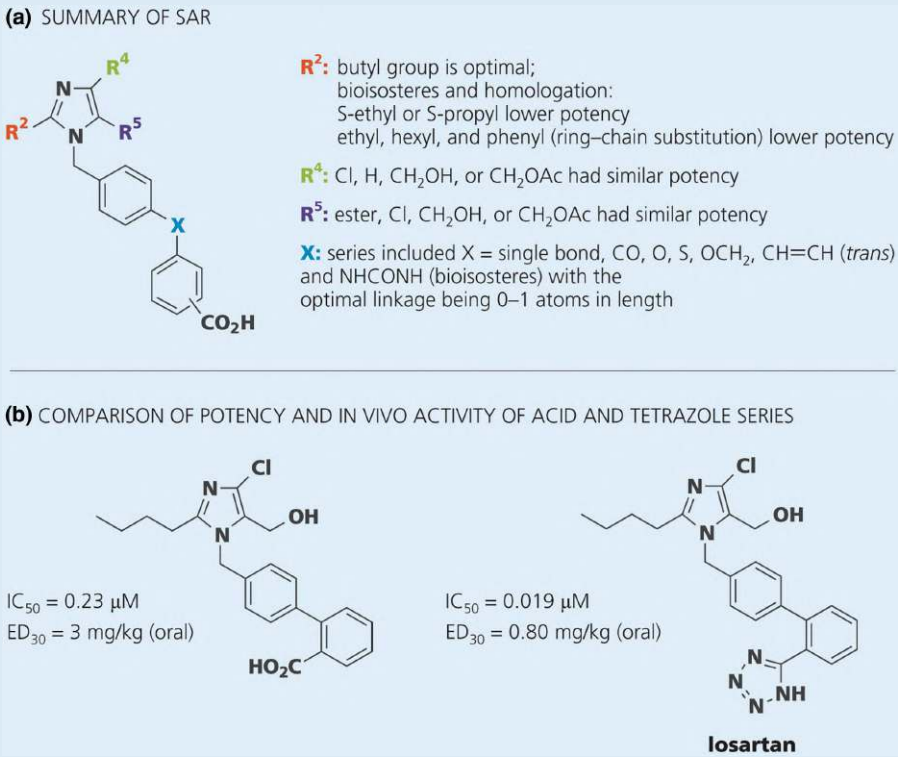


Figure 4.25 Summary of lead optimization to convert initial lead compounds to the optimized drug losartan. (a) Summary of SAR: The R^2 group was optimized to butyl. The R^4 position tolerated a variety of groups, as did the R^5 position; linkage (X) between the two aryl groups was best with zero or one atom. (b) Comparison of potency and ED_{50} of acid and tetrazole series. The incorporation of the tetrazole bioisostere improved both *in vitro* and *in vivo* potency.

Some of the medicinal chemistry strategies employed in this drug development program included the use of homologation, bioisosteres, and privileged structure (biphenyl).

CASE STUDY 2: REMDESIVIR, FIRST DRUG APPROVED TO TREAT COVID-19

With the emergence of viral infections caused by RNA viruses, such as HIV, hepatitis C, SARS, and Ebola, many organizations established efforts on the development of nucleoside analogs to identify broad-spectrum antivirals. This focus was based on the understanding that viral polymerases could be effectively targeted, as evidenced by the clinical success vs. HIV and herpesviruses, and that nucleoside inhibitors of one viral polymerase were also effective against polymerases from other viruses. Nucleosides such as the anti-HIV drug AZT and anti-herpes drug acyclovir are converted *in vivo* to the active corresponding triphosphate, which then inhibits the viral polymerase. This topic is covered in more detail in Chapter 11.

C-nucleosides are analogs in which the C-N bond between the furanose and the nucleobase is replaced with a more stable C-C bond. This change results in significant improvements in stability. A synthesis of such analogs in which there is substituent at the C1′ position reported by scientists at Gilead is outlined in **Figure 4.26**. The protected ribonolactone **1** was treated with lithiated aryl halides, such as **2**, to afford 1′-hydroxy 1′-C-nucleosides (hemiketals). Treatment of the hemiketals with alkylating agents, followed by deprotection of the benzyl ethers, afforded the 1′-substituted 1′-C-nucleoside analogs, such as **3a–3d**.

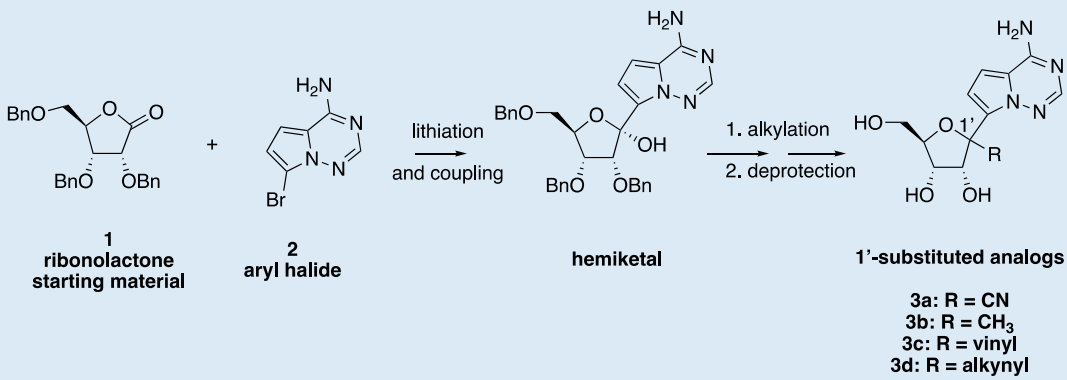


Figure 4.26 Synthetic approach to 1′-substituted, 1′-C-nucleosides. The addition of lithiated aryl halides, such as **2**, to protected ribonolactone **1** affords a hemiketal. Alkylation and deprotection of that intermediate gives 1′-C-nucleosides with 1′-substituents **3a–3d**.

Table 4.4 shows the antiviral activity of these 1′-substituted 1′-C-nucleosides against a series of RNA viruses, including HCV, SARS-CoV, and DENV-2 (dengue virus). The optimum compound in the series was the 1′-cyano analog (**3a** shown in **Table 4.4**), with broad activity (EC_{50}) and low cytotoxicity (CC_{50}).

Table 4.4 Antiviral activity of 1′-substituted 1′-C-nucleosides 3a–3d in cell-based assays measuring EC_{50}/CC_{50} (μM), where EC_{50} is a measure of antiviral activity and CC_{50} reflects cytotoxicity. Viruses are hepatitis C (HCV-subgenomic replicon), yellow fever (YFV), dengue (DENV-2), SARS-CoV (severe acute respiratory syndrome coronavirus A). Note: This is SARS-CoV (2002), not SARS-CoV2 (2019).

Virus	3a	3b	3c	3d
HCV	4.1/>89	39/>45	>89/>89	>89/>89
YFV	11/>30	ND	ND	ND
DENV-2	9.46/>30	>30/>30	>30/>30	>30/>30
SARS-CoV	2.24/>30	>30/>30	>14.0/>30	>30/>30

Use of Prodrugs for Nucleoside Antiviral Agents: ProTides (Pronucleotides)

Nucleosides polymerase inhibitors all must be converted to the corresponding triphosphate to generate the active species, which is recognized by the viral enzyme, as shown in **Figure 4.27**. Due to their charged nature, the triphosphate cannot be administered directly. In most cases, the first phosphorylation, catalyzed by kinases, is the most difficult. These kinases are more selective than the subsequent kinases that generate the triphosphate, and even when an unnatural nucleoside is recognized, the reaction can be slow. To address this, a prodrug strategy using derivatives called ProTides has been developed and applied to these nucleosides.

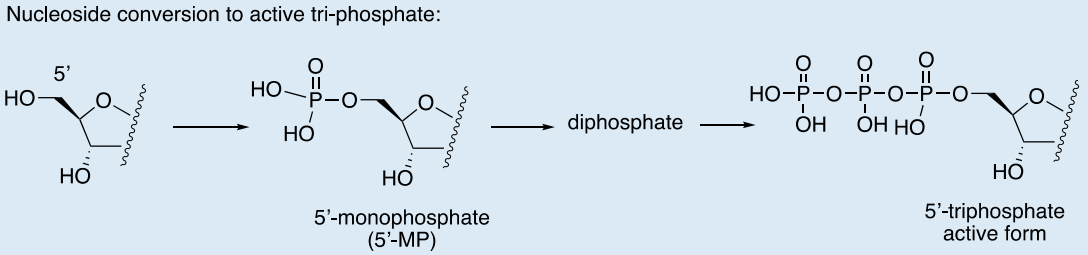


Figure 4.27 General structures showing 5'-phosphorylation of nucleosides in the formation of active triphosphates.

This ProTide strategy, introduced by Chris McGuigan in the 1990s, had been used with success in other drugs, such as sofusbuvir, approved to treat HCV. Amino acid esters of phosphoramidate amides, shown in **Figure 4.28**, undergo efficient ester hydrolysis *in vivo* to afford an intermediate that cyclizes and then undergoes hydrolysis, followed by enzymatic cleavage by the histidine triad enzyme Hint1, to afford the 5'-monophosphate (5'-MP). This strategy was applied to a number of the nucleoside antiviral agents.

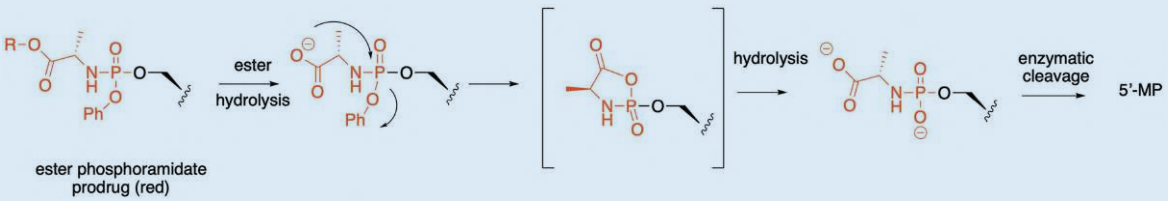


Figure 4.28 General structure of an ester phosphoramidate prodrug, and the mechanism of its conversion to the 5'-monophosphate. *In vivo* ester hydrolysis is followed by intramolecular attack on the phosphorous. Hydrolysis of that cyclic intermediate is followed by enzymatic cleavage to generate the monophosphate.

With their large library of nucleoside analogs, Gilead Sciences established a program in collaboration with the CDC (Centers for Disease Control and Prevention) and USAMRIID (US Army Medical Research Institute of Infectious Diseases) to identify potential drugs to treat RNA viral infections that had the potential to cause a pandemic, such as SARS, MERS, and Ebola, as well as other RNA viruses, such as RSV (respiratory syncytial virus) and HCV (hepatitis C virus).

One compound that progressed to Phase I clinical trials for HCV was the 1'-substituted 1'-C-nucleosides ProTide prodrug **GS-6620**. Although promising against HCV infection, the compound showed variable human pharmacokinetics and was dropped in 2014. A compound related to the GS-6620 parent nucleoside (without the C2'-methyl group), **GS-441524**, was also tested against HCV but was less potent than **GS-6620**. Concomitant with testing for HCV, Gilead also screened their library against other viruses. A high-throughput phenotypic screening of 400,000 compounds was



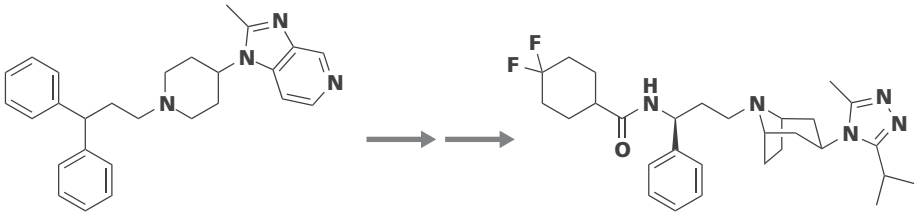
Figure 4.29 **Structures of GS-6620, GS-441524, and its prodrug, remdesivir (GS-5734).**

completed using a cervical adenocarcinoma cell line and assessed the inhibition of RSV A2–induced cytopathic effects also identified **GS-441524**, which targets the RSV polymerase RdRp (**Figure 4.29**). In subsequent screening, **GS-441524** was part of a library of about 1,000 ribose nucleosides and nucleotides that were screened in cell-based assays for antiviral effects vs. several other RNA viruses, including Ebola, MERS, SARS-CoV, and Dengue viruses. From that, remdesivir (**GS-5734**), the prodrug of **GS-441524**, showed potent activity against RSV, SARS-CoV, and Ebola.

Within Gilead's series of 1'-substituted 1'-C-nucleosides, remdesivir (GS-5734) had been patented in 2011 as a potential treatment for RSV and was reported in 2015 as a potential treatment against the Ebola virus, where it was given successfully to two patients under compassionate use. With COVID-19, one of the first approaches was to review the literature and legacy libraries for compounds that could be repurposed for COVID-19. It was through this effort that remdesivir was (re-)identified. Remdesivir had advanced to clinical trials for the treatment of Ebola and was found to be relatively nontoxic but was ultimately not approved. With the emergence of the RNA virus SARS CoV-2, remdesivir was well suited for testing. It had proven antiviral activity and a good safety profile. It was tested and received rapid approval, in October of 2020, less than a year after the initial COVID-19 outbreak.

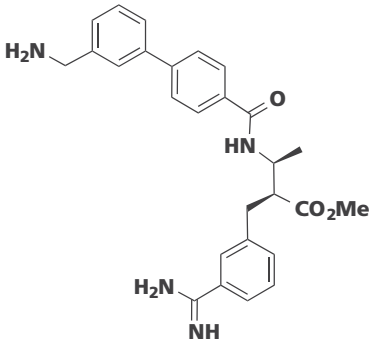
4.7 REVIEW QUESTIONS

1. The HIV drug maraviroc was developed by lead optimization of a hit found by high-throughput screening. Compare the structure of maraviroc to that of the original hit, and identify three strategies that were applied during lead optimization and incorporated into the drug molecule.

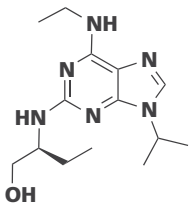


hit from HTS **maraviroc anti HIV agent**

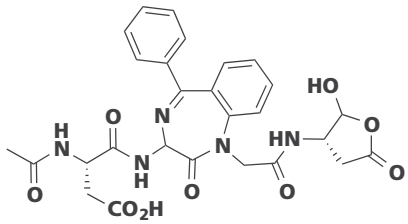
2. Each of the drugs shown contains a different privileged structure. Circle and identify that portion of each molecule. (Structures adapted from Constantino, L., and Barlacco, D. [2006]. "Privileged structures as leads in medicinal chemistry." *Curr Med Chem* 13:65–85.)



factor Xa inhibitor



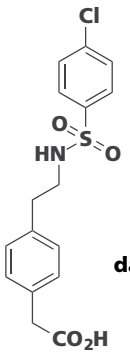
roscovitine CDK inhibitor



caspase inhibitor

3. Explain the difference between a prodrug approach and a targeting approach.

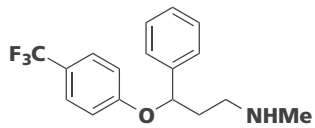
4. The compound shown, daltroban, is a thromboxane antagonist. Design two new compounds, each containing one fluorine atom. Explain what changes the incorporation of the fluorine atom should impart onto the new analog compared to daltroban.



daltroban

5. What important features should an acid isostere have in order to mimic a carboxylic acid?

6. Fluoxetine, shown in the following, is a highly prescribed antidepressant. It is a selective serotonin reuptake inhibitor and acts on serotonin transporters in the brain. It contains a CF_3 functional group. What properties might the CF_3 impart on this molecule that are important for its effects as a drug of this class?

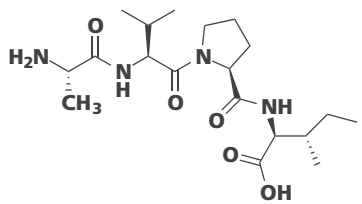


fluoxetine

7. Explain why lipophilicity is an important property for a drug candidate.
8. Explain why solubility is an important property for a drug candidate.
9. What are some reasons a prodrug strategy might be adopted (for example, what problems of the original active compound could be solved with a prodrug)?
10. Explain the difference between primary and secondary assays and some of the reasons a compound may be active in a primary assay but inactive, or much less potent, in a secondary assay.

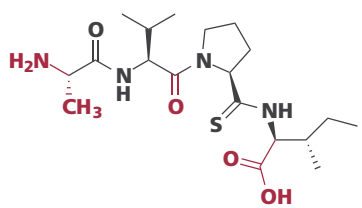
4.8 APPLICATION QUESTIONS

1. AVPI is a tetrapeptide that mimics the activity of the protein Smac. The Smac protein binds and inhibits the protein XIAP. Therefore AVPI is a potential starting point (hit) for a drug discovery effort targeting inhibitors of the protein XIAP. Identify two groups that could be replaced with isosteres, and design two different molecules that incorporate two isosteric replacements.

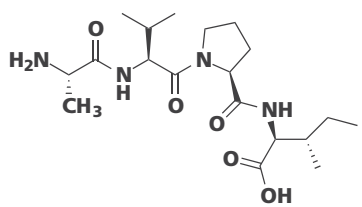


AVPI

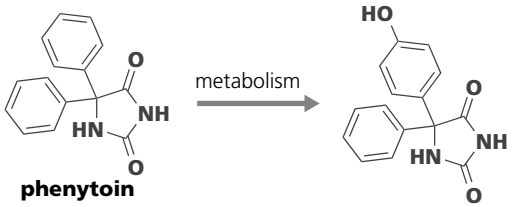
2. Using the same AVPI peptide, assume that the pharmacophoric elements are those in red. Design two different molecules that incorporate conformational constraints. The molecule is drawn in a way that is suggestive of the active conformation.



3. Design two different peptide bond isosteres for the peptide substrate AVPI.



4. Phenytoin is an anticonvulsant drug. It is metabolized to generate a product that is hydroxylated on one of the phenyl rings. Design two analogs of phenytoin that would reduce the metabolism at this position, and explain your rationale.

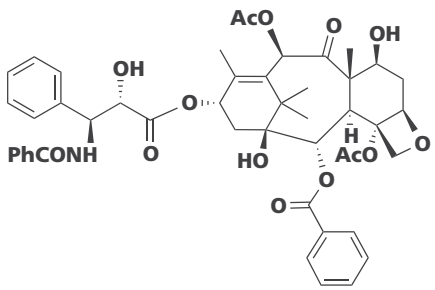


phenytoin

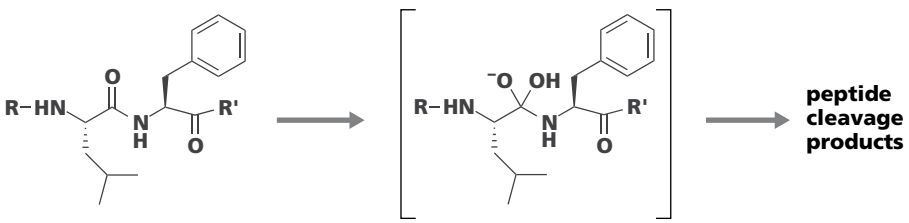
5. PubChem is a publicly available database sponsored by the National Center for Biotechnology Information: https://pubchem.ncbi.nlm.nih.gov. Perform a compound search for the three orally available drugs Valium, aspirin, and saquinavir. For each compound, compare the information in PubChem to the values in **Table 4.2** to determine whether each of the three compounds adheres to Lipinski's rule of five. Similarly, use the values to evaluate whether each of the three complies with the Veber rules.

6. Silicon is isosteric with carbon. It has been used in several drug candidates and some research compounds, but to date, no approved drugs have resulted. What similarities and differences will a silicon-containing compound have compared to the carbon analog? Research a drug for which the silicon analog has been prepared, and compare the properties.

7. Paclitaxel, a natural product, is an important anti-cancer drug, but it has poor solubility. Design two new analogs that you predict would be more soluble than paclitaxel itself, and rationalize your designs.



8. The structure shows a portion of one of the peptide substrates of an enzyme thought to contribute to the progression of Alzheimer's disease and the transition state through which it proceeds to generate the peptide cleavage products. An inhibitor of this protease has the potential to treat Alzheimer's disease. Design two peptide mimetics based on the transition state drawn.



transition state

9. A literature search for "conformational constraint" in the *Journal of Medicinal Chemistry* gives over 30 examples for the years 2016–2017. One article describes a series of conformationally restricted gamma secretase inhibitors that were evaluated as potential Alzheimer's drugs (*J. Med. Chem.* 2017, 60, 2,383–2,400). Identify the two different conformational constraints that were introduced to the lead compound. These compounds will need to penetrate the blood–brain barrier, so describe the effect of modifications on the cLog*P*.

10. The $-SF_5$ moiety is related to the $-CF_3$ group and has been incorporated into drug candidates as a CF_3 substitution. Research the properties of this group compared to the CF_3 group—size, lipophilicity, electronic character in the context of a drug discovery project—and find an example of where it was incorporated into a drug discovery project.

11. Use the publicly available database SwissBioisotere to:

a. Search for information on the bioisosteric replacement of indole with azaindole. What examples are highlighted?

b. Search for information on possible replacements for an aldehyde. What are ten different potential replacements?

FURTHER READING

Optimization of Physical and Pharmaceutical Properties

Bembenek SD, Tounge BA & Reynolds CH (2009) Ligand efficiency and fragment-based drug discovery. *Drug Discov Today* 14:278–283.

Hopkins AL, Keseru GM, Leeson PD et al. (2014) The role of ligand efficiency measures in drug discovery. *Nat Rev Drug Discov* 13:105–121.

Lipinski CA, Lombardo F, Dominy BW & Feeney PJ (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* 23:4–25.

Shultz MD (2019) Two decades under the influence of the rule of five and the changing properties of approved oral drugs. *J Med Chem* 62:1701–1714.

Veber DF, Johnson SR, Cheng H-Y et al. (2002) Molecular properties that influence the oral bioavailability of drug candidates. *J Med Chem* 45:2615–2623.

Wenglowsky S, Moreno D, Rudolph J et al. (2012) Pyrazolopyridine inhibitors of B-RafV600E. Part 4: an increase in aqueous solubility via the disruption of crystal packing. *Bioorg Med Chem Lett* 22:912–915.

Bioisosteres, Peptidomimetics, and Scaffold Hopping

Bursavich MG & Rich DH (2002) Designing non-peptide peptidomimetics in the 21st century: inhibitors targeting conformational ensembles. *J Med Chem* 45:541–558.

Eguchi M & Kahn M (2002) Design, synthesis, and application of peptide secondary structure mimetics. *Mini Rev Med Chem* 2:447–462.

Hu Y, Stumpfe D & Bajorath J (2017) Recent advances in scaffold hopping. *J Med Chem* 60:1238–1246.

Langdon SR, Ertl P & Brown N (2010) Bioisosteric replacement and scaffold hopping in lead generation and optimization. *Mol Inf* 29:66–385.

Meanwell NA (2011) Synopsis of some recent tactical application of bioisosteres in drug design. *J Med Chem* 54:2529–2591.

Patani GA & LaVoie EJ (1996) Bioisosterism: a rational approach in drug design. *Chem Rev* 96:3147–3176.

Pirali T, Serafini M, Cargnin S & Genazzani AA (2019) Applications of deuterium in medicinal chemistry. *J Med Chem* 62:5276–5297.

Subbaiah AM & Meanwell NA (2021) Bioisosteres of the phenyl ring: recent strategic applications in lead optimization and drug design. *J Med Chem* 64:14046–14128.

Sun H, Tawa G & Wallqvist A (2010) Classification of scaffold hopping approaches. *Drug Discov Today* 17:310–324.

Metabolic Stability

Kerekes AD, Esposite SJ, Doll RJ et al. (2011) Aurora kinase inhibitors based on the imidazo[1,2-*a*]pyrazine core: fluorine and deuterium incorporation improve oral absorption and exposure. *J Med Chem* 54:201–210.

St Jean DJ Jr & Fotsch C (2012) Mitigating heterocycle metabolism. *J Med Chem* 55:6002–6020.

Fluorine

Gillis EP, Eastman KJ, Hill MD et al. (2015) Applications of fluorine in medicinal chemistry. *J Med Chem* 58:8315–8359.

Losartan

Duncia JV, Chiu AT, Carini DJ et al. (1990) The discovery of potent nonpeptide angiotensin II receptor antagonists: a new class of potent antihypertensives. *J Med Chem* 33:1312–1329.

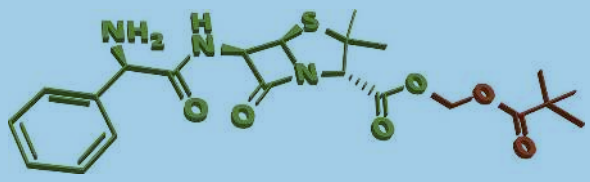
Larsen RD, King AO, Chen CY et al. (1994) Efficient synthesis of losartan, a nonpeptide angiotensin II receptor antagonist. *J Org Chem* 59:6391–6394.

Remdesivir

Eastman RT, Roth JS, Brimacombe KR et al. (2020) Remdesivir: a review of its discovery and development leading to emergency use authorization for treatment of COVID-19. *ACS Cent Sci* 6:672–683.

Mackman RL, Hui HC, Perron M et al. (2021) Prodrugs of a 1′-CN-4-Aza-7,9-dideazaadenosine C-nucleoside leading to the discovery of remdesivir (GS-5734) as a potent inhibitor of respiratory syncytial virus with efficacy in the African green monkey model of RSV. *J Med Chem* 64:5001–5017.

Siegel D, Hui HC, Doerffler E et al. (2017) Discovery and synthesis of a phosphoramidate prodrug of a pyrrolo [2,1-f][triazin-4-amino] adenine C-nucleoside (GS-5734) for the treatment of ebola and emerging viruses. *J Med Chem* 60:1648–1661.



The Process of Developing a Drug from an Optimized Lead

5

LEARNING OBJECTIVES

- Develop an understanding of the historical evolution of the drug approval process.
- Summarize the different types of patents important for drug discovery and development.
- Identify the characteristics of process chemistry synthesis.
- Become familiar with the pharmacokinetic parameters that are evaluated during the drug discovery and development stage, and how they are optimized.
- Understand some of the toxicity testing undertaken during the drug development stage.
- Summarize aspects of formulations that are critical for a drug development project.
- Learn the various phases of clinical trials and what is evaluated in each.

5.1 INTRODUCTION

Historically, the process of new drug evaluation and approval has evolved so that drugs are proven to be both safe and effective in order to comply with increasing levels of regulatory oversight. In the United States, the first regulations imposed were decreed by the Pure Food and Drug Act of 1906, which required accurate labeling of medicines and listing of "addictive and/or dangerous" substances, such as morphine, cannabis, and alcohol. The issue of safety was addressed further in 1938 with the passage of the Food, Drug, and Cosmetic Act. This law came about in response, in part, to a tragedy when approximately 100 people died in 1937 as a result of taking a diethylene glycol preparation of sulfanilamide, a drug used to treat streptococcal infections. Although there had been isolated reports of diethylene glycol's toxicity, it was not widely known at that time. The new act required testing of medicines in animals in order to prove safety. However, it was not necessary to prove efficacy of the drug until the Kefauver–Harris Amendment (Drug Efficacy Amendment) was added over 20 years later, in 1962. This amendment required clinical trials establishing efficacy in humans and mandated reporting serious side effects of a drug even after its approval and marketing. It also increased the power of the US Food and Drug Administration (FDA) by requiring that it approve a drug before the drug could be marketed. This amendment was enacted to some extent in response to the thalidomide tragedy.

DOI: 10.1201/ 9781003465607-6

Thalidomide was widely prescribed in the late 1950s in Europe and Canada as both a sedative and a treatment for morning sickness. Sadly, it was later realized to have caused serious birth defects in more than 10,000 children whose mothers had taken it. It is widely believed that more thorough testing (perhaps mandated by government agencies) prior to thalidomide's widespread use could have prevented this tragedy.

Today, there are two broad stages for developing a new drug. The discovery stage, discussed in prior chapters, is focused on hit and lead discovery and lead optimization to obtain a candidate that is predicted to be suitable for human testing. **Drug development** is the second stage of the process and consists of acquiring sufficient data to support the approval of the drug. It differs from the discovery stage in that all efforts are focused on characterizing and testing the one (or few) molecule that has evolved from the discovery stage and that exhibits all the appropriate pharmacological effects and desirable physical and pharmaceutical properties while, at the same time, being predicted to be safe. In the simplest cases, these two processes occur sequentially; however, most often, some of the specific activities overlap, and data generated during the development stage can often impact the direction of the discovery stage. Activities during this stage are numerous. One early activity, which may begin in the discovery stage, is patenting, which will give the developer exclusivity when marketing a drug. Process chemistry development concentrates on developing an efficient, safe, and cost-effective method to consistently prepare pure drug substances in large quantities. Pharmacokinetic studies evaluate how the body handles the drug and are used to predict human doses. Toxicology studies evaluate the safety of the drug in multiple species to give confidence of its safety in humans. Formulation development creates and tests combinations of the drug molecule and excipients and identifies the final form of the drug (capsule, tablet, injection, etc.) that will be used in patients.

Results of these studies, plus a plan for clinical evaluation, are required in order to file documents, such as an Investigational New Drug Application (IND) required by the US FDA, with regulatory agencies. In the United States, 30 days must pass after the submission before human (clinical) trials can be initiated. Clinical trials consist of four phases, and the first three take approximately 3–7 years for completion. During Phase I, Phase II, and Phase III clinical trials, the drug is tested for safety and efficacy, as well as for its potential to interfere with other drugs. It is also compared to other medications for the same disorder. In the United States, at the end of these three trials, a New Drug Application (NDA) is filed for review by the FDA. This review may take 1–2 years before final approval of the drug. An outline of the entire discovery and development process is shown in Figure 5.1.

In an industrial setting, the discovery stage typically lasts 1–3 years, and the development stage continues for many years beyond that. The odds of a compound proceeding from hit and lead discovery to an approved drug have been approximated at about 1 in 10,000. According to some estimates, the process of drug discovery and development for one drug takes an average of 12 years and costs up to $2.3 billion.

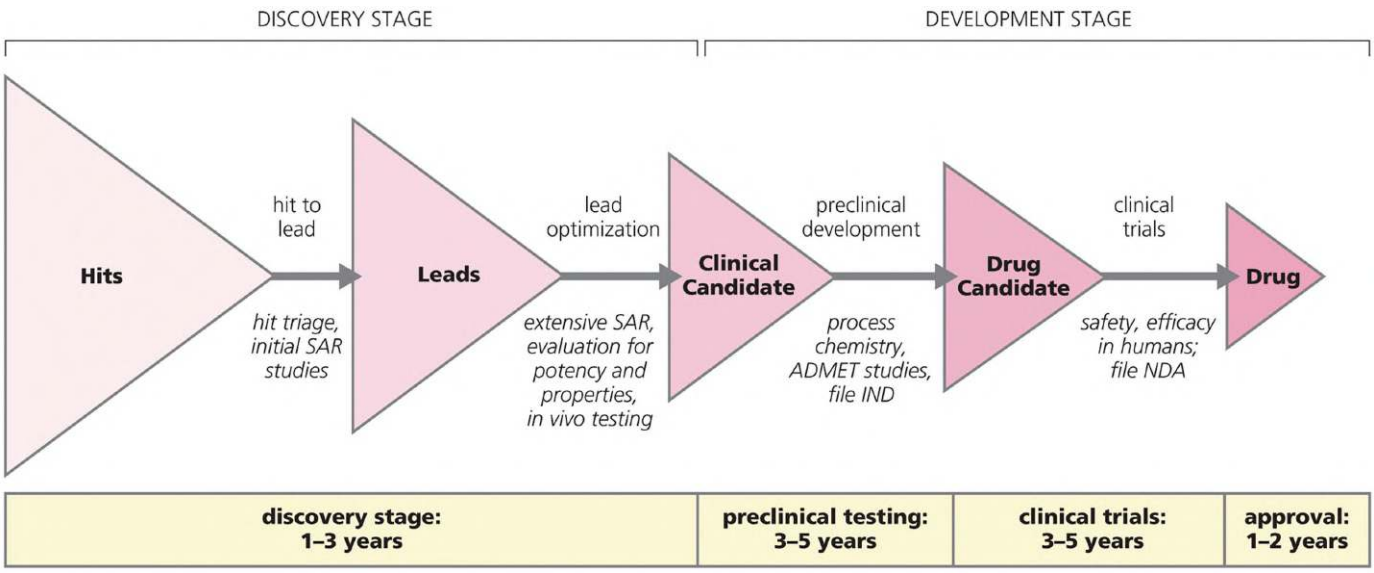


Figure 5.1 Overview of the drug discovery and development process. The discovery stage includes hit discovery, hit-to-lead activities, and lead optimization and results in the identification of a clinical candidate. This process usually takes between 1 and 3 years. The development stage focuses on the most promising compound(s). Clinical candidates are evaluated in extensive toxicity studies, efficient chemical synthesis and formulations are developed, and an IND is filed to advance drug candidates to clinical trial. At each stage, fewer and fewer compounds progress.

5.2 PATENTING AND DRUG DISCOVERY AND DEVELOPMENT

For all practical purposes, only commercial organizations have the capacity and resources necessary to complete the entire drug development process. Therefore, given the very high cost of drug development, protecting intellectual property through **patenting** is a high priority and often is a prerequisite for the molecule being advanced through clinical trials. A *patent*, according to the US Patent and Trademark Office (USPTO), is an intellectual property right granted by the government of the United States of America to an inventor "to exclude others from making, using, offering for sale, or selling the invention throughout the United States or importing the invention into the United States" for a limited time in exchange for public disclosure of the invention when the patent is granted. In other words, it gives the holder (or licensee) of the patent the exclusive right to sell a drug for a specified purpose in the country where the patent is held. This right allows the patent holder (or its licensee) the opportunity to recoup the investment made in the lengthy and expensive drug discovery and development process, as well as to make a profit. In some cases, a patent is held by one organization (such as a university) and is licensed by a commercial entity (licensee) to complete the drug development process. After a patent expires, the drug can be sold by anyone once studies showing bioequivalence of the product have been performed and regulatory agencies approve the application. Drugs that are sold by commercial organizations other than the original patentor or licensee, after patent expiration, are termed generic versions of the drug.

Producing and marketing **generic drugs** is a very active and lucrative business, in which many companies specialize. Generic firms do not need to support drug discovery research efforts, and therefore, their costs of developing a generic drug are much less than those of the pharmaceutical firms who develop new drugs.

A number of requirements must be met for a patent to be granted. The invention must be able to be used (i.e., not a theory), a clear description of how to make and use the invention must be provided, the invention must be new, and the invention must not be obvious. The patenting process usually takes between two and five years and is concurrent with drug discovery and early preclinical development stages. In the United States, the process starts by submitting some version of a patent application to the USPTO. The application contains claims of the invention that are reviewed by an examiner. Over the course of several years, the examiner's office will often request changes or even reject certain claims based on the examiner's analysis of the literature and other patents. Eventually, a patent may be awarded, but it is often very different from the application originally submitted. Today, in the case where multiple patents are filed on the same subject matter, virtually all countries recognize the first to file as the inventor, as opposed to the individual(s) who were the first to invent it. Once a patent is conferred, the government office is not responsible for enforcing the patent; that responsibility falls to the patent holder.

5.2.1 Three Versions of Patents Are Commonly Filed during the Drug Discovery Process

In the context of drug discovery and development, the most comprehensive kind of patent is termed a **composition of matter patent**. This protects the specific compounds claimed in the patent. As described earlier, a compound must be novel (not previously disclosed to the public), have proven utility, and "not be obvious to one skilled in the art." In other words, if similar compounds have been previously described as having the same or related utility, it would be obvious to one skilled in the art that the new compounds should have the same activity and therefore would not be patentable. The patent can cover a specific chemical entity and/or structurally related series with multiple modifications specified. The patent will also include a description of syntheses of the compounds claimed, physical data that characterize them to prove they were indeed prepared, and usually preliminary data (i.e., biological data) that support the utility of the compound(s).

A second type of patent important for drug discovery and development is a **method of use patent**. These patents are filed on known compounds for which a new, unobvious use has been discovered. For example, in Chapter 2, clinical observation

was one of the lead sources described, with minoxidil as an example. Upjohn originally patented this compound as a blood pressure drug. When clinical observations showed minoxidil's ability to promote hair growth, they filed and were awarded a method of use patent for it as a hair growth promoter. The granting of this new patent extended the time that Upjohn could exclusively market minoxidil as a hair growth promoter.

A third type of patent is commonly referred to as a **process patent**. These patents usually describe the synthesis that will be used to manufacture the drug, as well as other possible syntheses. These syntheses can differ significantly from the one that may have been included in the original composition of matter patent. Process patents prevent others from using the protocols described in the owner's patent and are an additional method to strengthen the intellectual property estate surrounding a drug. A process patent can prevent anyone except the patent holder from manufacturing the compound by the same methods claimed, regardless of who holds the patent for use of the compound.

The lifetime of US and most international patents is 20 years from the time of the application. This may seem like a long time, until you consider that the average time to bring a drug to market is 12 years. This leaves only about eight years for an organization to try to recover the enormous investment that was expended during the discovery and development stages. Deciding when to file patent applications, how to structure the application, and in what countries to file is a complex, strategic process. Applications that cover a large number of compounds (real and virtual) are often weaker than those that cover just a few compounds. However, patenting usually occurs early in the drug discovery and development process, well before the exact drug candidate is identified, so including multiple potential drug candidates is necessary. There is careful consideration as to when to file and what to file. Factors involved in the decision include the timing for exclusivity and when any product might be marketed, as well as the competitive landscape. In addition, fees to file and maintain patents can be very significant, and therefore, cost is also a factor.

5.3 PROCESS CHEMISTRY RESEARCH

Process chemistry research involves the development of a process to synthesize multi-kilogram quantities of a drug candidate for animal testing and eventually for human testing. The synthetic protocols developed during this process are typically very different from the synthesis used to prepare the molecule during lead optimization, when only small quantities are necessary. The criteria for acceptable synthetic procedures require the process to be safe, efficient, highly reproducible, and consistent in yield and impurity profile. It should also be cost-effective and capable of generating kilogram quantities of drug candidate product with a minimal number of operations. Furthermore, there is now considerable effort to incorporate principles of green chemistry into these processes. **Green chemistry** focuses on the design of chemical products and processes that reduce waste, conserve energy, and replace hazardous substances. At the latter stages of the drug development process and post-marketing, the production of drugs is monitored and regulated by government authorities, who ensure that a set of standards called current Good Manufacturing Practices (cGMP) protocols are followed.

During process chemistry development, high-throughput experimentation (HTE) relying on design of experiments (DOE) is carried out to rapidly optimize reaction conditions. In addition, it is typical to use advanced computational methods, such as machine learning, 2D and 3DQSAR methods, as well as artificial intelligence for process design, reaction optimization, and catalyst design. Two examples of the successful application of these techniques are shown in **Figure 5.2**. One is the invention of a catalyst using advanced computational chemistry that stereoselectively produces a single enantiomer at phosphorus for a ProTide drug candidate, MK-3682/uprifosbuvir (see Chapter 4 Case Study 2 for examples of ProTides). Another is the use of machine learning to optimize a phase transfer catalyst for an aza-Michael reaction to synthesize letermovir, an approved drug to treat cytomegalovirus infection. Other important aspects of process chemistry research include the use of directed evolution

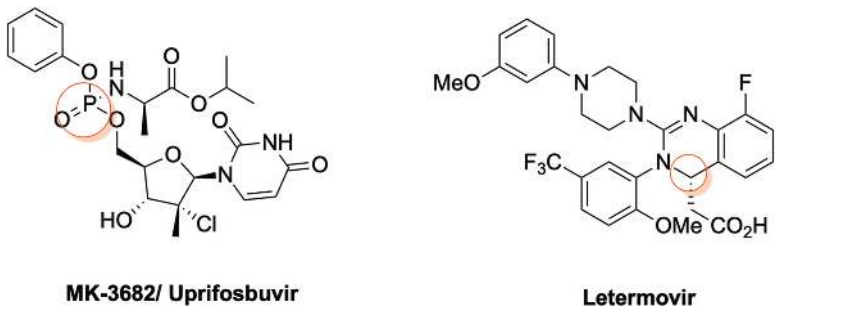


Figure 5.2 Structures of MK-3682/ uprifosbuvir and letermovir. Red circles indicate bonds formed based on the application of advanced computational methods/ML/AI.

to generate enzymes that can efficiently catalyze key steps, particularly those that involve installation of asymmetric centers, as highlighted in what follows in the synthesis of sitagliptin.

An excellent example of the differences between an early medicinal chemistry synthesis and a production-scale process can be highlighted by work on the anti-diabetic drug sitagliptin (Januvia). The original synthesis, first reported in 2005, was convergent, where each half of the molecule was synthesized and the halves (red and blue) were coupled to give the final product (**Figure 5.3**). Although fairly efficient and acceptable for the production of gram quantities, this route required 19 steps. Several reagents that are incompatible with a kilogram-scale synthesis due to safety issues were involved, one of which was diazomethane, used in the Arndt–Eistert homologation step. Chirality of the β-amino acid product originated from an α-amino acid that required four steps to prepare via alkylation of a valine-derived Schöllkopf reagent. The synthesis also relied on several expensive reagents, like 1-ethyl-[3–(3-dimethylamino)propyl]carbodiimide (EDC) used to couple the two halves, and silver benzoate (AgOBz), as well as multiple chromatographic steps.

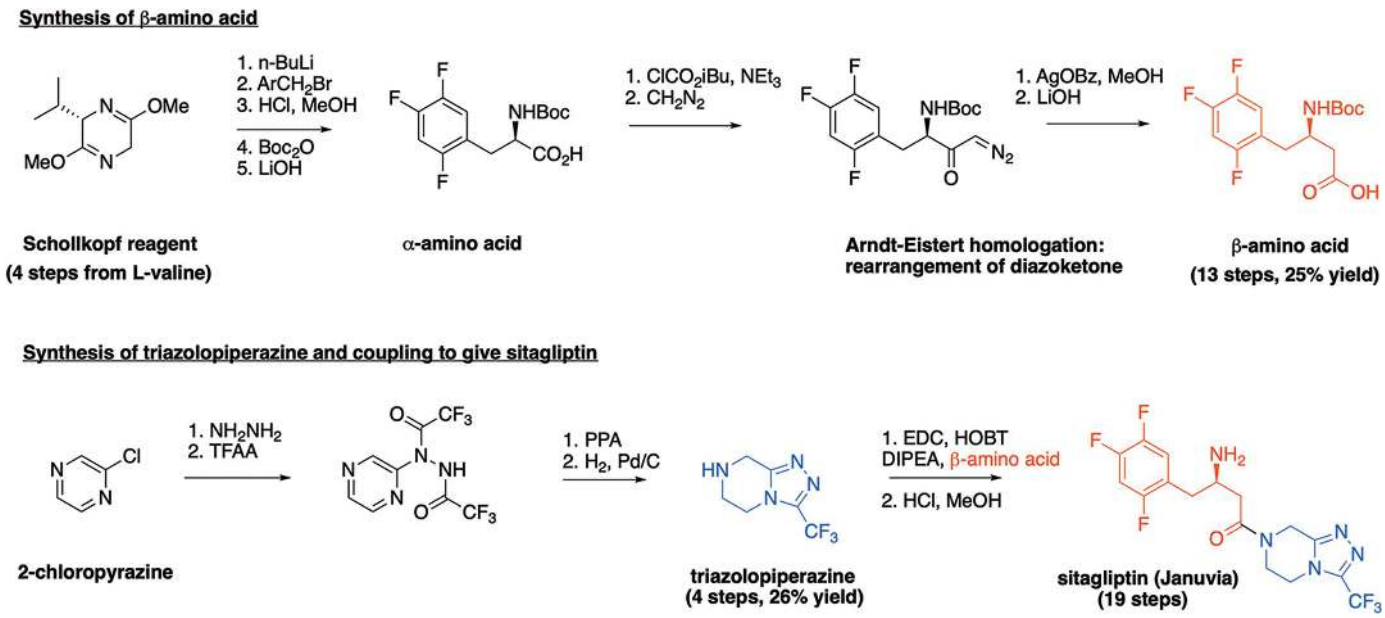


Figure 5.3 Originally reported synthesis of sitagliptin. *Top*, outline of synthesis of the β-amino acid component in nine steps from a valine-derived Schöllkopf reagent. Alkylation and release of the chiral auxiliary, followed by N-protection and ester hydrolysis, afforded the N-protected α-amino acid. Arndt–Eistert homologation, involving conversion to the acid chloride and reaction with diazomethane, followed by treatment with silver benzoate, and hydrolysis yielded the β-amino acid (red). *Bottom*, outline of the synthesis of triazolopiperazine and final coupling steps. Substitution of 2-chloropyrazine with hydrazine and N-acylation is followed by acid-catalyzed cyclization with polyphosphoric acid (PPA) and reduction to give the triazolopiperazine (blue). Amide formation between the β-amino acid (red) and triazolopiperazine (blue), followed by N-deprotection, afforded sitagliptin (Januvia). The sequence entails 19 steps.

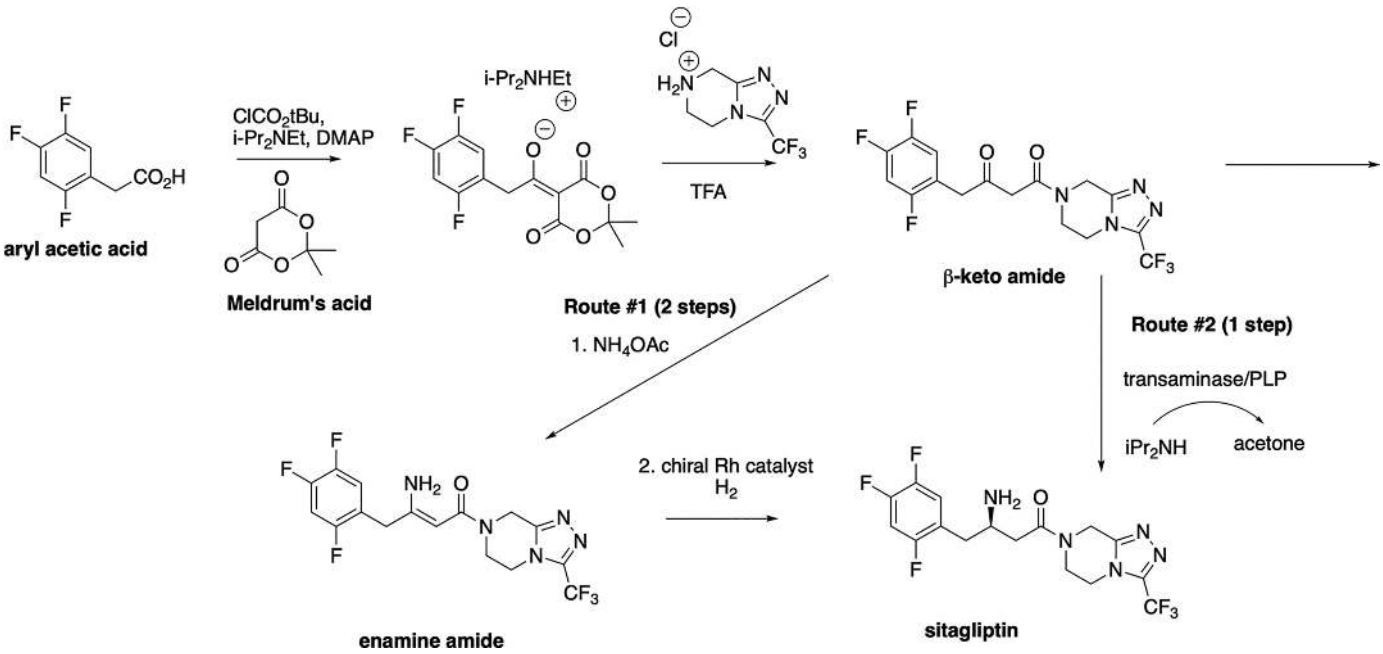


Figure 5.4 Process chemistry syntheses of sitagliptin. Acylation of Meldrum's acid is followed by coupling with the triazolopiperazine to give a β-keto amide. Route 1 uses NH_4OAc to give an enamine amide; reduction of the enamine by use of a chiral rhodium catalyst affords sitagliptin with 95% ee. This entire sequence entails 4 linear steps and proceeds in 65% overall yield. Alternatively, route 2 uses an engineered transaminase to proceed in one step from the β-keto amide to sitagliptin in 92% yield, with 99% ee.

Two routes were developed for a process synthesis (Figure 5.4), both proceeding through the same β-keto-amide. This β-keto-amide was generated by homologation of an aryl acetic acid using Meldrum's acid, followed by amide coupling with the triazolopiperazine (prepared separately). From this point, Route 1 involves two more steps: formation of the crystalline enamine amide and hydrogenation with an asymmetric rhodium catalyst to introduce the chirality in the final step with 95% enantiomeric excess. This extremely efficient process eliminated hazardous reagents and minimized both expensive coupling reagents and chromatography. In the final hydrogenation step, 95% of the chiral rhodium catalyst is recovered; this process generates 200 lb less waste per pound of sitagliptin manufactured than the original synthesis. A second route (Route 2) avoided the metal catalyst and proceeded directly from the β-keto-amide to sitagliptin, using an engineered transaminase enzyme with pyridoxal phosphate (PLP) as a coenzyme for the reductive amination. This final step afforded sitagliptin with 99.9% ee in 92% yield. It is one step shorter than Route 1 and eliminates the high-pressure Rh-catalyzed reduction that required the removal of trace amounts of rhodium in the final product. This example highlights some common strategies used in developing a process chemistry synthesis. These include using a minimal number of steps, employing inexpensive reagents, avoiding chromatography by utilizing crystallization for purification, and avoiding especially hazardous reagents. This process was recognized with a 2006 Presidential Green Chemistry Challenge Award for Greener Synthetic Pathways.

5.4 CHARACTERIZATION OF PHARMACOKINETIC PROPERTIES OF LEADS

During the drug discovery stage, experiments focus on effects of the molecule on protein targets (usually using *in vitro* assays) and in animal disease models (*in vivo*). However, unless appropriate concentrations of drug are present in the target tissue or organ, the drug will not be effective. The correlation between drug concentration at the site of action and effect (often efficacy) is termed **pharmacodynamics** (PD). An equally important consideration during the drug discovery and development process concerns how the body affects the drug after administration, or **pharmacokinetics** (PK). Four main processes contribute to the pharmacokinetics of a molecule: absorption, distribution, metabolism, and elimination (**ADME**). The safety of a drug is determined by **toxicity studies** in multiple animal species, and together these five components are called **ADMET** studies.

- **Absorption** is the mechanism by which the drug gets into the bloodstream. The specific route of administration directly affects this parameter.
- **Distribution** is the process through which the drug is distributed from the bloodstream to various compartments of the body, including tissues, brain, and other organs. This parameter is important to evaluate whether the drug reaches the specific organ or tissue that is targeted.
- **Metabolism** is the chemical modification of a substance in the body. This process typically occurs in the liver and generates more polar derivatives of the drug, called **metabolites**, often through oxidation or conjugation with small polar molecules. Metabolism is one aspect of the body's sophisticated and efficient method to remove toxic and/or foreign substances, termed **xenobiotics**.
- **Elimination**, or excretion, is the second part of the body's detoxification process. The intact drug or its metabolites are removed from the bloodstream, usually in urine or feces, but also through other modes, such as sweat or tears, and even exhalation.

The pharmacokinetics of a drug involves a series of complex processes; specific measurements corresponding to each of these are evaluated in both primary and secondary *in vitro* assays as well as *in vivo* assays. *In vitro* assays are incorporated early in the course of drug discovery and medicinal chemistry activities and are used to predict *in vivo* ADME results. For example, lipophilicity, permeability, and solubility measurements are used to predict absorption, and tests in isolated liver microsomes are used to evaluate how rapidly molecules are metabolized *in vivo*. Studies in animals are more complex and usually start with rodents and progress to animals such as dogs or rabbits, and then to primates. These animal studies, grouped under the umbrella term of *PK studies*, measure blood levels of the drug over a specific period of time and determine how the drug is metabolized, where it is distributed in the body, and how it is excreted. Data from these analyses allow comparison across several species and give an indication of what doses will be effective in humans.

5.4.1 Absorption Is One Determinant of How Much Drug Reaches Its Biological Target

Absorption is the process by which drugs pass through membranes, most often those in the gastrointestinal tract, into the bloodstream. For the sake of this discussion, the absorption of drugs administered orally (PO, or *per os*, by mouth) will be described. When a drug is ingested, it first passes through the highly acidic environment of the stomach (pH = 1–2.5). Only acidic drugs that have the appropriate balance of solubility and lipophilicity can be absorbed at this site. Acid-sensitive compounds will decompose here, and basic drugs, because they will be protonated in the acidic environment, will not be absorbed due to their charged nature. However, since the residence time in the stomach is only 20–60 minutes, there is relatively little absorption at this site. The second phase of absorption occurs in the small intestine, where the pH gradient moves from a partially acidic environment (pH ~5) to a neutral one (pH ~7) as the distance from the stomach increases. Most drugs are absorbed as they pass through the small intestine, due to its large surface area and the 4- to 6-hour transit time. The large intestine (pH 8–8.5) has a smaller surface area, and absorption here is less efficient. Upon absorption into the intestines, the drug passes through the portal vein into the liver (called the first pass), and then into general circulation, where it can be carried to the target tissue or directly eliminated in the feces. The overall process is outlined in **Figure 5.5**.

Drugs delivered via other routes of administration, such as intravenous (IV), intramuscular (IM), and transdermal (through the skin) delivery, enter circulation without first passing through the intestines and liver. Many routes of administration are possible, and each has its own absorption characteristics. Others include inhalation, sublingual (under the tongue), intraperitoneal (IP), rectal, and buccal (through the mucous membranes in the mouth) delivery.

As described in Chapter 4, three characteristics of a molecule are known to dramatically influence absorption: its solubility, lipophilicity, and permeability. Molecules must be soluble in their aqueous environment for absorption to occur, and polar molecules

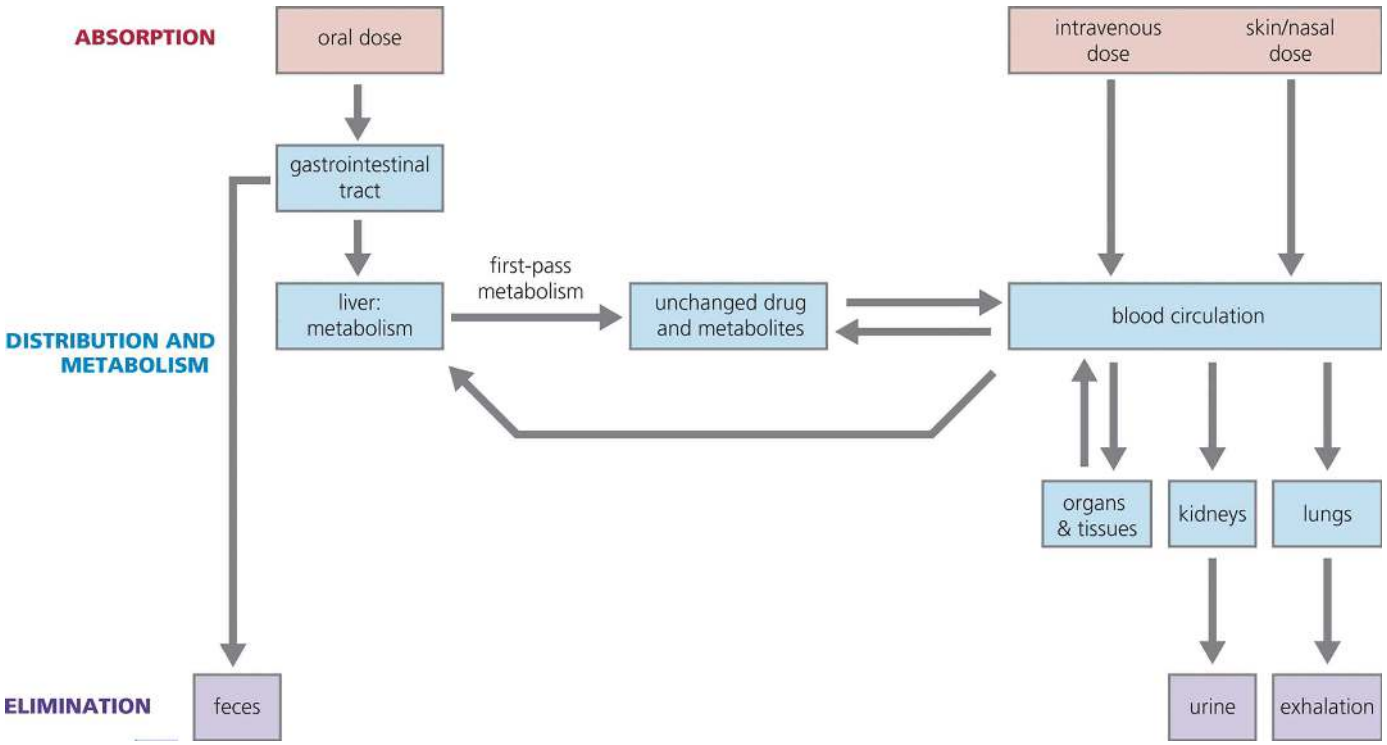


Figure 5.5 General process of absorption, distribution, metabolism, and elimination of drugs. An orally dosed drug passes through the gastrointestinal tract, where it may either be directly eliminated in the feces or pass through the liver. In the liver, the drug can be metabolized (called first-pass metabolism), and those metabolites and any unmodified drug enter into general circulation. An intravenously administered drug or one administered through the skin or nasal passage enters directly into general circulation (avoiding first-pass metabolism). Once in circulation by any route, the drug is distributed to various organs and tissues. The elimination of drugs occurs by metabolism in the liver to more polar derivatives, and then excretion in urine and feces, as well as other routes.

and those that contain charged functional groups favor excellent solubility. They must also be sufficiently lipophilic that they can pass through a cellular membrane, be it cellular membranes in the stomach and intestines that allow drugs to get into the bloodstream or, once in the bloodstream, the specific cell membrane that may contain the pharmacological target of the drugs. The permeability of a compound depends on both of these characteristics. Drugs permeate through cell membranes through two main routes. Most drugs rely on **passive diffusion**, driven by a concentration gradient, as a mechanism to enter cells. Molecules that are lipophilic, like the lipid membrane, or lipid-soluble are most effectively transported through the passive route, whereas charged molecules are not passively diffused. A second mechanism by which molecules can cross cell membranes is called **active transport**. This mechanism relies on protein transporters that recognize specific molecules, often biomolecules, such as amino acids or carbohydrates, but also some small-molecule drugs, and transport them across the lipid bilayer. This process is often difficult to predict; therefore, most drug discovery programs work on the assumption that molecules will be absorbed by passive diffusion.

In vitro assays to characterize solubility, lipophilicity, and permeability are used to predict absorption. To get a more complete assessment in an animal, **PK studies** are performed. They evaluate the concentration of the drug in plasma over a specific time frame after the administration of the compound at a certain dose (or doses). Since the dose administered is rarely reflective of the compound concentration in circulation due to all the processes outlined earlier, PK studies allow one to more fully understand the behavior of the compound in a living organism. They are an important bridge between the *in vitro* assays described in Chapter 4 and the results of studies in animal models of disease. By understanding the concentration of drug in circulation and other parameters after dosing, the behavior of the compound in animal models of disease becomes more interpretable, since one can then relate effects observed to a plasma concentration. As an example of the design of a single-dose PK study, animals are given one dose of a compound by the preferred route of administration. Concentrations (C)

of the drug are measured at various time points (t) over an extended period of time, often 24 hours, and then plotted vs. time. A typical PK graph is shown in **Figure 5.6**, plotting plasma concentration vs. time in hours. An intravenous (IV) dose will give a high concentration immediately and then decrease over time. An oral (PO) dose will exhibit low concentrations initially, then increase as the drug is absorbed and enters the bloodstream, and ultimately decrease again as it is eliminated.

A number of important parameters related to absorption are generated from these studies, including C_{max}, t_{max}, C_{ss}, AUC, and bioavailability, among others. $\mathbf{C}_{max}$ is the highest concentration the compound achieves after dosing, and $\mathbf{t}_{max}$ is the time that it takes to reach that concentration. Some *in vivo* effects require that a certain C_{max} threshold is reached, and therefore, this value is often a critical determinant for a compound to advance. Alternatively, some toxicities are the result of the compound's very high C_{max}, so this parameter must be monitored. $\mathbf{C}_{ss}$ is the steady state concentration of a drug, observed at equilibrium after a number of doses. This will give an indication of the drug concentration after a multiple-dose regimen and helps determine dosing in humans. The **area under the curve** (AUC) is calculated from the graph and reflects the total amount of drug absorbed or exposure of the drug. **Bioavailability** (or fraction absorbed, F, or FA) is also derived from a PK study and reflects the fraction of the molecule absorbed after a specific route of administration compared to what is theoretically possible. It is calculated by comparing the AUC when the drug is dosed by the preferred route (usually orally) to the AUC when dosed intravenously (considered 100% absorbed). Oral bioavailability is equal to AUC_{oral} divided by $AUC_{iv} \times 100$. High oral bioavailability values are >70%, with >50% preferred. Compounds with oral bioavailability of less than 50% may advance, but they often require additional formulation efforts due to their poor absorption characteristics.

Many different factors influence a drug product's PK properties. Some of these factors are intrinsic to the specific molecule and therefore cannot be changed. Others, however, such as formulations, can be modified to optimize parameters, such as C_{max}, exposure, and bioavailability. High bioavailability is an important goal for any drug discovery project for a number of reasons, including the following:

- *Safety.* The PK performance of a drug that exhibits excellent oral bioavailability is often less variable than those with poorer bioavailability. Therefore, unexpectedly high or low drug concentrations, which can result in over- or under-dosing, respectively, rarely occur. In addition, the lower doses that are possible when bioavailability is high may minimize off-target effects.
- *Cost.* Highly bioavailable drugs minimize the amount of active ingredient required in each drug substance, be it a tablet, capsule, or IV solution. The amount of active drug ingredient required will influence the cost of the materials of the drug as well as the cost to manufacture the drug.
- *Compliance.* Smaller doses (for example, tablet, pill sizes) taken less frequently are preferred by patients and improve compliance.

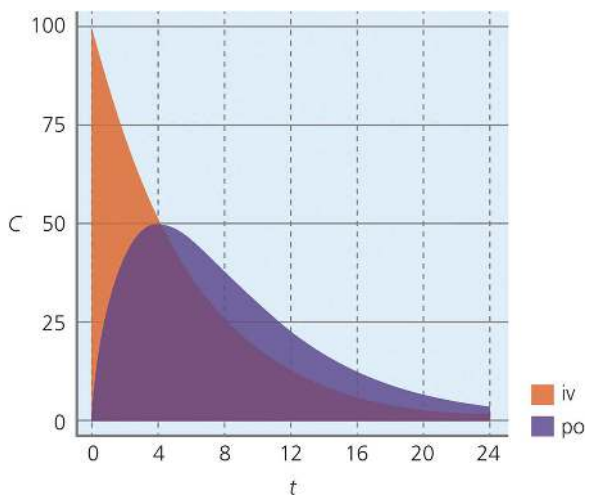


Figure 5.6 Graph of a theoretical PK study. Concentration (C) of drug measured over time (t, hours) after intravenous (IV, shown in orange) or oral (PO, shown in purple) administration. Various parameters are calculated from this graph: C_{max} is the peak concentration reached, t_{max} is the time required to reach C_{max}, *AUC* reflects overall exposure of the compound, and bioavailability is calculated as the ratio AUC_{oral}/AUC_{iv}.

5.4.2 Distribution Determines in Which Tissues Drugs Accumulate

The body has various compartments into which a drug can distribute. The blood is considered the first compartment; the second compartment is the rest of the body. Once in the bloodstream, drugs can further distribute and accumulate in specific tissues. The distribution may not be uniform and is influenced by characteristics of the drug and of the particular tissue. For example, drugs that accumulate in fat tissue are typically lipophilic. Differences in pH between blood and specific cells will also influence where drugs concentrate. Drugs can also build up in certain organs, such as the brain, for example, where central nervous system (CNS) drugs need to accumulate to be efficacious, or in tumors, where anti-cancer drugs must reach. Additional factors determining distribution include the ability of molecules to bind to plasma proteins or to be recognized by active transporters.

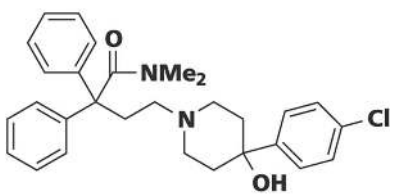
Drugs in circulation frequently bind to **plasma proteins,** such as albumin, to various degrees, and this phenomenon affects not only their distribution but also their *in vivo* potency. The unbound drug molecules are in equilibrium with plasma protein–bound drug molecules, and as the unbound molecules migrate into tissues, the equilibrium between bound and unbound drug in plasma is re-established. Therefore, drugs that are highly bound to plasma proteins usually are not widely distributed or only slowly distributed. Furthermore, when they are evaluated *in vivo*, highly bound drugs may not be as efficacious as molecules that are less tightly bound, because only the unbound molecule is available to bind to the target. On the other hand, highly bound drugs can act as a reservoir for free drug and therefore provide long-lasting concentrations of drug in plasma. The effects of a drug's binding to plasma protein are complex and difficult to predict. Assays to determine the extent of protein binding are often performed during lead optimization and can be done in plasma, tissue, or blood by dialysis, ultrafiltration, or ultracentrifugation methods. These secondary assays determine the **fraction unbound** (f_u) of compound and are often used to rank compounds. When the f_u value is low, it can be used to rationalize the unexpectedly poor behavior of a potent compound in *in vivo* animal models. An alternative way to assess the effects of plasma protein binding is to include serum (which contains those plasma proteins) in the original *in vitro* assay and determine if there is a shift in potency with and without serum.

A second determinant of drug distribution is the activity of transporter proteins, found in specific cells that recognize and transport distinct small molecules. When a small molecule is recognized by a transporter protein, it can be taken up by cells expressing that transporter and accumulate in those cells. Since transporters are responsible for both influx and efflux of compounds, they can also play a role in keeping molecules out of specific tissues. The most well-known example of an efflux pump is the P-glycoprotein transporter (P-gp) (also known as ABCB1), which functions by rapidly transporting substrate molecules of diverse structures out of cells against a concentration gradient. P-gp is widely expressed in the intestine, and its activity is one reason some drugs are poorly absorbed despite good solubility and permeability. It is also present in the brain, where it protects the CNS from xenobiotics and therefore is a powerful barrier to drugs reaching the brain. Loperamide is an over-the-counter treatment for diarrhea that works through the same mechanism as morphine: Both bind at the μ-opioid receptor. However, whereas morphine is brain-penetrant, loperamide is rapidly transported out of the brain by P-gp, making it safe and nonaddictive, in contrast to morphine (Figure 5.7). P-gp is also present in some cancer cells, where its expression or induction is a common mechanism of drug resistance caused by its transport of drugs out of the cancer cells. Assays to determine whether a compound is a substrate for an efflux pump, such as P-gp, or others, such as solute carriers (SLC) and breast cancer resistance protein (BCRP), are common during lead optimization.

The distribution of a drug affects its potency as well as its potential for toxicity. Certain drugs are highly effective and safe because they accumulate in the specific tissues, where they need to act. For example, iodine concentrates almost exclusively in the thyroid gland, and therefore, iodine and its derivatives are often used to treat thyroid disease. The apparent **volume of distribution** (V_d) is a value used to reflect the amount of drug that is distributed throughout the organism compared to the amount present in plasma; it is measured



morphine

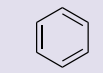
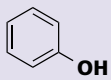
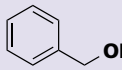
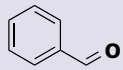
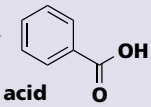
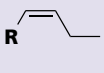
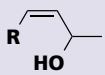
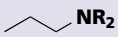
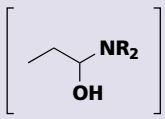
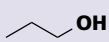
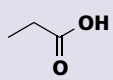
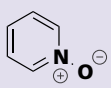


loperamide

Figure 5.7 Structures of morphine and loperamide. Both bind at the μ-opioid receptor, but in contrast to morphine, loperamide does not accumulate in the brain due to its efflux by the P-gp transporter and therefore has minimal CNS effects.

in a PK study. For molecules with high volumes of distribution, plasma levels will be low. Conversely, low volumes of distribution reflect high plasma concentrations. The specific disease targeted will dictate whether a high or low volume of distribution is desirable.

Table 5.1 **Oxidation Substrates and Products of Phase I Metabolism**

Substrate	Product
 arene	 phenol
 arene	 benzylic alcohol →  aldehyde →  acid
 alkene	 allylic alcohol
 carbonyl compound	 α-hydroxy carbonyl
 alkyl amine	 unstable hemiaminal →  aldehyde + amine
 alkane	 alcohol →  ketone
 alkene	 epoxide
 alcohol	 aldehyde →  carboxylic acid
 amine heterocycle	 N-oxide
 thioether	 sulfoxide →  sulfone

5.4.3 Metabolism and Elimination Are the Body's Method for Removing Drugs

One mechanism by which the body detoxifies itself is through metabolism, also termed *biotransformation*. Metabolism chemically modifies the xenobiotic, typically in two phases, I and II, to generate more polar, water-soluble derivatives called metabolites. Metabolites are more efficiently eliminated than the parent. Specific enzymes in the liver carry out most of these reactions, although metabolism also occurs in other organs and tissues. The degree of metabolism of a drug candidate and the specific metabolites formed affect its potency, dosing schedule, and toxicity. Molecules that are rapidly metabolized are then rapidly eliminated and never attain appropriate blood levels to exhibit the desired effects. Conversely, when drugs are metabolized very slowly, their concentration in the blood and in tissues can build up to unsafe levels. The specific metabolites of drugs need to be identified and studied, as they often exhibit their own pharmacological effects. In some cases, the metabolites are inactive; in other cases, they exert similar, beneficial effects as the parent; and in still other situations, the metabolites might exhibit some serious toxicity.

Phase I metabolic reactions usually occur in the liver and modify a functional group of the parent drug. Three of the most common Phase I reactions are oxidation, reduction, and hydrolysis. Oxidation is carried out chiefly by cytochrome P450 enzymes (CYPs), which are heme proteins associated with an NADPH-cytochrome P450 reductase. There are many different CYP variants (known as isozymes), each denoted with a three-figure notation of number-letter-number. The CYP isozymes responsible for most drug metabolism are CYP 3A4, CYP 2D6, CYP 1A2, CYP 2C9, and CYP 2C19. Additional enzymes that play a role in oxidative metabolism include alcohol and aldehyde dehydrogenases, aldehyde oxidase, amine oxidases, and lipoxygenases. The most common oxidation reactions are hydroxylation, oxidation, and epoxidation. Table 5.1 shows some

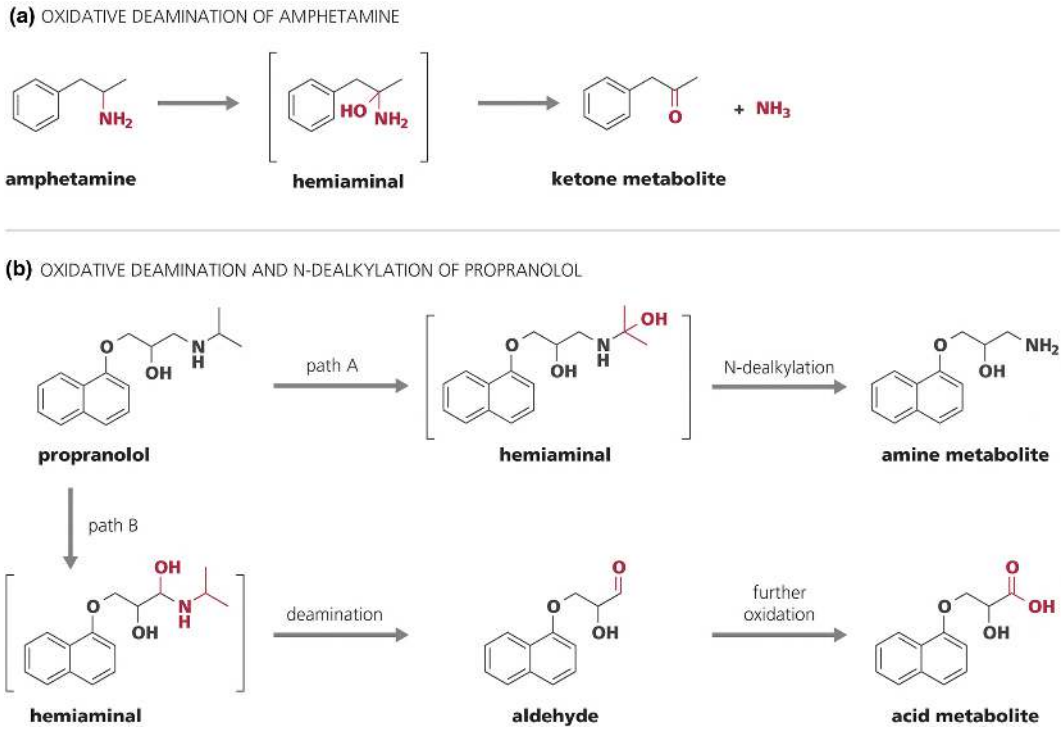


Figure 5.8 Oxidative N-dealkylation and deamination of amine-containing drugs. Fragments introduced and metabolized are shown in red. (a) Oxidative deamination of amphetamine gives a hemiaminal intermediate, which hydrolyzes to the ketone metabolite and ammonia. (b) Propranolol is metabolized by two oxidative pathways to give intermediate hemiaminals. Path A begins with hydroxylation of the isopropyl group, which is then converted to the N-dealkylated metabolite. Path B begins with hydroxylation of the methylene adjacent to the nitrogen; deamination affords an aldehyde, which is further oxidized to an acid.

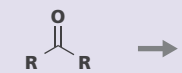
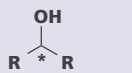
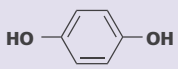
examples of typical oxidation processes with one-step processes on the left and multistep processes on the right.

Hydroxylation reactions can occur on arenes, at benzylic or allylic positions, at a position α to carbonyls, on carbons bearing a nitrogen, and even on unactivated alkyl groups. Epoxidation occurs on alkenes and also on arenes (not shown). Other oxidations include oxidation of alcohols and direct oxidation of nitrogen or sulfur to generate N-oxides and sulfoxides or sulfones, respectively.

Often, the first metabolite formed is also a substrate for the metabolizing enzyme, and secondary reactions occur. For example, when a primary alcohol is formed, it can undergo subsequent oxidation to the aldehyde and then to the carboxylic acid. The hemiaminal, $RC(OH)NR'_2$, formed from oxidation at a carbon α to a nitrogen atom, decomposes to give a carbonyl product and an amine product (NR'_2). Further reaction of the resulting carbonyl product is common. Some examples that follow this pathway are shown in Figure 5.8. Amphetamine undergoes oxidative deamination to give a ketone metabolite and ammonia. Propranolol, a β-adrenergic blocker, follows two pathways, based on the site of oxidation. Path A involves oxidation at the N-isopropyl site, followed by N-dealkylation, to generate the dealkylated primary amine metabolite and acetone (not shown). An alternative pathway is oxidative deamination at the internal carbon α to the nitrogen atom and further oxidation (path B), which ultimately gives a carboxylic acid.

Metabolic reductions (usually stereospecific) are carried out by α-keto reductases and flavin-independent NADPH-cytochrome P450 reductases, as seen in Table 5.2. Typical substrates are carbonyl-containing compounds (such as aldehydes and ketones, including quinones) and oxidized nitrogen compounds. Nitro, nitroso, or

Table 5.2 Reduction Substrates and Products of Phase I Metabolism

Substrate	Product
 ketone/ aldehyde	 chiral alcohol
 quinone	 hydroquinone
RNO_2 or RNO nitro/nitroso	RNHOH hydroxylamine
RNHOH hydroxylamine	RNH_2 amine
RN =NR azo	$2RNH_2$ amine
RX haloalkane	RH alkane

hydroxylamine groups are deoxygenated to generate either hydroxylamines or amines. Azo bonds are cleaved to produce two primary amines, and haloalkanes are typically reduced to alkanes. The stereoselectivity of the reduction of carbonyl groups may be species-dependent, as shown for the metabolism of naltrexone, an opioid antagonist (Figure 5.9a). Reduction of the ketone in chickens gives the *S*-alcohol, whereas reduction in either rabbits or humans gives the *R*-configuration at the carbinol. These results highlight an important caveat when animal data are used to predict human metabolism.

A third type of phase I transformation is a hydrolytic reaction, typically carried out on esters and amides by nonspecific esterases and amidases, respectively. The substrates may be aromatic or aliphatic, but hydrolysis may sometimes be highly chemospecific. For example, esterases hydrolyze the methyl ester (red) of cocaine in the presence of a benzoate ester (blue; Figure 5.9b). Hydrolysis may also be stereoselective for one enantiomer over another, such as the hydrolysis of the amide of prilocaine (Figure 5.9c). Both enantiomers of this local anesthetic are active, but only the amide (red) of the *R*-isomer is hydrolyzed to give the products toluidine and the *R*-carboxylic acid. *o*-Toluidine is toxic, so it is safest to administer only the *S*-isomer of prilocaine, which is not metabolized by this route.

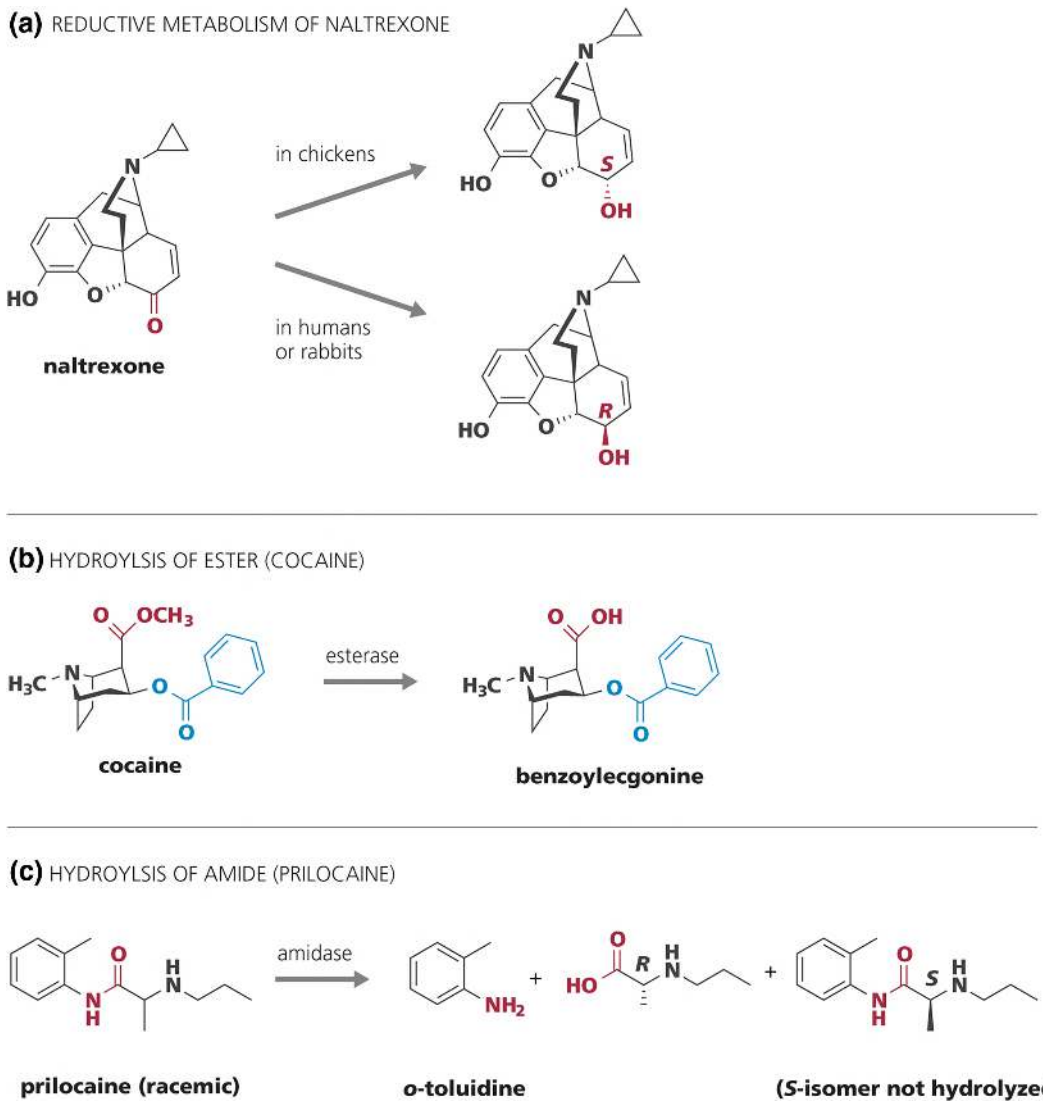


Figure 5.9 Examples of reductive and hydrolytic metabolism. (a) Reduction of the ketone of naltrexone affords the *S*-alcohol in chickens but the *R*-alcohol in rabbits and humans. (b) Hydrolysis of the methyl ester of cocaine (shown in red) occurs without hydrolysis of the benzoate ester (shown in blue). (c) Hydrolysis of the amide of prilocaine (shown in red) is specific for the *R*-enantiomer.

Prodrugs, discussed in Chapter 4, are administered as inactive forms of the drug and are metabolized, usually through phase I processes, to generate the active species. Esters are a common form of prodrugs, and hydrolysis by an esterase to the corresponding acid provides the active form. Examples, shown in **Figure 5.10**, include the antibacterial drug pivampicillin, an ester prodrug of ampicillin, and the blood pressure drug enalapril, an ester prodrug of enalaprilat. Biotransformation or metabolism played an important role in the history of the non-sedating antihistamine terfenadine (Seldane). Upon administration, it is rapidly oxidatively metabolized in the liver to produce another active molecule, fexofenadine. At high exposure levels, achieved when patients are also taking drugs or foods like grapefruit that inhibit cytochrome P450s, terfenadine is cardiotoxic, but fexofenadine is not. This observation led to the decision to replace the allergy drug terfenadine with fexofenadine (Allegra).

After phase I functional group transformations occur, there are further **phase II metabolic reactions**. The drug or metabolite is conjugated (connected) to some small endogenous molecule in order to expedite elimination or to reduce reactivity. These phase II reactions are carried out by various transferase enzymes. Some examples, illustrated in **Table 5.3**, are as follows:

- *Glucuronidation.* In this reaction, glucuronic acid is conjugated to alcohols, acids, and other nucleophilic groups. A functional group of the drug or metabolite attacks the anomeric carbon of UDP-glucuronic acid to give the glycoside, or conjugate, of the drug or one of its Phase I metabolites.

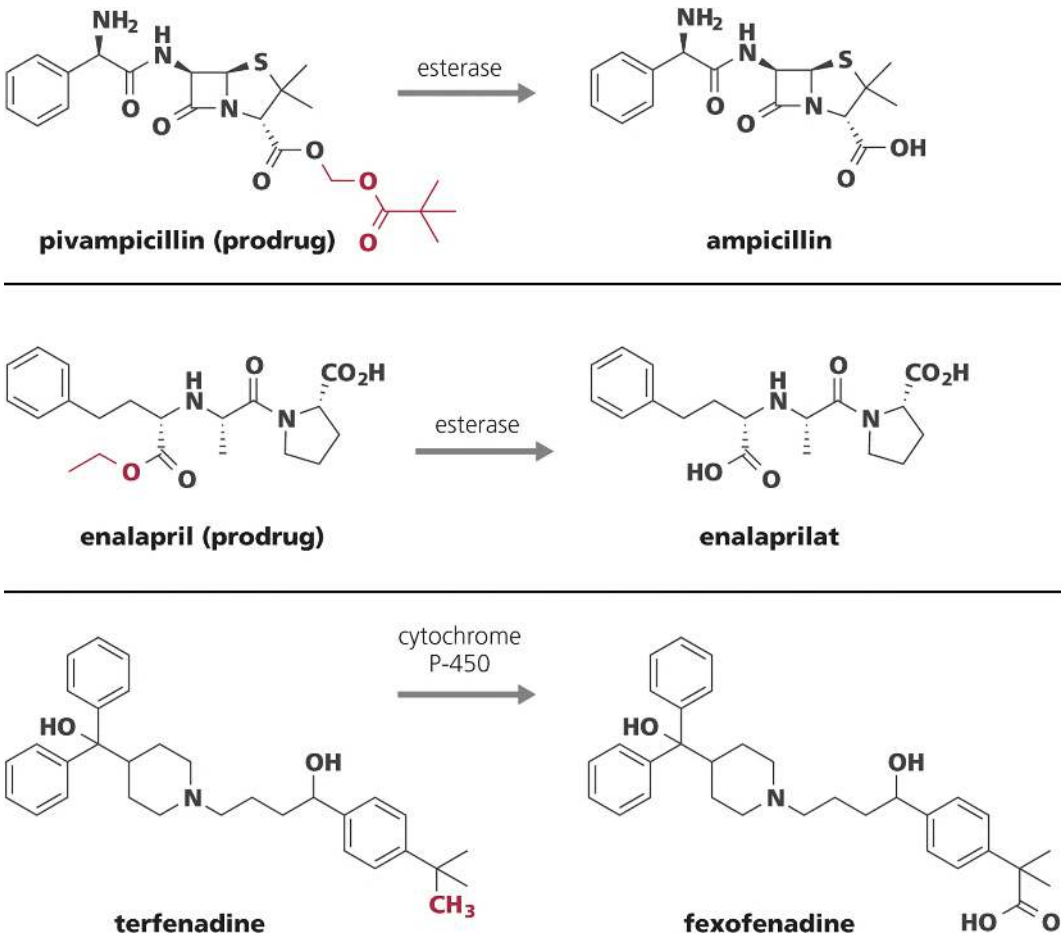
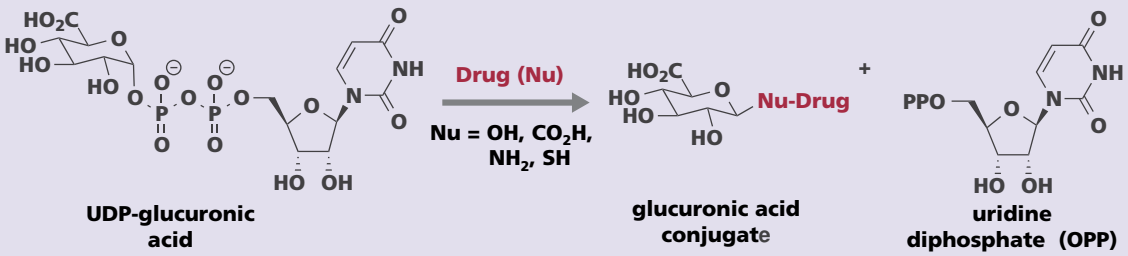
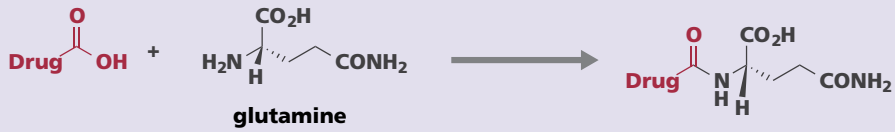
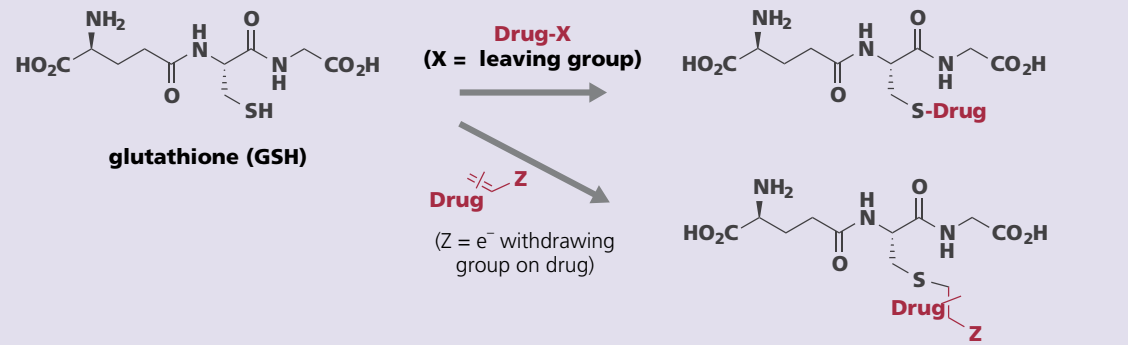


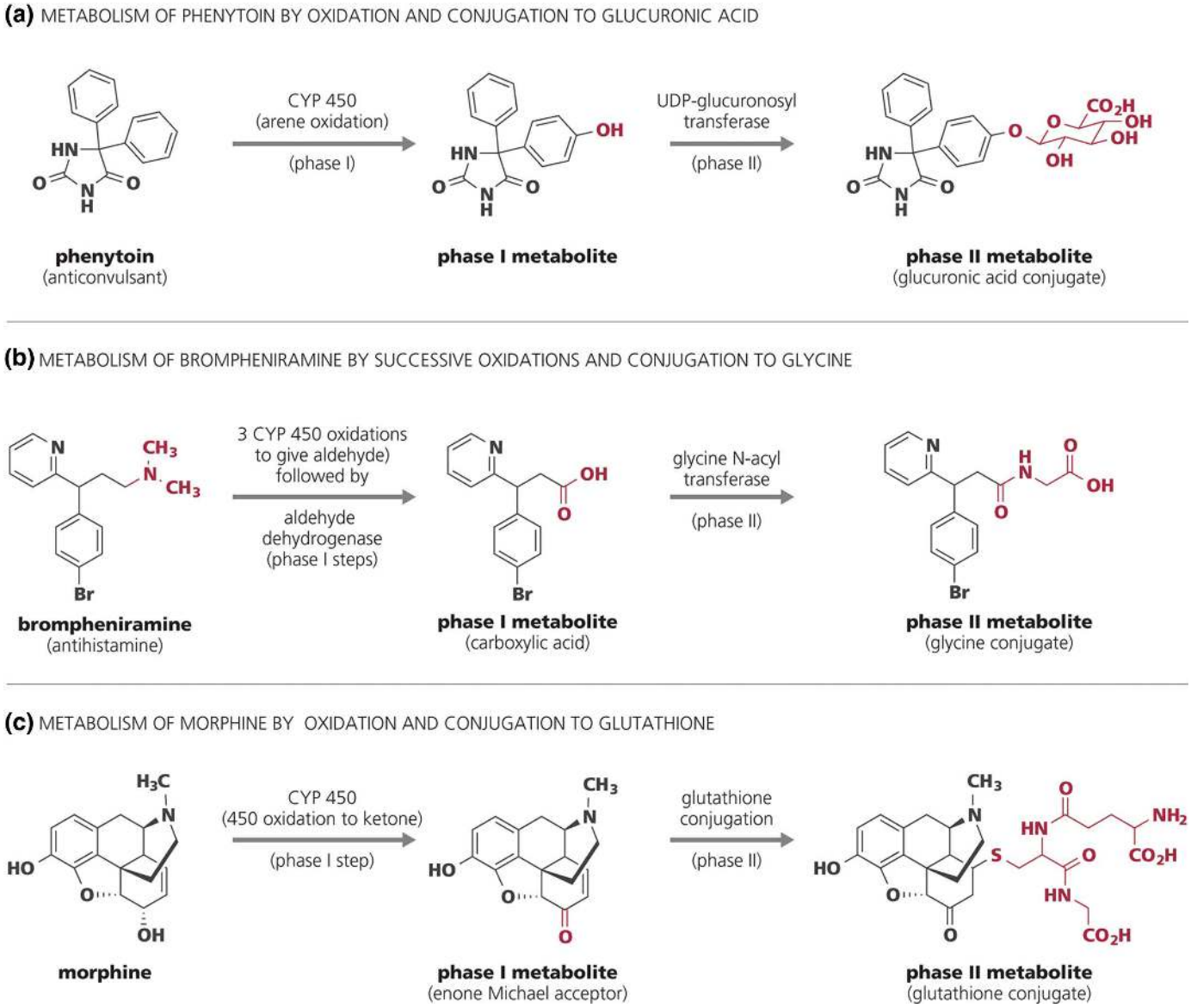
Figure 5.10 Examples of drugs generated by metabolism. The active forms of pivampicillin (ampicillin) and enalapril (enalaprilat) are formed by ester hydrolysis. Fexofenadine is a metabolite of terfenadine formed by cytochrome P450–mediated oxidation of a methyl group to a carboxylic acid.

Table 5.3 **Phase II Metabolic Conjugation Reactions**

Phase II transformation	Reaction
Glucuronidation: Nucleophile present in drug adds to anomeric carbon of UDP-glucuronic acid	
Amino acid conjugation: amide forms between acid in a drug and amine group of glutamine	
Glutathione conjugation: thiol of glutathione adds to electrophile present in a drug	
N-acetylation: acetyl group is transferred to nitrogen atom of a drug	

- *Amino acid conjugation.* An amide is formed between an amino acid (usually L-glutamine in primates) to a carboxylic acid of the drug or one of its metabolites.
- *Glutathione conjugation.* The thiol group of glutathione, a tripeptide found in all tissues, attacks a variety of electrophiles, such as aldehydes and Michael acceptors of the drug or drug metabolite, to provide the thiol adduct.
- *N-acetylation.* Amines are often conjugated during Phase II metabolism to generate the N-acetate derivatives.

Some examples of complete metabolic processes are shown in **Figure 5.11**. Metabolism of phenytoin, a drug to treat epilepsy, involves an oxidative phase I transformation, where one of the aromatic rings is hydroxylated at the para position. Phase II conjugation of that metabolite by glucuronidation gives a final metabolite. In the second example, the antihistamine brompheniramine proceeds through four different metabolic reactions to generate the carboxylic acid: two successive N-demethylations



generate the primary amine; oxidative deamination of that metabolite generates the aldehyde, which is further oxidized to the carboxylic acid metabolite. Phase II conjugation of the acid to the amine of glycine gives an amide as a final metabolite. Finally, the alcohol of the analgesic morphine is oxidized to a ketone, generating an electrophilic enone (Michael acceptor). Conjugation to glutathione affords a final metabolite.

During lead optimization, the potential for compounds to be metabolized is evaluated in an assay that measures the stability of the compound in liver microsomes or hepatocytes of multiple species, including human and animals in which the *in vivo* efficacy studies are performed (e.g., mouse, rat). A half-life is determined to indicate which compounds will be stable to metabolism and which are likely to be prone to rapid degradation. This data is often converted to a measure of *in vitro* intrinsic clearance (CL_{int}), described in greater detail later.

These assays report on Phase I metabolic processes only. During the drug development stage, metabolites of the drug candidates are identified and quantified through sophisticated mass spectrometric analysis. Metabolites that are formed in significant quantities are synthesized and tested in the same battery of assays as the parent to evaluate whether they exhibit pharmacological activity and/or toxicity. Predicting the identity of metabolites from the parent structure relies on understanding

Figure 5.11 Examples of drug metabolism showing both Phase I and Phase II metabolites. (a) Phenytoin undergoes phase I metabolism to a phenol and then phase II conjugation to the glucuronide. (b) Brompheniramine is converted by four successive phase I reactions to the acid, and then phase II amino acid conjugation affords the glycine conjugate metabolite. (c) Phase I oxidation of morphine generates an α,β-unsaturated Michael acceptor, which reacts with glutathione in a phase II reaction to generate the glutathione adduct as the metabolite.

the metabolic reactions and is fairly reliable. For example, alkyl chains and neutral or electron-rich aromatic groups are considered metabolic hot spots for hydroxylation. Hydroxyl groups and acids are often glucuronidated. Medicinal chemistry strategies are frequently applied in order to limit metabolism, either through introducing steric hindrance or blocking functional groups or isosteric replacement, as discussed in Chapter 4. Computational chemistry programs are available to predict both the sites and the extent of metabolism; however, predicting the extent of metabolism, in contrast to the identity of metabolites, is much less precise.

Key parameters that are extracted from *in vivo* PK studies that reflect metabolism include measurements of the clearance of a drug and its half-life. **Clearance** (CL) is the volume of plasma cleared of the drug per unit time and reflects not only metabolism but also excretion. High clearance rates indicate that the compound is rapidly eliminated from the body and may not have sufficient exposure for *in vivo* effects to be observed. The drug's **half-life** ($t_{1/2}$) is the time it takes for 50% of the drug to disappear from circulation. Like CL, $t_{1/2}$ informs the appropriate dosing schedule. Drugs with longer half-lives can be dosed less frequently than those with short half-lives. However, drugs with very long $t_{1/2}$ (>24 hours) may result in serious side effects.

The final step in the elimination of xenobiotics is **excretion**. Drug molecules are most often eliminated through urine or feces, but other routes are also known, for example, through saliva or exhalation. While some molecules are excreted unchanged, most are excreted as metabolites. Evaluating the route of excretion is done in animal models, by detecting the concentration or the presence of the parent or metabolite in urine or feces.

Prediction of *in vivo* PK properties begins in the drug discovery phase, as *in vitro* data is being collected. Methods that have been traditionally employed include INIVe (*in vitro* to *in vivo* extrapolation) and PBPK modeling (physiologically based pharmacokinetics) using *in vitro* data and early *in vivo* data to predict human pharmacokinetics. For example, hepatic clearance from the liver (CL_H) is typically extrapolated from *in vitro* incubation of the potential drug in a preparation of human liver microsomes or hepatocytes. The percent of drug remaining is measured over time ($T_{1/2}$) is used to calculate an *in vitro* intrinsic clearance value (CL_{int}). This is then extrapolated through calculations to an *in vivo* value in the liver (CL_H). These calculations take into account factors that are not captured in a half-life measurement, such as differences in assay cell count and liver cell count and predicted distribution in the liver. As with other aspects of drug discovery, AI methods are now being used to predict ADME properties.

5.5 DRUG TOXICITY DETERMINATION

According to Paracelsus, "[a]ll things are poison, and nothing is without poison; only the dose permits something not to be poisonous." This statement is particularly relevant to drug discovery and development, where identifying doses and conditions at which toxicity caused by a drug is observed, and then understanding the nature and cause of the toxicity, is essential. As the overall goal of drug discovery and development is to identify molecules that are safe and effective, toxicity testing plays a major role during the development stage. The **therapeutic index** (TI), the ratio between the highest tolerated dose (or exposure) and the minimally efficacious dose (or exposure), should be as high as possible. However, the specific disease will determine an acceptable TI. For non-life-threatening diseases, high TI values are required. For life-threatening disorders, such as cancer, TI values of only 2 are sometimes acceptable.

5.5.1 Indications of Toxicity Are Identified Early in the Drug Discovery Stage by *In Vitro* Assays

During lead optimization, there are a number of *in vitro* assays that are employed to give an early indication of potential toxicity. These assays are used so that potential toxicity issues are identified early, and the most promising leads are progressed to animal efficacy testing. Some common strategies and assays include the following:

- *Predictive software.* Multiple software packages and new machine learning algorithms are available to predict toxicity of small molecules. This effort includes

"structural alerts" and the identification of "toxicophores." *Toxicophores* are molecular fragments that are known to have high reactivity or that are prone to biotransformation that generates a reactive metabolite and therefore are associated with toxicity. Examples of fragments that are flagged by structural alerts include nitro groups, 2-halopyridines, anilines, diaminobenzenes, hydroquinones, Michael acceptors, and polycyclic aromatic hydrocarbons, among others.

- *Cytotoxicity assays.* These assays involve treating cells with a small molecule for some period of time and then measuring specific biological readouts of cell viability to determine if the compound causes cell death. No mechanism can be inferred, but this assay provides an indication of potential toxicity.
- *Cytochrome P450 inhibition or induction assays.* One important activity evaluated predicts the potential for **drug–drug interactions**. When one drug prevents the metabolism of another drug by inhibiting a cytochrome P450 (CYP) enzyme, the second drug can reach dangerously high, toxic drug levels because it is not eliminated. For example, ingredients in grapefruit are known to inhibit certain CYP isozymes, and therefore, patients are often told to avoid consuming the fruit while taking certain medications. Drug–drug interactions can also be exploited. The HIV protease inhibitor ritonavir is part of a multidrug regimen prescribed to AIDS patients. In addition to its desired pharmacological effect, ritonavir inhibits CYPs and therefore causes a desirable increase in blood levels of the other protease inhibitors in the drug cocktail, which typically have low bioavailability. Paxlovid, a drug that treats COVID-19, is a combination of the SARS-CoV2 protease inhibitor nirmatrelvir and ritonavir. The urgent nature of the COVID-19 pandemic led to the decision to co-administer a CYP inhibitor (ritonavir) with the anti-COVID compound as a means to increase the exposure of the antiviral agent. Alternatively, a molecule may induce a CYP enzyme, causing an overall increase in the amount of a particular CYP isoform, and therefore its metabolizing ability. This induction can result in the rapid metabolism of a second drug, making that second drug ineffective. Another important reason for evaluating CYP inhibition or induction is the large genetic variability of CYP enzymes among the human population. Certain genetic determinants may influence how a particular drug is metabolized; therefore, knowing that a subpopulation may exhibit drastically different pharmacokinetics than the general population allows safer patient selection. Assays to assess drug–drug interactions are done early in the drug discovery process by determining the effects a molecule has on the ability of a specific CYP isozyme (for example, CYP 3A4, CYP 2D6, or CYP 2C9) to metabolize a specific substrate. When CYP inhibition is observed, it is important to also characterize if the inhibition is reversible or irreversible, as in some instances (e.g., when there is a large differential between the efficacious dose and the CYP inhibition dose), reversible inhibition can be acceptable but irreversible inhibition is not.
- *hERG channel.* The human ether-à-go-go-related gene (hERG) encodes an ion channel that transports potassium ions in heart muscle and is critical for coordinating the heart's beating. When its activity is compromised, a fatal condition called QT prolongation syndrome can result. Recently, several well-known drugs (for example, Seldane) previously considered safe were found to cause sudden death in some patients because the drug inhibited the hERG channel. Since then, testing the activity of molecules in *in vitro* assays for hERG activity is a necessary part of the drug discovery and development process.
- *Genetic toxicity.* Genetic toxicity studies establish the mutagenic potential (whether or not it causes DNA damage) of a drug, and therefore its potential to cause cancer. The **Ames test** is the gold standard for *in vitro* genetic toxicity testing. In this assay, a mutant form of the bacterium *Salmonella* is treated with the test agent. This specific mutant requires histidine in the growth medium in order to replicate. The mixture is then placed in a low-histidine medium. If there are no changes to the DNA, the bacteria will not grow. If the DNA has mutated, the bacteria replicates and suggests that the test agent has the potential to be mutagenic.

5.5.2 Toxicity Testing Is Done with *In Vivo* Assays in Multiple Animal Species

Once a lead has been optimized and *in vivo* PK studies and other studies to evaluate its effects in animal models of disease have been carried out, there are several different types of *in vivo* toxicity testing that are done. These are used to identify a therapeutic index (ratio between maximum tolerated dose and minimal efficacious dose), find the safest dosing levels, and uncover any unexpected effects of the compound. These studies will also show effects of long-term dosing that may not be evident otherwise.

Acute toxicity studies are done in animals at increasing doses until toxicity is observed. Studies generally begin in rodents and progress to dogs or rabbits. Acute testing begins with a single dose, and the animal is observed for signs of toxicity, such as vomiting, diarrhea, weight loss, lethargy, or death over the course of 7–14 days after the dose. Blood levels of the drug are also monitored to determine at what drug concentration toxicity is observed. The relationship between toxicity and concentration of drug in circulation is termed **toxicokinetics** (TK).

Range-finding studies are then carried out in which several different doses (usually a high, medium, and low dose) are administered daily over 2–4 weeks. These studies establish the range of doses that can be administered without significant toxicity, as well as the maximum tolerated dose (MTD), and they dictate the doses that will be used in subsequent subchronic and chronic toxicity testing.

Subchronic testing and **chronic testing** comprise the next level of *in vivo* toxicity testing and involves much longer duration of dosing. During subchronic testing, the animals (either rodents or non-rodents) are dosed daily for three months and monitored for signs of toxicity, including weight loss or gain. After completion of the study, tissues and organs are harvested and examined for histological changes compared to control animals. Chronic testing lasts for longer than three months and is usually of at least the same duration as the proposed dosing in clinical trials. These are carried out in at least one rodent and at least one non-rodent species. Animals are monitored after dosing stops to identify effects that appear after long-term dosing. Chronic testing has two components: first is 6–12 months of observation, looking for signs of toxicity, and second is 12–24 months of observation for signs of carcinogenicity.

These acute and chronic toxicity studies are used to establish the no-observed-adverse-effect level (NOAEL) of the drug. This value is the highest dose of the drug that can be administered without serious toxicity and helps define the tolerable dose for clinical trials.

Reproductive toxicity determines whether the drug will have any effects on fertility, the ability to maintain a pregnancy, and the health of the progeny of patients. It is completed in three parts:

1. **Segment I testing** looks at fertility and reproductive performance in both male and female animals that are administered the drug.
2. **Segment II testing** evaluates embryonic toxicity (teratogenicity) during early pregnancy of mothers being dosed with the drug.
3. **Segment III testing** determines if there are any effects of the drug on the final stage of pregnancy, as well as on delivery and lactation (milk production), when administered to a pregnant female or mother.

Most drug candidates that have unacceptable levels of toxicity, whether acute, chronic, or reproductive, will not proceed to clinical trials. In these cases, lead optimization is re-initiated in order to find an acceptable candidate. Depending on the therapeutic area, however, different levels of toxicity are acceptable. For example, several of the coxib anti-inflammatory drugs were withdrawn from the market after exhibiting cardiac liabilities. However, doxorubicin, an anti-cancer drug, is still in use despite its cardiotoxicity because its benefits outweigh the toxicity. In another example, thalidomide, the teratogenic sedative, has found new use to treat both leprosy and multiple myeloma and is now marketed with extensive warnings of its teratogenicity.

A summary of some important parameters that are commonly measured or calculated during ADMET studies is given in **Table 5.4**.

Table 5.4 Summary of Pharmacokinetic and Safety Parameters and Measurements

Absorption Parameters, Measurements, and Relevance	
C_{max}, maximum concentration; t_{max}, time to reach maximum concentration	• Peak plasma concentration of a drug after administration and the time it is reached. • Measured in *in vivo* PK assays in concentration and time units; C_{max} values should be sufficient for activity, depending on drug potency (C_{max} should exceed *in vitro* IC_{50}). • Important to understand pharmacodynamics and toxicity effects; used to determine most appropriate dose and dosing schedule.
C_{ss}, steady state concentration	• Equilibrium concentration of drug after a number of doses (when the amount of drug absorbed/time = amount eliminated/time). • Measured in *in vivo* PK assays in concentration units; based on AUC taken over different times; dependent on dosing. • Used to determine most appropriate dose and dosing schedule.
AUC, area under the curve	• Integral of plasma concentration over a set time period. • Measured in *in vivo* PK assays; values should be sufficient for activity, depending on drug potency and pharmacodynamic requirements. • Related to the required dose and frequency of dosing.
F or *FA*, oral bioavailability	• Fraction of drug that reaches systemic circulation upon oral administration compared to an IV dose; equal to $AUC_{oral}/AUC_{iv} \times 100$. • Measured in *in vivo* PK assays; greater than 20% is considered developable, albeit low, with >50% usually considered acceptable. • Related to dose required and efficiency of absorption.
Metabolism, Distribution, and Elimination Parameters, Measurements, and Relevance	
Fraction of drug unbound (f_u); plasma protein binding	• Degree of binding to plasma proteins. • Measured in secondary *in vitro* assays; fraction unbound (f_u) >99% usually undesirable, as is irreversible binding. • Reflective of how much drug is available to bind to the target.
V_d, distribution volume	• Volume (in liters or liters per kilogram) in which drug is distributed, if homogeneous distribution is assumed. • Measured in *in vivo* PK studies; drug may be in plasma and/or tissue. • When $V_d < 6$ L, most of the drug is in the blood or plasma; when $V_d > 42$ L, all the drug is in the tissue. • Indicates how much drug is distributed outside of plasma; specific disease target dictates desirable values.
CL, clearance	• Rate at which a drug is eliminated, calculated as dose/AUC or volume of plasma cleared of drug per unit time. • Measured in *in vivo* PK assays; can also be predicted from *in vitro* ADME assay data. • One parameter that contributes to effective *in vivo* drug concentration; used to determine appropriate dose and dosing schedule.
$t_{1/2}$, half-life of drug	• Time required to halve the plasma drug concentration. • Measured in *in vivo* PK assays. • Related to required dose and frequency of dosing.
Metabolic stability ($t_{1/2}$)	• Ability of compound to withstand metabolic transformation by liver enzymes. • Measured in primary or secondary assays, such as liver microsomes or hepatocytes, to detect half-life of compound. • $t_{1/2}$ > 60 min is ideal, although very long $t_{1/2}$ may be unsafe; values used to calculate CL_{int}. • Related to achieving high compound concentrations (bioavailability) required for *in vivo* biological activity.
Safety and Toxicity Parameters	
Cytochrome P450 (CYP) inhibition or induction	• Ability of compound to inhibit or induce metabolism of specific compounds by specific CYP isoenzymes. • Measured in primary assays; no/low inhibition preferred. • Acceptable values are related to the differential between CYP affinity (K_i), C_{max}, and IC_{50} (or EC_{50}) and whether inhibition is reversible or irreversible. • Indication of potential to interfere with metabolism of other administered drugs and is a safety consideration.
hERG, human ether-à-go-go gene, channel activity	• Potential of compounds to inhibit the hERG ion channel. • Measured in primary and secondary assays as a concentration; inhibition of >10 μM usually acceptable, with 30-fold safety margin desirable. • Reflection of safety of the compound.
MTD, maximum tolerated dose	• Highest dose that is tolerated without significant toxicity. • Measured in *in vivo* toxicity assays; higher doses are better. • Indication of safety of the drug.
LD_{50}, lethal dose	• Dose that is lethal to 50% of animals tested. • Measured in *in vivo* toxicity assays; higher values are safer. • Indication of safety of the drug.
TI, therapeutic index	• Ratio between maximum tolerated dose (or exposure) and minimal efficacious dose (or exposure) = MTD/MED. • Measured in animal efficacy and toxicity studies (*in vivo*). • Higher TI values are preferable, but the specific condition dictates a tolerable TI: For example, drugs to treat cancer often exhibit a TI between 2 and 10, while drugs that treat less serious disorders, childhood diseases, and chronic diseases require a much higher TI (>10). • Indication of safety of the drug.

5.5.3 Formulation of the Final Drug Product

The active drug substance is termed the **active pharmaceutical ingredient** (API); however, the medication itself contains additional components, and together these make up the pill, capsule, liquid medication, or other drug substance. The combination of the API and the other components, known as excipients, is the **formulation** of the drug. The specific formulation can affect the site and degree of absorption, half-life, bioavailability, and dosing schedule, among other factors. In addition, the formulation must be stable over long periods of time, be inert to heat and humidity, and consistently deliver the same drug levels. Specific disease states and patient populations are additional factors considered during formulation development. In infant and pediatric patients, for example, liquid formulations for oral administration are preferred over capsules or pills.

Many of the factors that influence the specific formulation developed are related to the physical properties of the drug molecule: solid-state form (for example, the specific crystalline polymorph), particle size of solids, pK_a, solubility, and dissolution rates are key determinants. The bioavailability of compounds of low solubility and high permeability can be improved with formulations; however, it is difficult to improve PK characteristics through formulation of compounds with low permeability, regardless of their solubility.

While complex formulation science is beyond the scope of this text, the generation of salts of drugs that are acids or bases is an important component of formulation studies. The free acid or base of the API is often evaluated in the earliest stages of the drug discovery process, but evaluation of various salt forms can occur during lead optimization to identify products that have improved solubility, crystallinity, stability, and permeability. These characteristics will all affect exposure and bioavailability of the compound.

5.6 CLINICAL TRIALS IN HUMANS

After the 4–8 years it takes to complete the discovery and preclinical development studies, a clinical candidate is identified from the optimized leads. This compound will exhibit the best combination of potency, physical and pharmaceutical properties, and low toxicity. The preclinical data, including method of synthesis, formulation plan, and results of pharmacokinetic and toxicity studies in multiple animal species, as well as a plan for clinical trials, is submitted to the US FDA as an **Investigational New Drug Application** (IND). This proposal needs to be approved by the FDA and by an Institutional Review Board (IRB), where the human studies will take place before proceeding with trials in humans. Testing of drugs in humans usually occurs in four phases, but only the first three are required for approval:

1. During **Phase I clinical trials**, the drug is tested in a very small number of healthy volunteers (20–80) for a short duration, often in a hospital, so subjects can be closely monitored, with the goal of uncovering unexpected toxic effects and gathering data on pharmacokinetics in humans. According to a review spanning 2010–2023, approximately 48% of the drugs that enter phase I trials move onto phase II studies.
2. If no unacceptable toxicity is observed, the efficacy of the drug in patients is then evaluated in **Phase II clinical trials**. These studies involve a larger number of carefully selected subjects (~100–300), and they compare the efficacy of the drug candidate with either a placebo or a drug that is already proven to be effective. These studies can be **double-blind**—meaning that neither the patients nor the clinicians administering the drug are aware of which agent (drug candidate or comparator) is being given to each patient. At the same time, safety is continually monitored. These studies can last for a few months to several years, and only ~35% of phase II–tested drugs move into the next phase.
3. If efficacy is observed in Phase II, the program progresses to **Phase III trials**. These studies are more complex and involve an even larger number of patients

(~200 to >3,000), often in different stages of the disease, to study efficacy, different dosages, different patient populations, combinations with other drugs, tolerability, and safety. Only 50% of drugs successfully complete phase III studies.

4. Additional studies that further evaluate efficacy, safety, and optimal use after drug approval are termed **Phase IV trials**.

Once the Phases I–III studies are completed and the drug candidate shows statistically significant effects and acceptable side effects/toxicity, an extensive report is filed with the regulatory agency. This is called a **New Drug Application** (NDA) in the United States. This report contains not only data from the clinical trials on efficacy, safety, and PK but also all previously acquired data in animals, as well as specifics on how the drug will be manufactured after approval. Prior to 1992, the average time for the FDA to grant approval was 2–3 years, although this could be considerably shorter for drugs to treat a life-threatening disease. In 1992, the Prescription Drug User Fee Act (PDUFA) was passed, in which the applicant pays a fee to the FDA. It is estimated that the FDA receives approximately $100 million per year from applicants, part of which has been used to increase FDA staffing. Since then, the time for approval has shortened to about 6 months to 2 years. Once approved, the drug can be marketed for a specific use, based on the regulatory agency's review and requirements.

After the drug is approved, additional studies are usually required that further evaluate efficacy, safety, and optimal use; these are termed Phase IV trials, which include **post-marketing surveillance**. The regulatory agency continues to monitor reports on side effects and any results from post-marketing studies, as well as results on new uses for the drug. Should new data arise, the regulatory agency can change the labeling of the drug, add new uses, add specific warnings, change the suggested dosing, or remove the drug from the market.

Once a molecule enters clinical trials, the probability that it will become an approved drug is approximated at only 11%. The entire process, from the earliest stages to drug approval, takes an average of 12–14 years, and costs are estimated to be well over $2 billion, according to FDA figures (https://www.fda.gov and https://clinicaltrials.gov). Out of the 5,000–10,000 molecules that might be evaluated in the earliest stages of the drug discovery project, chances are that only one will be approved and marketed.

5.7 SUMMARY

As described in the first five chapters of this text, the overall process of developing a drug is a complex and lengthy process. Hit and lead identification, followed by lead optimization, and then extensive testing are required to identify a viable clinical candidate. The optimization process requires a cycle of design, synthesis, testing, and data analysis to achieve the optimal balance of potency, properties, and low toxicity. This is all a part of the preclinical discovery process, which also includes the important process of patenting viable leads. As an optimized lead emerges from preclinical *in vivo* assays, attention turns toward complete understanding of that compound. Evaluation of the pharmacokinetics of the compound in multiple *in vitro* assays and in several animal species determines whether the compound is suitable for further advancement and helps predict human dosages. The development of a process synthesis to produce kilogram quantities of a drug is an important activity during this phase, as is formulation of the drug substance to the eventual final form of the medicine. Extensive testing in *in vitro* assays as well as in animals to uncover any potential for toxicity is also done in preparation for human trials. For this final stage of development, a drug needs to pass three phases. Phase I clinical trials establish safety and dosing levels, as well as human pharmacokinetics. Phase II trials test efficacy, and then Phase III trials, in larger populations, verify efficacy and particular aspects of safety. Even after many years are invested into a drug discovery program, there is still only about a 10% chance that a drug candidate will progress through all phases of clinical trials and receive approval.

CASE STUDY 1: DISCOVERY AND DEVELOPMENT OF UBROGEPANT (UBRELVY)

Migraine, a common neurological condition that affects about 11% of the population, is characterized by recurrent debilitating headaches, many times with photophobia and phonophobia (aversion to light and sound) and nausea. It is associated with exaggerated pain signaling as well as dilation of blood vessels in the brain. The widely used anti-migraine drugs called triptans, such as sumatriptan and rizatriptan, are vasoconstrictors that work through selective activation (agonism) of the serotonin 5-$HT_{1B/1D}$ receptor, thereby reversing the blood vessel dilation. However, the triptans are limited in that they cannot be used in patients with cardiovascular conditions. An alternate approach began in the 1990s with the observation that CGRP (calcitonin gene–related peptide) release is increased in patients undergoing a migraine attack. CGRP is a 37-amino acid neuropeptide that, after binding to its receptor, causes vasodilation and is involved in pain signaling. The hypothesis was that preventing increased levels of CGRP; therefore, preventing the increased vasodilation produced during migraine attacks would alleviate the symptoms. The receptor is a GPCR and a member of the secretin family of receptors. Ubrogepant (MK 1602) is a migraine medication that was approved in 2019 as a first-in-class, orally active calcitonin gene–related peptide (CGRP) receptor competitive antagonist to treat the acute phase of migraine, and the closely related atogepant was approved two years later for the prevention of migraine.

Early Clinical Candidates Olcegepant and Telcegepant

Scientists at Boehringer-Ingelheim and Merck began a search for oral CGRP antagonists (blockers) in the early 2000s. CGRP had been discovered in 1982, and its C8-C37 C-terminal fragment had been reported as an antagonist, although with low potency and no oral activity. At Boehringer-Ingelheim, high-throughput screening measuring displacement of radiolabeled human CGRP in a human neuroblastoma cell line led to a series of dipeptide-like compounds, which were then optimized. One of the initial hits, the modified dipeptide based on Tyr-Lys (red), had an IC_{50} of 17 μM (**Figure 5.12**). Lead modification of this hit compound included chain-ring modifications

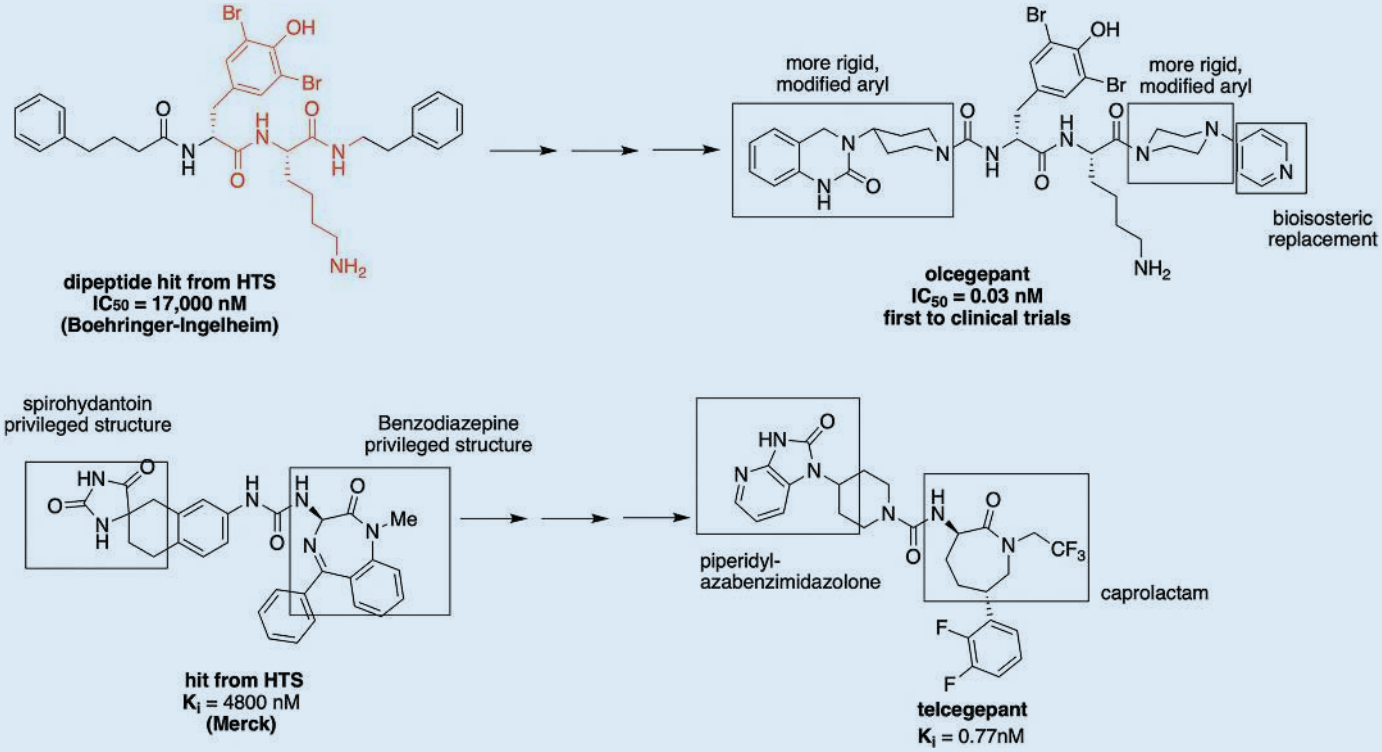


Figure 5.12 HTS hits and CGRP receptor antagonist clinical candidates to treat migraine. *Top*, a dipeptide identified via HTS (modified tyrosine and lysine highlighted in red) was optimized using conformational constraint and bioisosteric replacement strategies to generate olcegepant. *Bottom*, an HTS hit containing privileged scaffolds was optimized to generate telcagepant.

to make the structure more rigid, as well as using bioisosteric replacement of one of the terminal aryl rings with a pyridine group. This optimization resulted in olcegepant, the first CGRP antagonist to proceed to clinical trials. One important potential advantage of olcegepant over the widely used triptans was a lack of cardiovascular events. However, olcegepant had a high MW (869), was highly polar, had low oral availability, and had to be dosed intravenously. Concomitantly, Merck had identified several hits from HTS, one of which contained spirohydantoin- and benzodiazepine privileged scaffolds. Changing the spirohydantoin to a piperidyl-azabenzimidazolone as well as simplifying the benzodiazepinone to a caprolactam improved potency, ultimately leading to orally active telcagepant. This compound progressed to clinical trials, but those were suspended in 2011 due to patients' exhibiting elevated liver enzymes.

Development of Ubrogepant and Atogepant

Pursuing another avenue, scientists at Merck modified their original hit, this time maintaining a spirocycle. The spirohydantoin privileged structure was modified to a spiroazaindane, which molecular modeling showed mimicked the C-terminal Phe-NH_2 of CGRP. In addition, the benzodiazepinone was replaced with a ring-contracted 2-piperidinone, resulting in a spiroazindane with improved binding (K_i = 0.04 nM). Further optimization led to MK-3207 containing an additional NH, which led to improved solubility, increased polarity, and avoidance of first-pass metabolism in monkeys (**Figure 5.13**). Conformational constraint was also introduced via connecting the gem-dimethyl group into a ring. MK-3207 had improved binding (0.021 nM) and was selective *vs.* other calcitonin family receptors. Pharmacokinetic studies showed about 70% oral availability in rats and dogs and 41% in monkeys (at 20 mg/kg), with a rapid T_{max} of 1–2 hours in humans. The mean plasma concentration was 40 nM at 2 hours, higher than the projected 14 nM EC_{90} level. Capsaicin-induced dermal vasodilation (CIDV) assay is a non-invasive test that involves the application of capsaicin to the forearm of a rhesus monkey and measurement of blood flow using laser Doppler imaging. The difference between the blood flow in untreated vs. treated animals gives a measurement of pain reduction. MK-3207 reduced blood flow by 89%. MK-3207 progressed to phase II clinical trials, was effective at pain reduction, and was well-tolerated; however, elevated liver enzymes were observed in patients, and the compound's development was stopped.

After licensing arrangements and further optimization, including reversing the amide in MK-3207, ubrogepant and atogepant were developed and approved for acute and preventative treatment, respectively. These compounds differ only in the

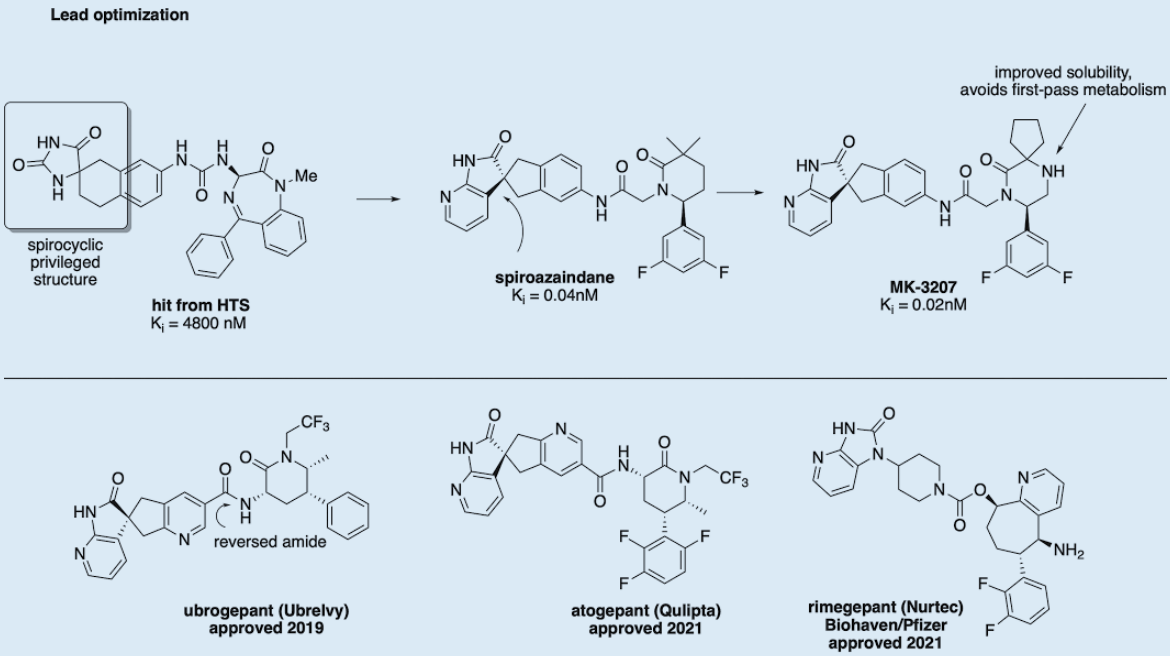


Figure 5.13 Lead optimization of HTS hit to clinical candidate MK-3207, and structures of the approved drugs derived from this series, ubrogepant and atogepant. Also shown is rimegepant, another CGRP antagonist, structurally related to telcagepant.

substitution on the pendant aromatic ring. Rimegepant, related to the earlier candidate telcagepant, was also approved for treatment of migraine.

Synthesis of Ubrogepant

Key challenges in the synthesis of ubrogepant (and atogepant) include setting the four stereogenic centers. Analysis of the medicinal chemistry route to ubrogepant identified areas for improvement for manufacturing, such as poor enantioselectivity requiring the need for expensive chiral chromatography, complications from polymerization during simultaneous formation of both bonds of the spirocycle, and use of an expensive SEM nitrogen protecting group.

The process synthesis, described in what follows, is convergent and combines two elaborated segments, **1** and **2**, via amide bond formation as the final step. Synthesis of the acid **1** began with 2,3-dibromo-5-chloropyridine and proceeded through trans-metalation and diacylation first at position 3 and then at position 2. Reduction of the 3-formyl group and protection of the resultant alcohol as a THP ether afforded **3**. Condensation of **3** with the protected azaindolone afforded an alkene (not shown). Reduction, and removal of the THP protecting group was followed by chlorination to afford compound **4**. This key intermediate was cyclized using a chiral phase transfer catalyst to set the stereogenic center, which provided the spirocycle **5** in high yield and with 99.5% ee. Carbonylation and removal of the N-tert-butyl protecting group gave the acid **1** in 11 steps and 43% yield from 2,3-dibromo-5-chloropyridine. The second key intermediate, the amine **2**, began with conjugate addition of 4-bromophenylacetone to an α,β-unsaturated ester, followed by reduction of the bromine atom to afford the cyclization substrate **6**. Note that the 4-bromo substituent is unnecessary and is removed by reduction. The unsubstituted phenylacetone starting material would have been a preferred starting material but is a Schedule II controlled substance and hence was not used. The key cyclization

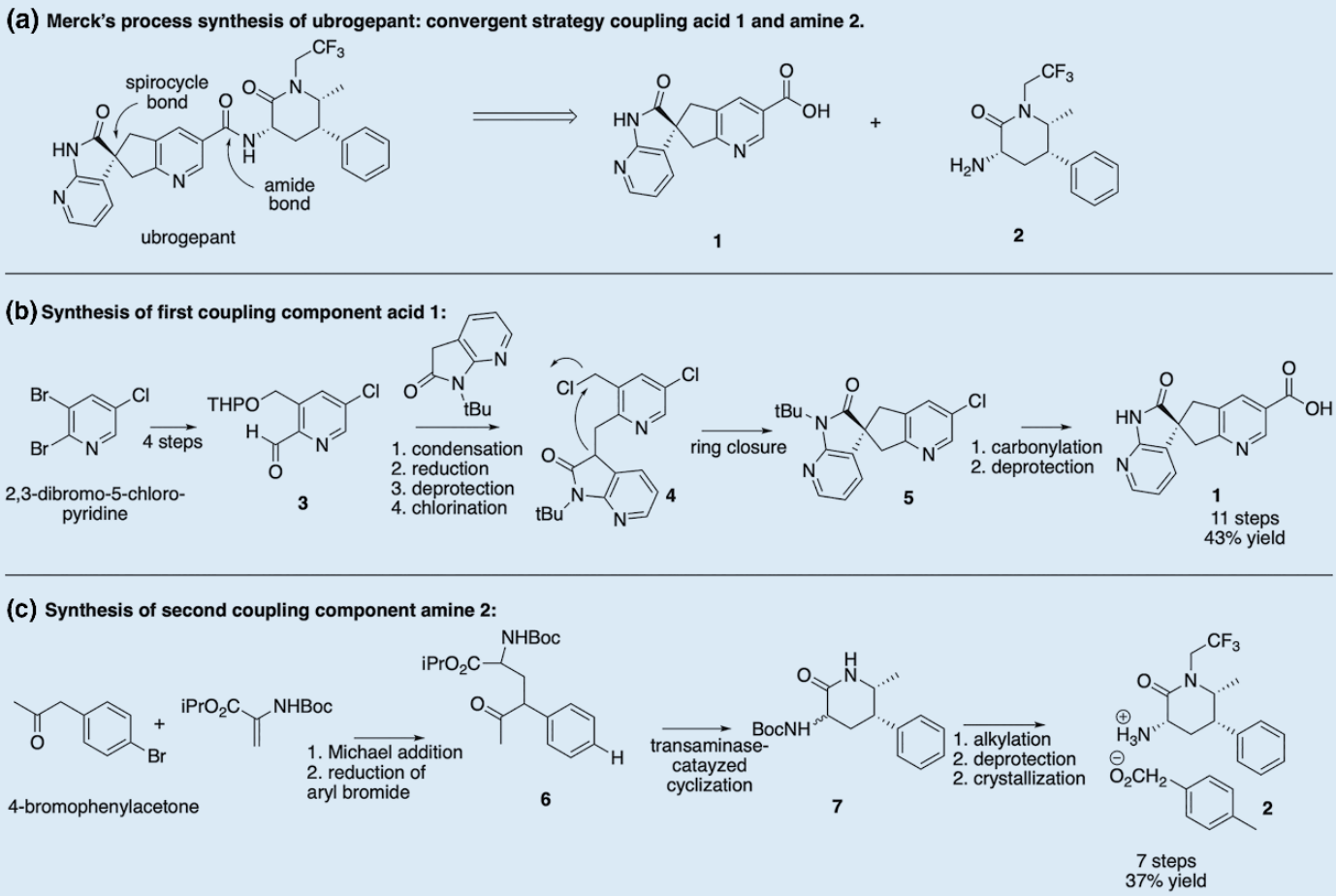


Figure 5.14 (a) Overall convergent strategy coupling acid 1 with amine 2 to form the amide bond of ubrogepant. (b) Synthesis of the acid 1: the aryl dibromide is diacylated, mono-reduced, and protected to give 3. Condensation of 3 with the bicyclic amide, followed by alkene reduction, THP deprotection, and chlorination, provides 4. Stereoselective ring closure gives 5, and carbonylation and N-deprotection afford the acid 1. (c) Michael addition, followed by reduction of the aryl bromide, generates compound 6, and then stereoselective cyclization using an evolved transaminase enzyme gives the lactam 7. Alkylation and deprotection, followed by crystallization in the presence of an aldehyde, afford the salt of amine 2.

step used a dynamic kinetic transamination with an engineered transaminase enzyme as catalyst to set the two contiguous stereogenic centers with 61:1 stereoselectivity to afford the lactam **7** as a mixture of diastereomers at the position alpha to the ketone. Alkylation of the nitrogen was followed by Boc deprotection and crystallization, using p-toluic acid and a catalytic amount of 3,5-dichlorosalicylaldehyde. The aldehyde mediated the crystallization-induced diastereoselective transformation (CIDT), which afforded only one crystalline diastereomer, the amine **2**, as the toluic acid salt.

Ubrogepant (Ubrelvy) received FDA approval in 2019, and atogepant (Qulipta) in 2021. These drugs are structurally similar to MK-3207, with the key change being the incorporation of a reversed amide. This change addresses the structural alert that flags the aniline functional group while maintaining the spiroazaindane of MK-3207. Both also have vestiges of telcegepant and are ring-contracted versions of that compound and maintain the trifluoroethyl N-alkyl group. Also approved in 2021 was rimegepant (Nurtec), developed at Bristol Myers Squibb, licensed to Biohaven, and marketed by Pfizer, which has structural similarities to telcegepant. Since their introduction to the market, these drugs have rapidly become preferred treatments for migraine.

CASE STUDY 2: DEVELOPMENT OF RIVAROXABAN (XARELTO)

Rivaroxaban (Xarelto) is an anticoagulant drug that was developed by scientists at Bayer AG in Germany. Older drugs to treat blood clots include heparin, which must be given intravenously, and warfarin, which requires regular blood monitoring due to its low therapeutic index. Rivaroxaban resulted from the search for an alternative to these drugs that was safer and orally active. One proposed novel target was activated serine protease factor Xa (FXa), which plays a role in the blood coagulation cascade. This enzyme converts prothrombin to thrombin. Thrombin then converts fibrinogen to fibrin, activates platelets, and through feedback activation of other coagulation

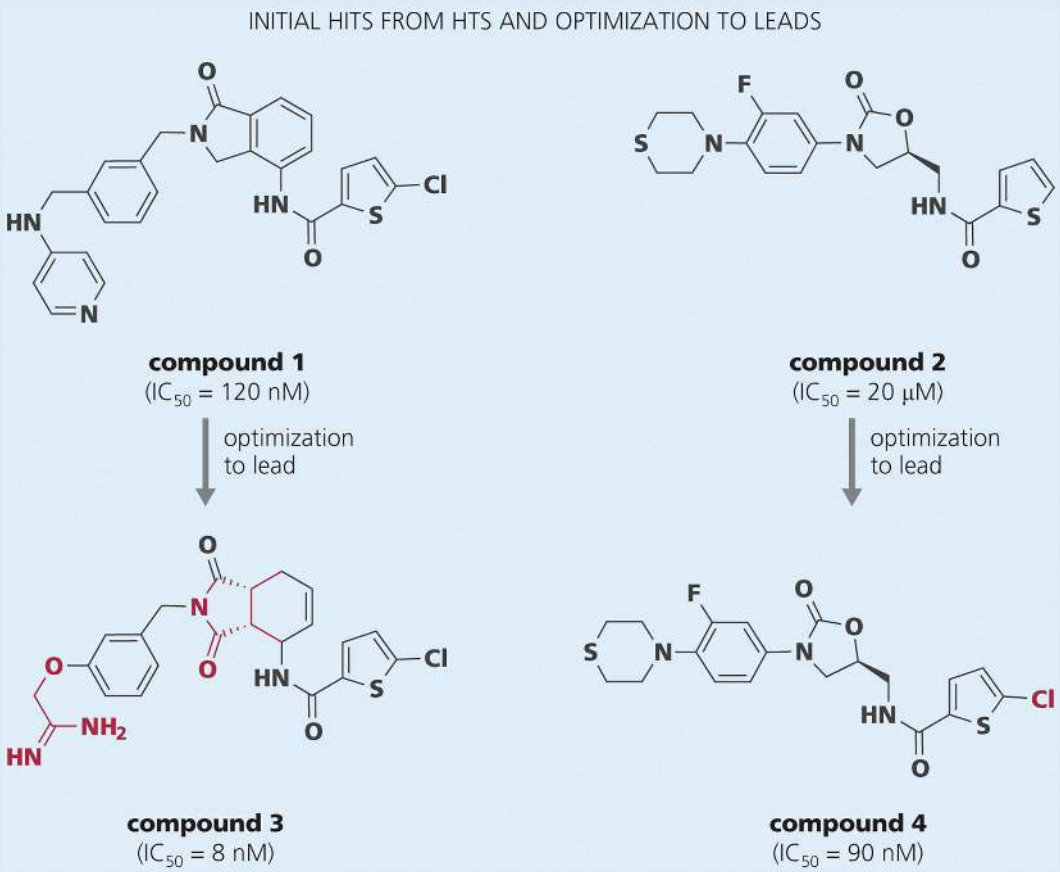


Figure 5.15 Structures of the hits and leads that led to the discovery of rivaroxaban. Compounds **1** and **2** were initial hits and optimized to lead compounds **3** and **4**. Changes are shown in red.

factors, increases its own levels. By inhibiting FXa, coagulation would be inhibited, and blood clots would be prevented.

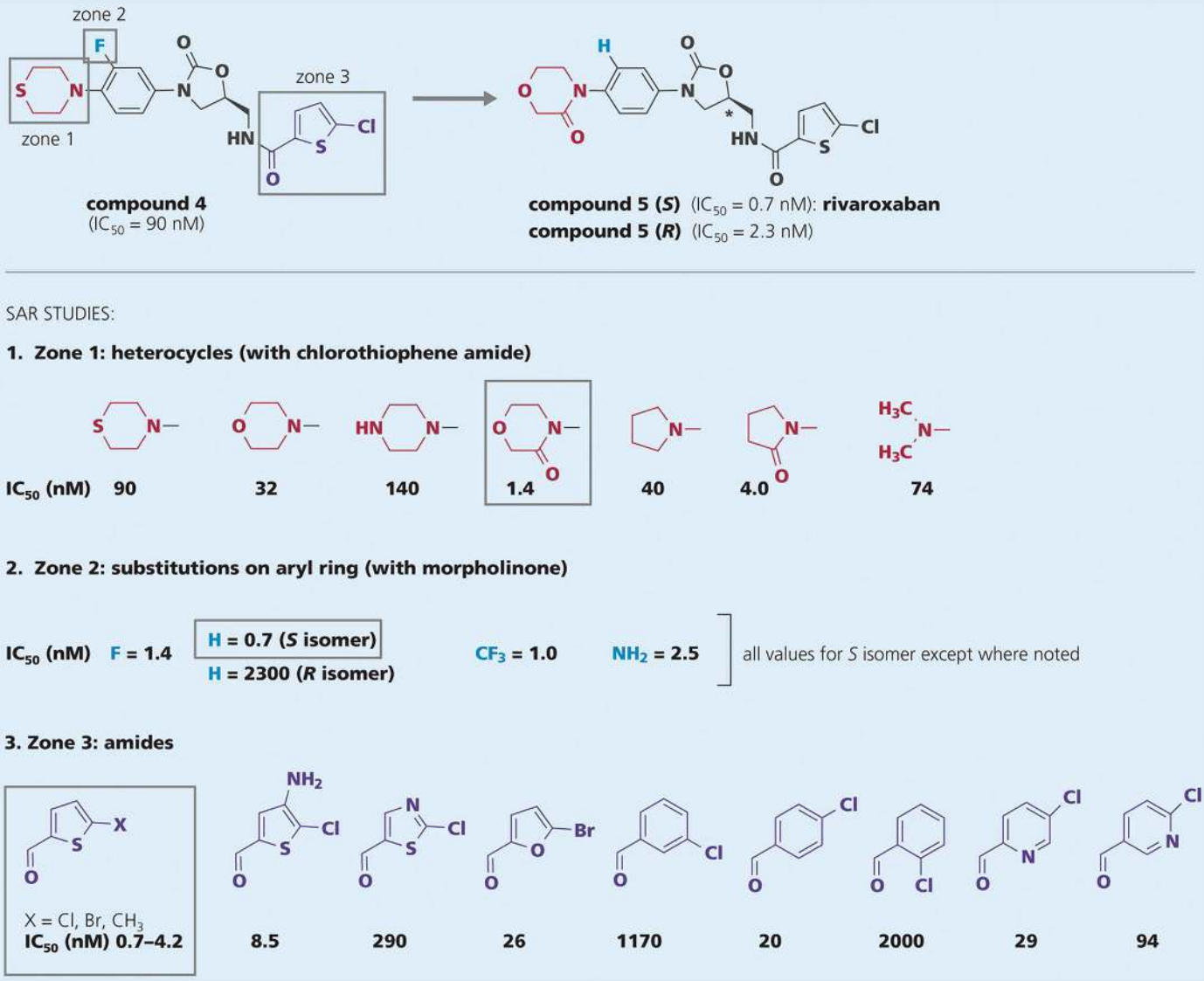
As a starting point to identify inhibitors of FXa, the initial hits **1** and **2** were discovered by high-throughput screening (HTS) of over 200,000 compounds (Figure 5.15).

The assay used at the discovery stage detected cleavage products of fluorogenic substrates of FXa. Optimization of hit compound **1** led to **3**, which had improved potency, but the series exhibited poor pharmacokinetics due to low absorption.

The SAR developed during the investigation of hit **1** suggested that the 5-chlorothiophene-2-carboxamide was necessary for activity. On the basis of that information, the lower-potency hit **2** was re-examined. Optimization to lead compound **4** simply involved the addition of the chlorine atom to the thiophene moiety, which afforded a nearly 200-fold improvement in potency.

Further optimization of **4** first involved the replacement of the thiomorpholine group (zone 1). The investigation of a number of different heterocycles led to morpholinone as the preferred group. The *S*-configuration at the oxazolidinone (indicated by *) was required for optimal activity, and in the *S*-isomers, replacement of the fluorine atom of the aryl ring with hydrogen (zone 2) doubled activity. These changes resulted in **5** (rivaroxaban). The lipophilic 5-chlorothiophene of **5** is responsible for low aqueous solubility (8 mg/L); however, investigation of other heterocycles and substituted thiophenes (zone 3) gave a sharp decrease in potency. For example, the introduction of an amino group at the 3-position of the thiophene resulted in a 10-fold loss of potency (IC_{50} = 8.5 nM). The SAR studies are summarized in Figure 5.16.

Figure 5.16 Summary of SAR studies from optimization of lead compound 4 to rivaroxaban. Zone 1 was optimized to the oxomorpholine. Zone 2 optimization led to the unsubstituted aryl ring (H) with the *S*-configuration at the stereogenic center. Zone 3 was found to be optimal with chlorothiophene.



The solution of the X-ray co-structure of rivaroxaban bound to FXa shows several key binding interactions, as seen in Figure 5.17, where rivaroxaban is shown in orange and the enzyme is shown in green. The molecule adopts an L shape that is optimal for binding, with the chlorothiophene amide extending back into the binding site in the S1 pocket. The oxazolidinone ring is coplanar with the aryl ring; two hydrogen bonds to Gly219 are observed—one from the oxazolidinone carbonyl and another from the amide nitrogen. The morpholinone ring, which is orthogonal to the aryl ring, fits into the S4 pocket.

Compound **5** (rivaroxaban) is synthesized by the efficient route shown in Figure 5.18, where the substituted aniline adds to the chiral epoxide to give the amino alcohol oxazolidinone precursor. Cyclization with carbonyldiimidazole (CDI) is followed by cleavage of the phthalimide group to afford a primary amine, which is then coupled to the chlorothiophene.

In addition to the inhibition assay, further characterization of **5** included secondary *in vitro* assays that measured prolongation of prothrombin time in both rat and human plasma. This assay measures the inhibitor concentration required to double the time to fibrin formation, compared to no inhibitor being present. Compounds, all containing the oxomorpholine ring but with varying aryl substituents (X, Figure 5.19), were tested, with **5** and the F (**6**), CF_3, and NH_2 analogs giving comparable results. Testing *in vivo* in a rat model to examine anti-thrombotic effects showed that **5** and **6** had equivalent intravenous potency (ED_{50} = 1 mg/kg); however, rivaroxaban (**5**) was twice as potent as **6** (ED_{50} = 5 vs. 10 mg/kg) when dosed orally.

Further testing of **5** and **6**, as well as the pyrrolidone analog **7**, for their pharmacokinetic profile in rats gave the results summarized in Figure 5.19.

Compound **5** was tested for selectivity against related serine proteases thrombin, trypsin, plasmin, FVIIa, FIXa, FXIa, urokinase, and activated protein C and had no effect at concentrations up to 20 μM.

On the basis of its potency, effects in animal models, and pharmacokinetic profile, **5** entered clinical trials. In Phase I trials, it was found to be well-tolerated,

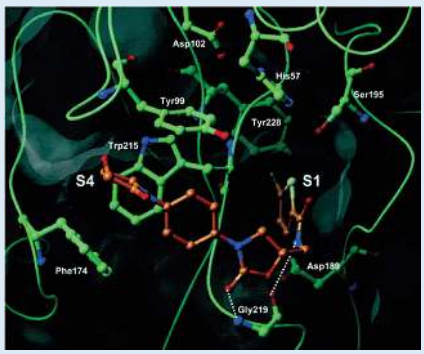


Figure 5.17 X-ray crystal co-structure of rivaroxaban (orange) bound to FXa. Rivaroxaban adopts an L shape in the binding site. Hydrogen bonds to Gly219 from the oxazolidinone carbonyl and amide nitrogen are shown as dotted lines. The S4 pocket is filled by the morpholinone ring; the S1 pocket is filled by thiophene.

Source: From Roehrig, S.; Straub, A.; Pohlmann, J.; et al. (2005). *J Med Chem* 48:5,900–5,908. With permission from the American Chemical Society.

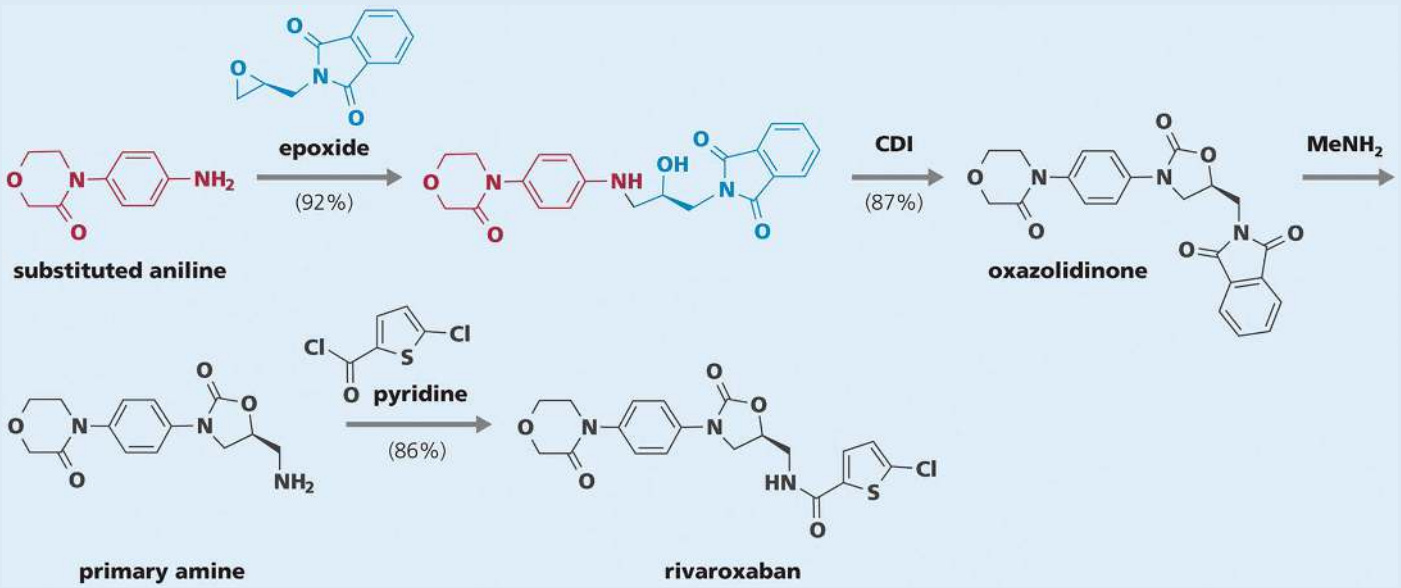
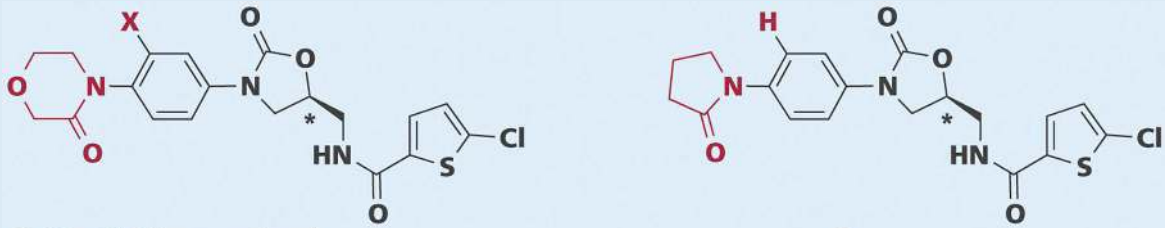


Figure 5.18 Synthesis of rivaroxaban (5). The synthesis begins with substituted aniline (shown in red) and epoxide (shown in blue). Epoxide opening to generate the amino alcohol is followed by the formation of the oxazolidinone. Cleavage of the phthalimide generates a primary amine, which is coupled to an acid chloride to afford rivaroxaban.



5: X = H (rivaroxaban)
6: X = F

7

compound	CL [L·h^{-1}kg^{-1}]	V_{ss} [L·kg^{-1}]	$t_{1/2}$ [h]	bioavailability [h]	f_u [%]
5	0.4	0.3	0.9	60	1.3
6	1.09	0.32	0.3	n.d.	n.d.
7	0.17	0.21	1.1	65	0.2

f_u = fraction unbound in plasma

Figure 5.19 Summary of pharmacokinetic data of compounds 5, 6, and 7 in male Wistar rats. Compound **5** (rivaroxaban) had the best combination of low clearance, appropriate volume of distribution, acceptable half-life and plasma protein binding profile, and high bioavailability compared to **6** and **7**.

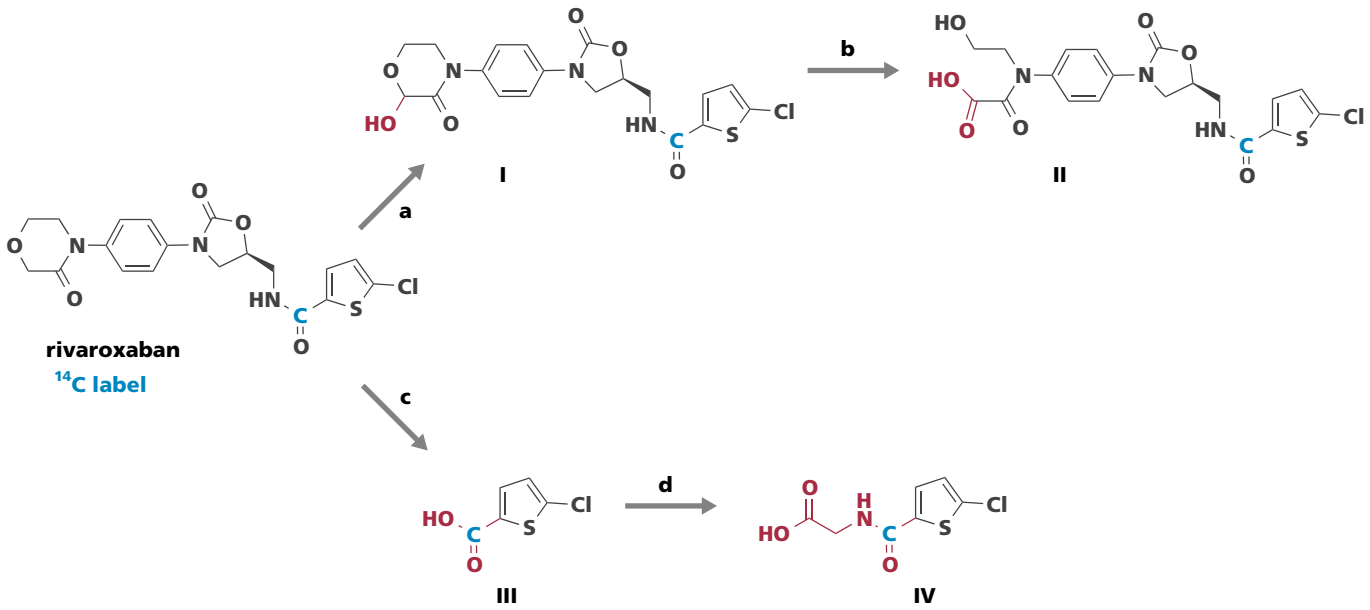
with predictable pharmacokinetics. In Phase II trials, over a range of 5–80 mg, it exhibited anticoagulant effects with no bleeding. It was examined for the prevention of venous thromboembolism following surgery and was effective at a 10 mg dose. Rivaroxaban received FDA approval in 2008 for the prevention of blood clots in surgical patients.

5.8 REVIEW QUESTIONS

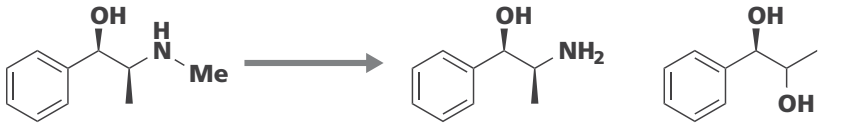
1. Describe several US federal regulations that have affected the drug industry.
2. Explain the difference between Phase I and Phase II metabolism.
3. Describe three different types of *in vitro* toxicity assays that are done during the early drug discovery stages and what they predict.
4. Give at least four characteristics of a process synthesis, and explain why it may differ from the original synthesis of a drug candidate.
5. Describe the three types of testing done for reproductive toxicity.
6. Describe three types of patents that are used to protect drug and drug candidates.
7. What are some of the properties and characteristics of a molecule that can affect its distribution in the body?
8. Explain why a highly bioavailable drug is preferable to a poorly bioavailable drug.
9. Define pharmacokinetics, pharmacodynamics, and toxicokinetics.
10. Explain several reasons a drug might have excellent potency in *in vitro* assays but poor activity in *in vivo* animal models.

5.9 APPLICATION QUESTIONS

1. In spite of toxicity screening in both preclinical and clinical trials, sometimes serious side effects do not become apparent until after a drug has been approved and marketed and has been taken by large numbers of patients. This was the case with several selective COX-2 inhibitors, such as rofecoxib and valdecoxib, that were in use as anti-inflammatory agents. Research these two drugs, and identify the side effects that were discovered, leading to their voluntary removal from the market.
2. Two metabolic pathways for cimetidine involve S-oxidation to the sulfoxide and hydroxylation of the 5-methyl group on the imidazole ring. Draw structures of these two metabolites.
3. Studies on ^{14}C-labeled rivaroxaban (Case Study 2 in this chapter) led to the identification of the four metabolites (I–IV) shown. Classify each metabolic steps (a)–(d) by reaction type as well as whether it represents phase I or phase II metabolism.



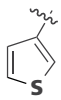
4. The metabolism of brompheniramine is shown in **Figure 5.11b**. Draw the reaction pathway of the first phase I reactions to generate the carboxylic acid metabolite.
5. Two metabolites of ephedrine are shown here (major and minor).



ephedrine **major metabolite** **minor metabolite**

   a. What type of phase I transformations led to each metabolite? The diol (minor) metabolite is excreted as a glucuronide. Draw that phase II metabolite.
   b. Using some of the strategies described in Chapter 4, design a molecule that could slow or modify the formation of each metabolite, and explain the rationale of your design.

6. *Green chemistry* is "the design of chemical products and process that reduce or eliminate the generation of hazardous substances," according to the US Environmental Protection Agency. The concepts of green chemistry are an important component of process chemistry research, and a number of processes for the synthesis of drugs have received the Presidential Green Chemistry Challenge Award. Research one of these award winners.
7. Formulation and final drug delivery form is an important consideration during the drug development stage. Consider what disease or circumstances would warrant a pill vs. an intravenous formulation. What physical characteristics of the drug substance would be necessary for an intravenous formulation?
8. During your drug discovery program, you learn that the thiophene fragment of your lead is extensively metabolized, resulting in very poor oral bioavailability. Unfortunately, thiophene contributes to the potency of the molecule. On the basis of the concepts described in Chapter 4, design two new replacements or changes to the thiophene that would address the metabolism issue but would likely maintain potency.



9. Developing new formulations of existing drugs has been one way to extend the patent life of certain drugs. These new formulations often benefit patients since they are often easier or more convenient to take. Research one of the particularly well-known examples of this successful strategy.
10. Many drugs progress to Phase II clinical trials and then fail for a variety of reasons. Two examples are preladenant and darapladib. Research one of the two drugs:
   a. Identify its intended use and drug target.
   b. Summarize preclinical and Phase I clinical data.
   c. Why did it fail in clinical trials?

FURTHER READING

Drug Development Costs

Dickson MD & Gagnon JP (2004) Key factors in the rising cost of new drug discovery and development. *Nat Rev Drug Discov* 3:417–429.

DiMasi JA, Grabowski HG & Hansen RW (2016) Innovation in the pharmaceutical industry: new estimates of R&D costs. *J Health Econ* 47:20–33.

IQVIA (2024) *Global Trends in R & D: Activity, Productivity, and Enablers*. IQVIA Institute Report.

Process Chemistry and Sitagliptin Synthesis

Desai AA (2011) Sitagliptin manufacture: a compelling tale of green chemistry, process intensification, and industrial asymmetric catalysis. *Angew Chem Int Ed* 50:1974–1976.

Griffin DJ, Coley CW, Frank SA et al. (2023) Opportunities for machine learning and artificial intelligence to advance synthetic drug substance process development. *Org Process Res Dev* 27:1868–1879.

Hansen KB, Hsiao Y, Xu F et al. (2009) Highly efficient asymmetric synthesis of sitagliptin. *J Am Chem Soc* 131:8798–8804.

Kim D, Wang L, Beconi M et al. (2005) (2*R*)-4-Oxo-4–[3-trifluoromethyl)-5,6-dihydro[1,2,4]triazolo[4,3-*a*]pyrazin-7(8*H*)-yl]-1–(2,4,5-trifluorophenyl)-butan-2-amine: a potent, orally active dipeptidyl peptidase IV inhibitor for the treatment of type 2 diabetes. *J Med Chem* 48:141–151.

Lexa KW, Kelyk KM, Henle J et al. (2021) Application of machine learning and reaction optimization for the iterative improvement of enantioselectivity of cinchona-derived phase transfer catalyst. *Org Process Res Dev* 26:670–682.

Savile CK, Janey JM, Mundorff EC et al. (2010) Biocatalytic asymmetric synthesis of chiral amines from ketones applied to sitagliptin manufacture. *Science* 329:305–309.

Pharmacokinetics and Toxicity

Jang GR, Harris RZ & Lau DT (2002) Pharmacokinetics and its role in small molecule drug discovery research. *Med Res Rev* 21:382–396.

Parasuraman S (2011) Toxicological screening. *J Pharmacol Pharmaceut* 2:74–79.

Food and Drug Administration and Clinical Trials

Adams CD & Brantner VV (2003) *New Drug Development: Estimating Entry from Human Clinical Trials*. Working Paper 262. Federal Trade.

The biopharmaceutics classification system (BCS) guidance (2016) http://www.fda.gov/AboutFDA/CentersOffices/OfficeofMedical-ProductsandTobacco/CDER/ucm128219.htm

Commission, bureau of economics. https://www.ftc.gov/reports/new-drug-development-estimating-entry-human-clinical-trials

The FDA's drug review process: ensuring drugs are safe and effective (2014) http://www.fda.gov/drugs/resourcesforyou/consumers/ucm143534.htm

Case Studies

Bell IM, Gallicchio SM, Wood MR et al. (2010) Discovery of MK-3207: a highly potent, orally bioavailable CGRP receeptor antagonist. *ACS Med Chem Lett* 1:24–29.

Dubowchik GM, Conway CM & Xin AW (2020) Blocking the CGRP pathway for acute and preventive treatment of migraine: the evolution of success. *J Med Chem* 63:6600–6623.

Perzborn E, Roehrig S, Straub A et al. (2011) The discovery and development of rivaroxaban, an oral, direct factor Xa inhibitor. *Nat Rev Drug Discov* 10:61–75.

Roehrig S, Straub A, Pohlmann J et al. (2005) Discovery of the novel antithrombotic agent 5-chloro-*N*-({(5*S*)-2-oxo-3–[4-(3-oxomorpholin-4-yl) phenyl]-1,3-oxazolidin-5-yl}methyl)thiophene-2-carboxamide (BAY 59–7939): an oral, direct factor Xa inhibitor. *J Med Chem* 48:5900–5908.

Yasuda N, Cleator E, Kosjek B et al. (2017) Practical asymmetric synthesis of a calcitonin gene-related peptide (CGRP) receptor antagonist ubrogepant. *Org Process Res Dev* 21:1851–1858.

PART II

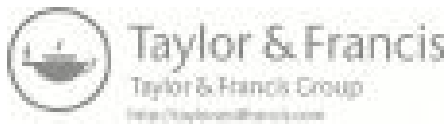
CLASSES OF DRUG TARGETS

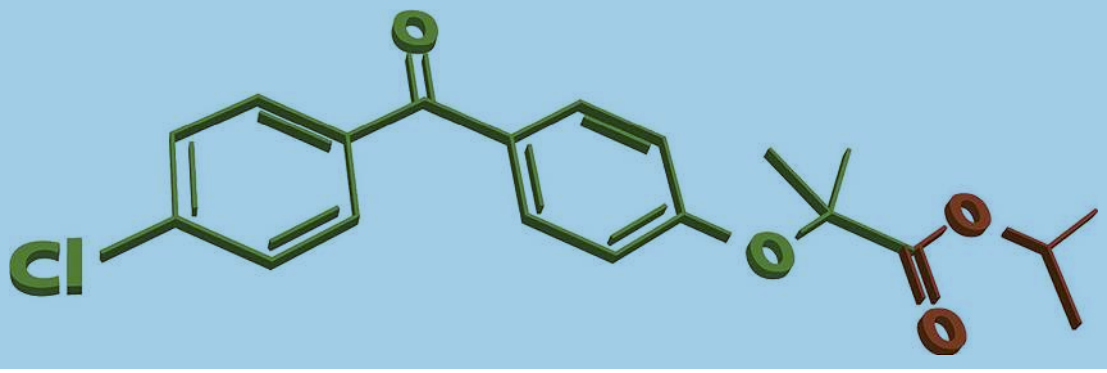
CHAPTER 6
Receptors, Ion Channels, and Transporters as Drug Targets

CHAPTER 7
Enzymes as Drug Targets

CHAPTER 8
Protein–Protein and Lipid Structure Interactions as Drug Targets

CHAPTER 9
DNA and RNA as Drug Targets


Taylor & Francis
Taylor & Francis Group



Receptors, Ion Channels, and Transporters as Drug Targets

6

LEARNING OBJECTIVES

- Describe the basic structural features of different classes of receptors, ion channels, and transporters.
- Understand the various ways small molecules can affect the activity of receptors, ion channels, and transporters.
- Become familiar with different types of assays used to categorize drug activity of G protein–coupled receptors, ion channels, nuclear receptors, and transporters.
- Develop a basic understanding of receptor theory.
- Learn examples of various drugs that act at G protein–coupled receptors, nuclear receptors, ion channels, and transporters.

6.1 INTRODUCTION

Based on current knowledge, there are approximately 3,000 genes in the human genome that code for proteins considered druggable, that is, capable of having their activity modulated by the binding of druglike compounds. This number is expanding, however, as medicinal chemistry evolves to develop and optimize methods and strategies aimed at interfering with historically undruggable targets, such as proteins that mediate physiological processes through protein–protein interactions. It has been estimated that between 600 and 1,500 proteins are involved in disease processes, and the actual number of human targets of approved drugs has recently been estimated to be over 650. Target proteins that are unique to pathogens number almost 200. Of the approximately 1,000 small-molecule drugs on the market, the vast majority bind to only approximately 600 macromolecular targets. Some specific anti-tumor and antiviral agents, as well as a few small-molecule therapeutics that treat genetic disorders, bind to nucleic acids. Of the human proteins targeted by drugs, G protein–coupled receptors (GPCRs) account for 12%, ion channels 19%, and nuclear receptors 3%. Kinases, which have both receptor and enzyme functions, account for 10% of the protein targets, according to a 2017 report. Another way to assess these figures is to consider the proportion of drugs that act at each protein target type. This analysis, which takes into consideration that there are often multiple drugs that modulate the same protein target, finds that 33% of drugs target a G protein–coupled receptor, approximately 24% act on enzymes (including kinases), 18% act at ion channels, 16% modulate nuclear receptors, and 6% target **transporters**. This analysis shows that over two-thirds of drugs bind to receptors as defined in the broadest sense. These target classes (Figure 6.1), as well as some historically undruggable targets, will be described in this and the following three chapters.

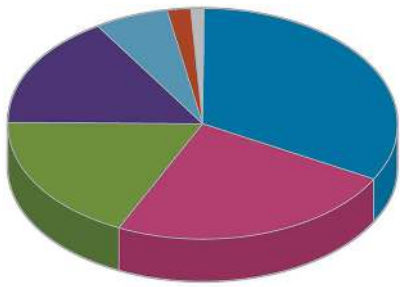


Figure 6.1 Distribution of macromolecular drug targets based on 2016 analysis. The majority of drugs bind to proteins, such as enzymes, GPCRs, ion channels, nuclear hormone receptors, and transporters, with a small percentage binding to nucleic acids.

DOI: 10.1201/ 9781003465607-8

In the most general sense, receptors, ion channels, and transporters are macromolecular protein complexes that propagate a signal or initiate subsequent events in response to some stimulus. The specific stimuli include binding of small molecules, hormones, or proteins, as well as changes in electrical membrane potential. The effects initiated can be as diverse as activation or deactivation of other proteins, up- or down-regulation of gene expression, protein localization, or release and uptake of chemical messengers, such as ions, small organic molecules, and even gases like nitric oxide. These macromolecules are a critical component of highly complex systems and are involved in physiological processes as diverse as neurotransmission, muscle contraction, and embryonic development. As such, they are frequent targets of drug discovery activities. Receptors can broadly include GPCRs, ion channels (ligand- and voltage-gated), transporters, and nuclear receptors. Receptor tyrosine kinases have both receptor and enzyme functions and will be discussed in Chapter 7. Among drugs that act at these receptors, the majority target GPCRs, followed by ion channels and nuclear hormone receptors (in similar proportion), and transporters. While these drug targets are similar in that they are proteins involved in signal transduction, GPCRs, ion channels, nuclear receptors, and transporters are distinct classes of drug targets, with each having specific structural features and functioning via specific mechanisms.

6.2 G PROTEIN–COUPLED RECEPTORS AS DRUG TARGETS

G protein–coupled receptors (**GPCRs**) span the cell membrane and act to relay messages from outside the cell to the inside. They are typically linked to an **effector system** that is usually an enzyme or ion channel. Occupation of the receptor via binding of its ligand (usually a small molecule, protein, or lipid) results in the activation or deactivation of the effector system, which, in turn, ultimately results in a change in concentration of a **second messenger**. This second messenger is often an ion or a small molecule. The second messenger then activates other enzymes or affects ion channels, resulting in a specific physiological effect (Figure 6.2). This entire process of an extracellular event (for example, binding of a small molecule) activating a membrane receptor and ultimately altering intracellular concentrations of specific molecules is termed **signal transduction**, where the **transducer** is typically the G protein or an arrestin.

The initial stimuli (the natural ligands) are typically small endogenous molecules, such as neurotransmitters (like dopamine and histamine), steroid hormones, small peptides, proteins, lipids, light, or even odorant molecules. The small-molecule second messengers are compounds whose levels are changed by the action of enzymes or by the result of ion channels opening or closing (effector systems) after binding of the ligand to the receptor. A common messenger, cyclic adenosine mono-phosphate (cAMP), is formed by the action of the enzyme adenylate cyclase on adenosine triphosphate (ATP). Two other common messengers are diacylglycerol (DAG) and inositol trisphosphate (IP_3), formed by phospholipase C (PLC)–catalyzed hydrolysis of a diacylphosphate (Figure 6.3).

It is estimated, on the basis of DNA sequences, that there are nearly 800 proteins that are considered GPCRs; however, only about 550 have an identified natural ligand. Those remaining, whose natural ligands are unknown, are termed **orphan GPCRs**.

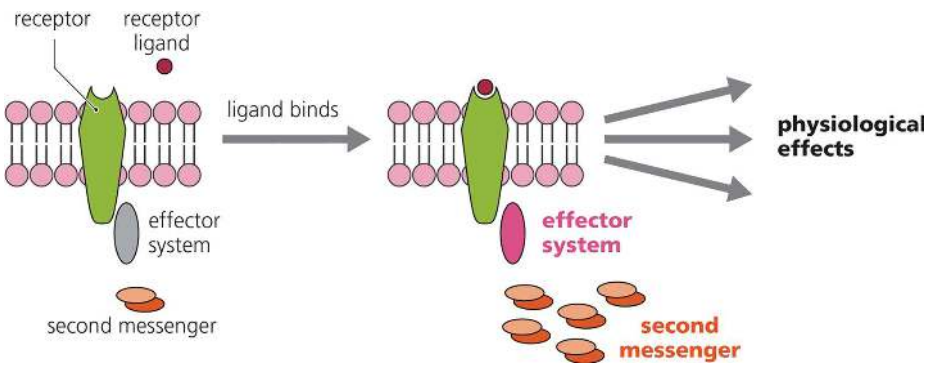


Figure 6.2 General scheme for ligand binding to GPCR and resultant signal. Unoccupied GPCR (green) is embedded in the plasma membrane (light pink); an effector system (gray) and second messengers (orange) sit inside of the cell. Binding of a ligand (red) at the extracellular ligand binding site results in a change in the effector system (now dark pink), and those changes, in the case illustrated, increase levels of second messengers. Changes in the concentration of second messengers result in physiological effects.

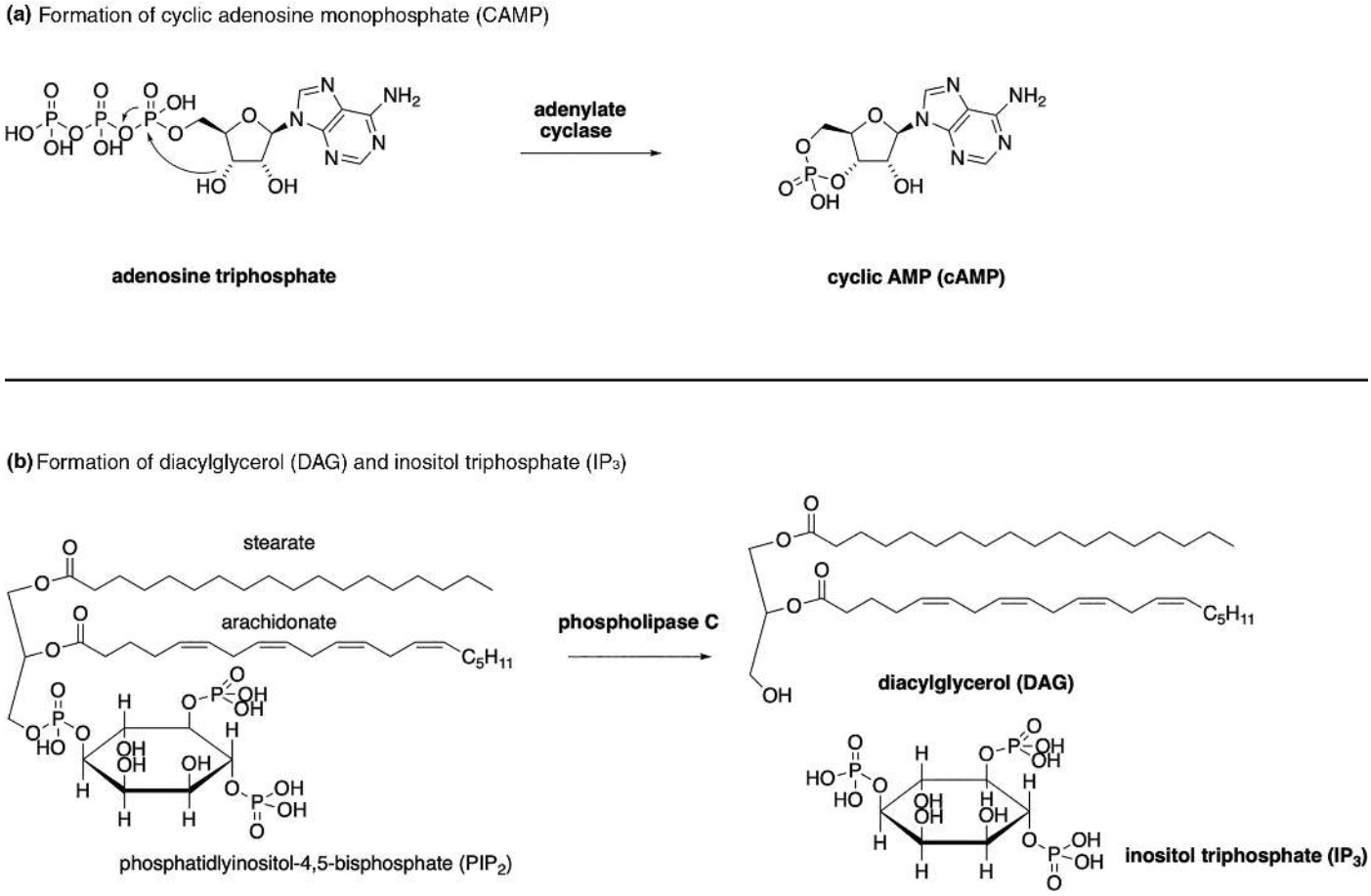


Figure 6.3 Formation of second messengers by effectors. (a) Generation of cyclic AMP (cAMP) from adenylate cyclase–catalyzed reaction of adenosine triphosphate (ATP). (b) Generation of diacylglycerol (DAG) and IP_3 (phosphatidyl inositol trisphosphate) from phospholipase C-catalyzed hydrolysis of phosphatidyl inositol bisphosphate (PIP_2).

6.2.1 GPCRs Are Defined by Structural Features

Human GPCRs fall into four classes based on sequence similarity and characteristics of their ligands: class A, B, C, and F. Most drugs on the market affect class A receptors, but other classes are also represented as drug targets. Of note, class B GPCRs were previously resistant to small-molecule drug discovery efforts; however, the recently approved "gepants" for migraine target the class B CGRP (calcitonin gene–related peptide) receptor (see Case Study 1 in Chapter 5). Other class B GPCRs are targets for peptide drugs like semaglutide and terzepatide that treat diabetes and obesity and are likely to succumb to small-molecule drug discovery efforts in the near future. Several important GPCRs in terms of drug action are β-adrenergic, opiate, serotonin, chemokine, angiotensin, histamine, and dopamine receptors. **Table 6.1** lists some examples of G protein–coupled receptors and their corresponding drugs and/or drug candidates.

Structurally, GPCRs are characterized by seven membrane-spanning helical domains arranged in a semicircular fashion. In most cases, it is believed the intracellular portion of the protein is associated with a GDP-bound form of a G protein (GDP-binding protein). The G proteins are composed of three subunits (Gα, Gβ, and Gγ), and multiple subtypes of each subunit exist. Binding of ligands usually occurs within this transmembrane array of helices and results in a conformational change of the receptor. Binding of the natural ligand and coupling to the G protein is illustrated in **Figure 6.4**. In this simplified model, binding of the ligand results in the α subunit of the G protein releasing a molecule of guanosine diphosphate (GDP) and then binding a molecule of GTP, resulting in the Gα subunit dissociating. Upon interaction of the Gα subunit or the Gβγ subunit with their effector systems, the signaling cascade is initiated and propagated, and downstream proteins are activated or deactivated. Enzyme effector systems include adenylate and guanylate cyclase, phosphodiesterases, phospholipases A_2 and C, and phosphoinositide-3-kinases. The activation or inactivation of these enzymes results in

Table 6.1 Examples of G Protein–Coupled Receptors, Their Natural Ligands, Drugs Targeting It, and the Approved Indication

Natural Ligand (Agonist)	Receptor	Drug or Ligand Examples	Indication
Class A			
Acetylcholine	Muscarinic (M_{1-5})	Pilocarpine	Glaucoma
Adenosine	A_1, A_{2A}, A_{2B}, A_3	Tecadenoson[a]	Atrial fibrillation
Angiotensin II	AT II	Losartan, valsartan	Hypertension
Anandamide	Cannabinoid (CB_1, CB_2)	δ^9-Tetrahydrocannabinol, rimonabant[a]	Weight loss
Chemokine	CCR5	Maraviroc	HIV/AIDS
Dopamine	D_{1a}, D_{1b}, D_{2-5}	Chlorpromazine, haloperidol	Psychosis
Histamine	H_1, H_2, H_3	Diphenhydramine (H_1)	Allergy
		Cimetidine (H_2)	Gastric ulcers
Adrenaline (epinephrine)	α_1 and α_2 adrenergic	Clonidine	Hypertension
		Phenylephrine	Nasal congestion
	β_1-, β_2-, β_3-adrenergic	Propranolol (β_1)	Hypertension
		Albuterol (β_2)	Asthma
Enkephalin, endorphins	μ-, δ-, κ-opioid	Morphine (μ)	Pain
		Loperamide (μ)	Diarrhea
Serotonin (5-HT)	5-HT_{1a-d}, 5-HT_2	Sumatriptan	Migraine
		Buspirone	Anxiety
	5-HT_4	Tegaserod	Irritable bowel
Somatostatin	SSTR	Pasireotide	Cushing's disease
Sphingosine 1-phosphate	S1PR1–5	Fingolimod	Multiple sclerosis
Vasopressin	V_1, V_2, V_3	Conivaptan	Hyponatremia
Gonadotropin-releasing hormone	GnRH-R	Elagolix	Endometriosis
Class B			
Calcitonin	CGRP receptor	Ubrogepant, atogepant	Migraine
Corticotropin-releasing factor	CRF_1	Antalarmin[a]	Antidepressant
Gastric inhibitory peptide	GIP receptor	Terzepatide	Diabetes, weight loss
Glucagon-like peptide	GLPR	Semaglutide	Diabetes, weight loss
Parathyroid hormone	PTH1R, PTH2R	Abaloparatide	Osteoporosis
Class C			
Glutamate	Metabotropic $GluR_{1-8}$	Eglumegad[a]	Anxiety
γ-aminobutyric acid	$GABA_B$	Baclofen	Alcoholism

[a]Compound is either in development or in clinical trials or has been removed from the market.

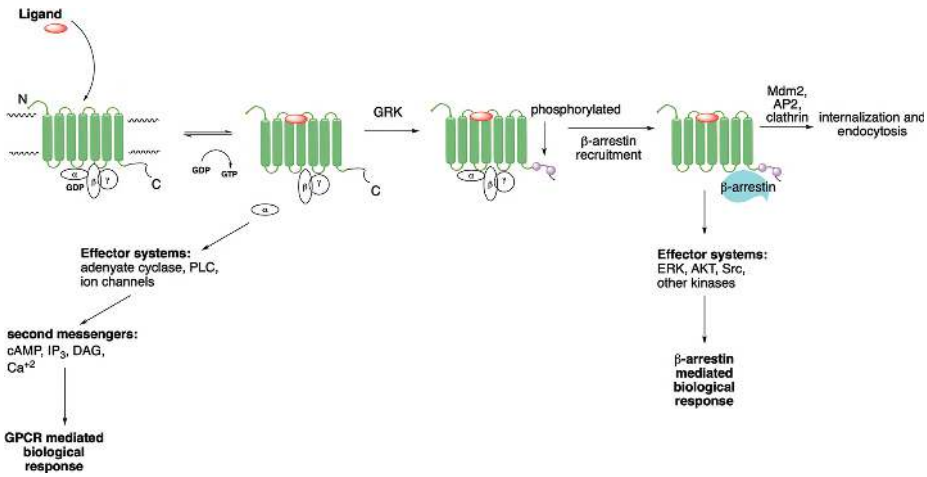


Figure 6.4 GPCR mechanism. Simplified GPCR signaling: GPCR (green), consisting of 7 membrane-spanning domains, is often associated with a GDP-bound form of a G protein, consisting of Gα, Gβ, and Gγ subunits (3 ovals). Upon ligand binding, the α subunit of the G protein releases GDP, which is then phosphorylated. The G protein binds GTP, and the Gα subunit dissociates. The individual subunits of the G protein interact with effector systems to change the levels of second messengers, which result in a GPCR-mediated biological response. Once the ligand is bound, GRK and β-arrestin signaling is also initiated: GPCRs are phosphorylated (purple) by GPCR kinases (GRK), and the phosphorylated receptor recruits β-arrestin (blue). Once β-arrestin is recruited, it blocks the G protein from further binding, resulting in either internalization, typically mediated by the proteins Mdm2, AP2, or clathrin, or the β-arrestin complex may initiate its own signaling pathway, mediated by kinase effector systems, such as ERK, AKT, or Src. This leads to a β-arrestin-mediated biological response.

changing levels of the second messengers, which then go on to elicit further biological responses, such as activating or deactivating transcription factors in the nucleus or changing the levels of phosphorylated proteins. This traditional view has been complicated by the fact that some GPCRs have been identified that do not bind G proteins and also that some cellular responses are not mediated by second-messenger levels.

Once activated, negative feedback loops turn off signaling and maintain balance. This process involves receptor phosphorylation on the cytoplasmic loops and C-terminus by G protein receptor kinases (GRKs). The phosphorylation enhances the binding of a protein called β-arrestin, which acts as a negative regulator of G protein–mediated signaling by way of sterically blocking G protein binding and also by receptor internalization (endocytosis). While the negative feedback by β-arrestins has been known for many years, more recently it has been found that they also can mediate their own signaling events through various kinases and other signaling proteins, and small molecules can dictate which signaling pathway prevails. This has implications for drug discovery and development, with much effort now placed on **biased ligands** that may affect either the G protein pathway or the β-arrestin pathway.

The ligand binding sites of GPCRs vary with the class of receptor. A model of each of the three major classes (A–C) is shown in **Figure 6.5**, where the cylinders represent transmembrane domains (TMD) 1–7. In class A receptors, the natural ligand binds in the TMD region. In class B receptors, which contain a large extracellular domain (ECD), ligands bind by the "hot dog in a bun" model in both the ECD and the TMD. In class C, which also has a large extracellular domain, the ligand binds only to the ECD by the "Venus flytrap" model.

An important advance came in 2000, when the first X-ray structure of a GPCR (rhodopsin) was solved. Prior to 2000, many computational homology models of various GPCRs had been developed on the basis of site-directed mutagenesis studies.

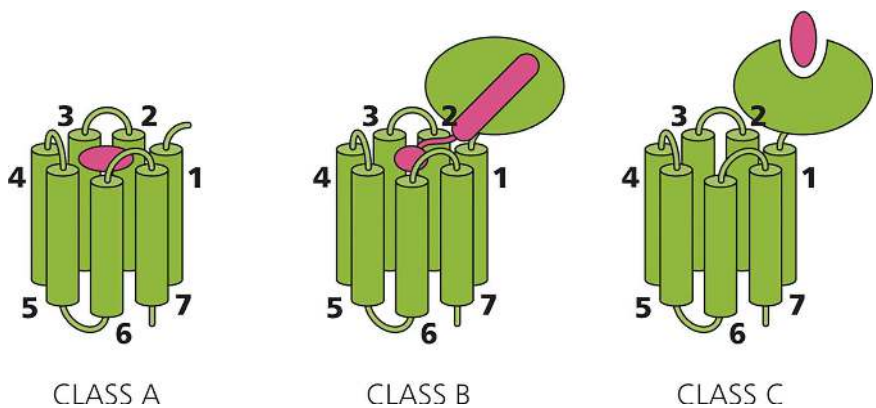


Figure 6.5 Binding sites in the major GPCR classes A, B, and C. Cylinders 1–7 (green) denote helical transmembrane domains (TMD) that all GPCR families have in common. Classes B and C contain large extracellular domains (ECD) (green ovals) compared to class A members. The ligand-binding domain (pink) is in the TMD for class A receptors, in both the ECD and TMD (linked) for class B receptors, and in the ECD in class C receptors.

Source: Adapted from Congreve, M.; Langmead, C. J.; Mason, J. S.; and Marshall, F. H. (2011). *J Med Chem* 54:4,283–4,311. With permission from the American Chemical Society.

Although very similar to most drug-targeted GPCRs, rhodopsin differs in that it binds its natural ligand 11-*cis*-retinal through covalent rather than noncovalent interactions. Since that time, over 100 X-ray and cryo-EM structures of GPCRs have been reported. The β_2 adrenoceptor bound to the partial inverse agonist carazolol was the first non-covalent ligand-mediated GPCR structure published and is shown in **Figure 6.6a**. The

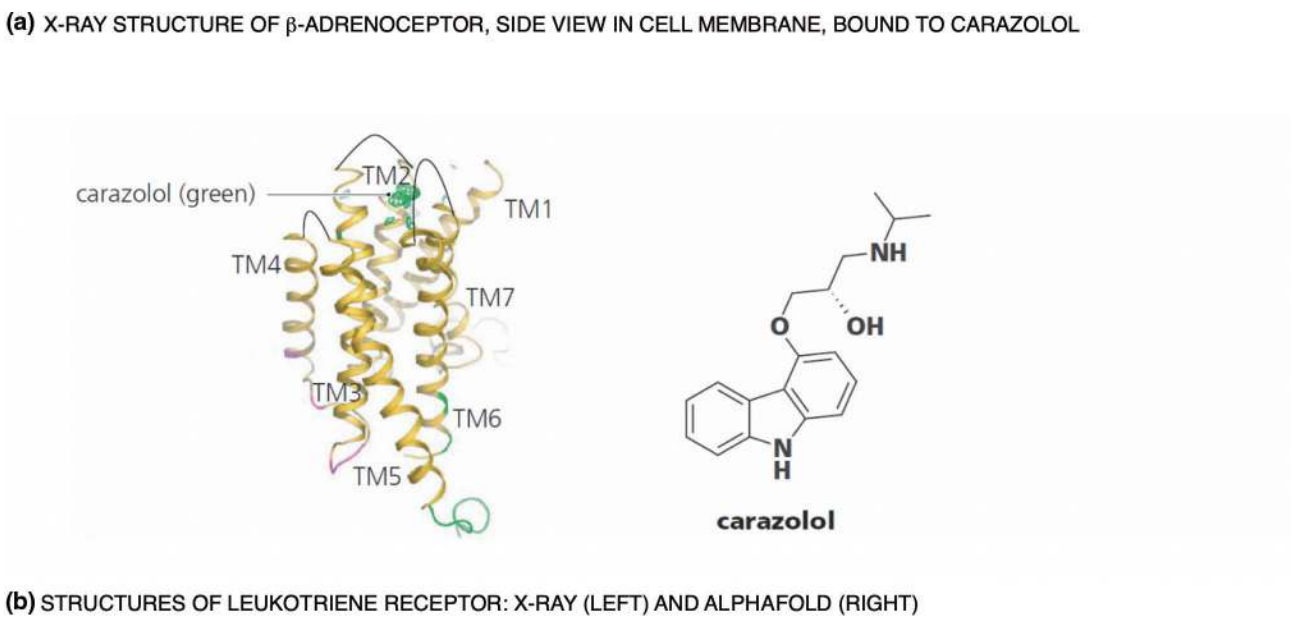


Figure 6.6 Structures of two GPCRs. (a) X-ray structure of the β_2-adrenoceptor bound to carazolol (*structure on right*), showing α-helices of transmembrane domains in yellow (TM1–TM7). Carazolol is in green at the top of the structure. (b) *Left*, structure of leukotriene receptor B1 (LTB41) bound to antagonist MK-D-046 (space-filling model). *Right*, structures of AlphaFold2 models (green and orange) superimposed on experimental leukotriene receptor X-ray (gray).

Source: (a) From Rasmussen, S. G. F.; Choi, H.-J.; Rosenbaum, D. M.; et al. (2007). *Nature* 450:383–387. With permission from Macmillan Publishers Ltd. (b) *Left*, from Michaelian, N.; Sadybekov, A.; Besserer-Offroy, E.; et al. (2021). *Nat. Comm.* 12:2,971. https://doi.org/10.1038/s41467-021-23149-1. *Right*, from Sala, D.; Hildebrand, P.W.; Meiler, J. (2023). *Frontiers in Molecular Biosciences*, DOI: *10.3389/fmolb.2023.1121962*.

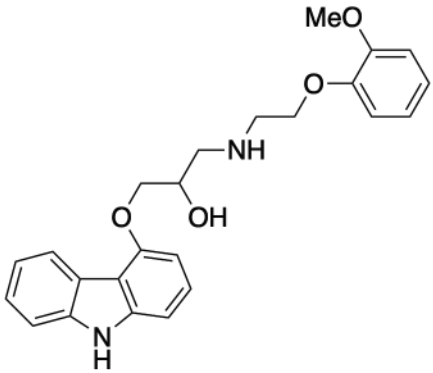
transmembrane domains (TM1–TM7) are evident, as are the extracellular and intracellular loops. More recently, structures of β_1 adrenoceptor, A_{2A} adenosine receptor, and a chemokine receptor have been solved, all of which are class A GPCRs. Recently, a class B GPCR, the calcitonin receptor, bound to a G protein and its natural ligand, calcitonin, was solved by cryo-EM. In addition, an X-ray structure of the smoothened receptor (SMO) has been solved, which is a class F GPCR, a small group of unique GPCRs.

While to date only some GPCRs have succumbed to structure elucidation, many others have been cloned, expressed, and subjected to site-directed mutagenesis, allowing further understanding of the overall structure and function of the receptor based on amino acid sequences. The availability of high-resolution structures as well as homology models provides some additional insight into drug design and medicinal chemistry strategies. In addition to experimental X-ray and cryo-EM structures, AlphaFold2 has become a valuable computational method to accurately predict multiple conformations of these membrane-bound receptors. The process uses databases of amino acid sequences to generate multiple sequence alignments (MSAs), then uses neural networks to generate a structure, and molecular dynamics simulations to resolve steric clashes and optimize bond geometries. An example of how accurate this predictive approach can be is shown in **Figure 6.6b**, which shows an X-ray crystal structure of the class A leukotriene B41 (LTB41) receptor bound to MK-D-046, an antagonist (A). **Figure 6.6b** shows two AlphaFold2-generated structures, in green and orange, superimposed on the actual X-ray structure, to exemplify the accuracy of the computational methods. Hit identification typically relies on ligand-based drug design derived from known small molecules or screening of focused libraries; however, with recent advances, structure-based design is becoming more feasible.

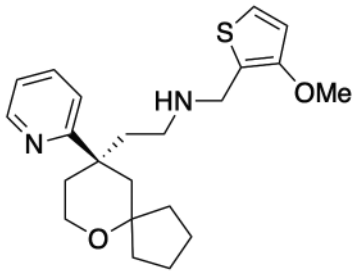
A particular challenge in GPCR-based drug discovery is the issue of selectivity. GPCRs are widely expressed throughout the body, and many different receptors respond to the same ligand. For example, there are currently six types of GPCRs that use 5-hydroxytryptamine (serotonin) as a ligand, and each of those includes multiple subtypes. Within just the 5-HT_1 GPCR are five subtypes: 5-HT_{1A}, 5-HT_{1B}, 5-HT_{1D}, 5-HT_{1E}, and 5-HT_{1F}. Other GPCRs have similarly large numbers of subtypes. Drug discovery activities often focus on optimizing selectivity on the basis of differences among the specific receptors and receptor subtypes. In addition, the location of the receptor in the body can also influence specificity.

In addition to selectivity of ligands for a particular GPCR, ligand selectivity for the G protein signaling pathway vs. the β-arrestin signaling pathway is now a focus of drug development. These **biased ligands** are in development for a number of conditions. In fact, some drugs that were classified as antagonists are now recognized as biased ligands. One example is the third-generation heart medication carvedilol. Classified as a member of the family of β-blockers, approved in 1995, carvedilol (seen in **Figure 6.7**) was shown in 2007 to stimulate the β-arrestin-mediated signaling pathway. An example of a drug intentionally developed as a biased agonist is oleceridine, a μ-opioid agonist that binds to the GPCR but has low β-arrestin recruitment. This pain medication, approved in 2020, has reduced respiratory suppression in comparison to morphine.

Because GPCRs are membrane-bound proteins, testing often relies on cellular assays or membrane preparations, where the particular receptor of interest is overexpressed. Prior to the availability of genetically engineered cells, specific tissue samples were utilized. Quantification of receptor binding affinity is done through **competition binding assays**, in which a labeled compound is incubated with a receptor preparation. The label can be either a radiolabel (for example, ^3H or ^{125}I) or a fluorescent label. After incubation, the ligand–receptor complex is isolated and quantified. For radiolabeled ligands, quantification may be done by filtration and radioactivity counting or by scintillation proximity assay (SPA), among other methods. Fluorescent techniques include fluorescence polarization assays and FRET assays. These binding assays provide no information on the ultimate effect of the molecule, and therefore, functional assays are also used to detect downstream effects. The activation of the G protein (a very early event) by detection of products of GDP/GTP exchange, second-messenger levels, and even further downstream events, such as translocation or transcription,



carvedilol



oleceridine

Figure 6.7 Structures of biased ligands carvedilol (activates β-arrestin signaling through β_1-adrenoceptor) and oleceridine (selectively activates G protein signaling through μ-opioid receptor).

can all be measured as part of a functional assay for GPCR activity. A frequently used technology involves detecting changes in the level of Ca^{2+} through the use of fluorescent or chemiluminescent dyes that are responsive to Ca^{2+} concentrations. Similarly, assays to assess signaling through β-arrestin can include proximity-based assays, which detect changes in the distance between the GPCR and β-arrestin, or functional assays to detect β-arrestin-specific effects, such as receptor internalization or other signaling. Ligand bias is then measured by comparing the effects from both G protein and β-arrestin pathways.

6.2.2 Ligands Bind to GPCRs and May Activate or Inactivate a Receptor

Ligands, typically small molecules, can bind to GPCRs and either activate or inactivate a receptor, and drugs acting at these receptors are broadly classified as either agonists (activators) or antagonists (blockers). An **agonist**, like the natural ligand, binds at the receptor, and the binding results in a physiological response. A **partial agonist** binds to the receptor but gives less than a full response in relation to some defined response by a full agonist or the natural ligand, regardless of increasing concentration. An **antagonist** (or a **neutral antagonist**) binds to the receptor and prevents the natural ligand from eliciting a response. When no agonist is present, a neutral antagonist will have no effect. An **inverse agonist**, first described by Tommaso Costa and Albert Herz for GPCRs in 1989, binds and results in a decrease in the response of a constitutively active receptor: A constitutively active receptor transduces a small amount of signal even in the absence of an activating ligand. In the presence of an agonist or the natural ligand, an inverse agonist will decrease the response, even below basal levels. The effect of these types of ligand is shown in Figure 6.8. A **biased ligand**, unlike a

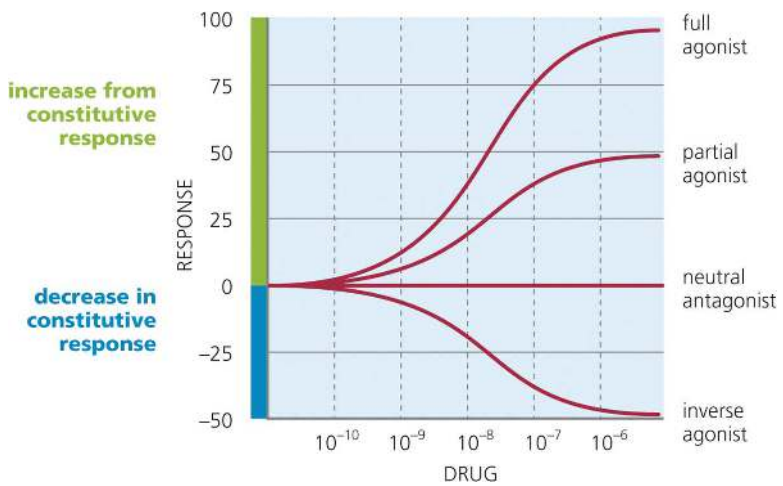


Figure 6.8 Effect of increasing concentrations of various types of ligands. The biological response of a receptor that has constitutive activity is set at zero in the graph. A full agonist, at some concentration, will afford a complete (100%) response; a partial agonist will give a less-than-full response, despite increasing concentrations; a neutral antagonist will cause no response and will prevent an agonist from activating the receptor; and an inverse agonist will decrease the constitutive or basal response.

Source: Adapted from Boghog/Wikimedia Commons, CC BY-SA 4.0.

pure agonist or a **balanced ligand**, results in the activation of some of the GPCR's signaling cascade. When it activates the β-arrestin signaling pathway, it may mimic an agonist; when it activates the internalization pathway, there is a decrease of the GPCR signal, and it appears to be an antagonist. Given the complexities of these signaling pathways, a battery of assays is required to correctly characterize compounds' effects. Highlighted in what follows are some classical ways to characterize GPCR ligands.

Drugs may bind either at **orthosteric sites** (same as natural ligand), where they compete with the natural ligand for binding, or at **allosteric sites** that are remote from where the natural ligand binds. Drugs that bind at allosteric sites may act as either **positive allosteric modulators** (PAMs) or **negative allosteric modulators** (NAMs), depending on whether they increase or decrease activity of the natural ligand or the agonist. For any type of drug, there are two important components to bioactivity: First, the drug must have binding affinity for the receptor, and second, upon binding, it must have efficacy or an effect on signal transduction. Typically, a PAM will increase the affinity and/or the efficacy of the agonist, and a NAM will decrease the affinity and/or the efficacy of the agonist. However, another possibility is a **PAM antagonist**, which will increase affinity of the agonist but decrease efficacy.

The **affinity**, or binding strength, of a drug that acts at a receptor can be measured in a binding assay. This information can reflect both strength of binding and site of binding (orthosteric vs. allosteric). The strength of binding (affinity) depends on the dissociation of the drug–receptor complex, as outlined in Figure 6.9a, where A = drug, R = receptor, and AR = drug–receptor complex (see Chapter 3). In practice, this can be determined

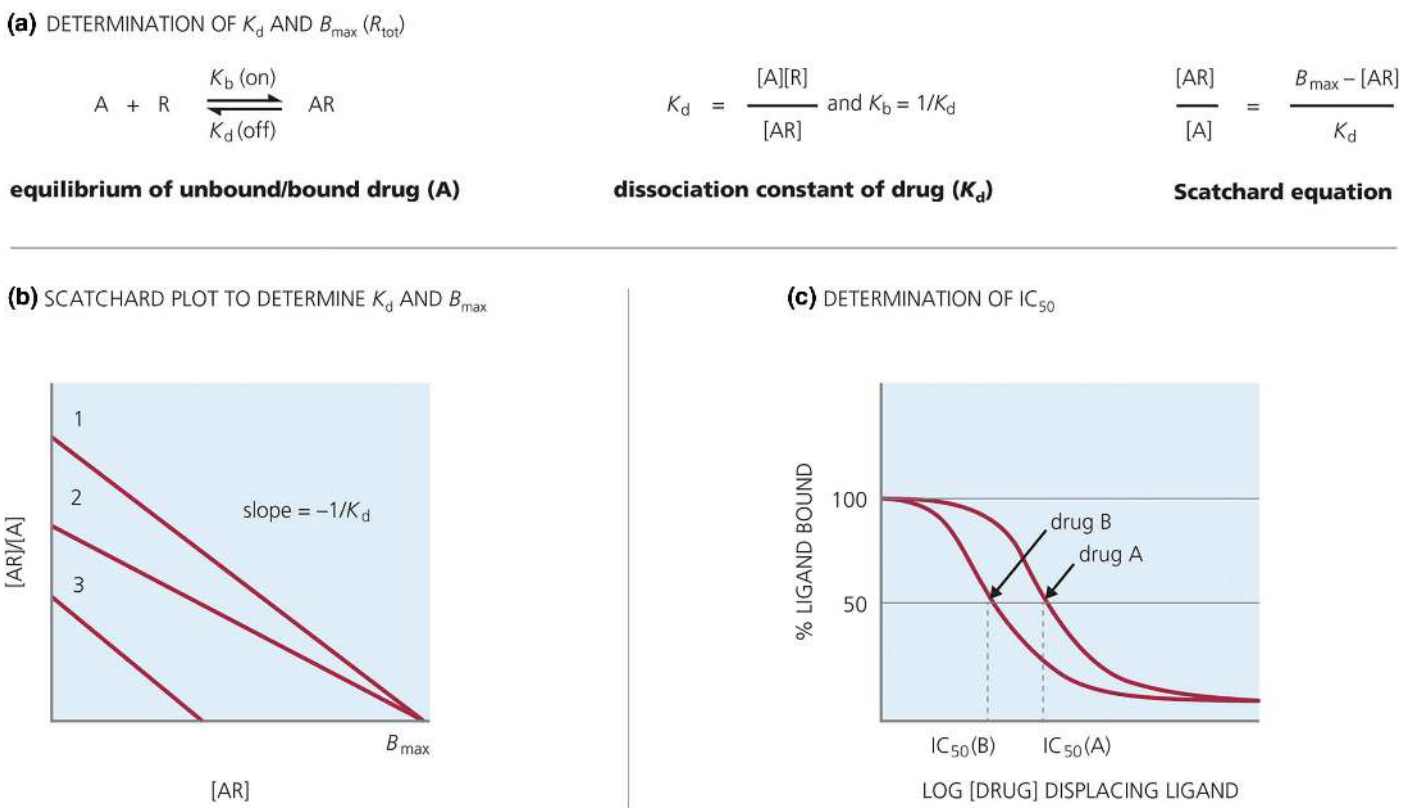


Figure 6.9 Binding affinity determinations. (a) Equations used to determine K_d and B_{max} for a ligand, A binding to receptor R. The dissociation constant K_d can be derived from a series of experiments measuring the amount of bound drug in the presence of various concentrations of a known ligand. The binding constant K_b is the inverse of the dissociation constant ($1/K_d$). (b) Scatchard plots of [AR]/[A] vs. [AR] for (1) a known radiolabeled ligand, (2) a compound binding at an orthosteric site, and (3) a compound binding at an allosteric site. The slope of each line = $-1/K_d$, and the x-intercept = B_{max}. The line for the allosteric compound (3) will not intersect B_{max} (total number of binding sites) since it binds at a different site than in cases 1 and 2. (c) Graph of concentration of known bound ligand vs. concentration of two drugs tested on a log scale. The concentration of drug that displaces 50% of the labeled ligand is known as IC_{50}. Drug B is more potent than drug A because it displaces 50% of the labeled ligand at a lower concentration than drug A.

by using a known radiolabeled ligand, allowing it to equilibrate with the receptor, and then washing to remove unbound ligand. While the concentrations of [A] (free ligand) and [AR] (bound ligand) can be measured because A is labeled, [R] (unoccupied receptor) cannot be measured and is equal to $[R]_{tot}$ – [AR], or the total number of receptors (R_{tot}) minus those bound to the ligand (R_{tot} is also known as B_{max}). The dissociation constant (K_d) is determined by measuring [AR]/[A] at varying concentrations of the known ligand and plotting the ratio vs. [AR] in a Scatchard plot. The binding constant (K_b) is the reciprocal of the K_d. Repeating the experiment in the presence of the unlabeled ligand indicates if the drug competes with the known ligand, and therefore, whether it works via an orthosteric or allosteric mechanism. In a Scatchard plot, the slope of a line reflects affinity of the drug for the receptor, and the *x*-intercept is equal to B_{max}.

In the Scatchard plot in **Figure 6.9b**, line 1 is characteristic of an agonist and was generated from a known radiolabeled ligand. Line 2 is characteristic of a drug binding at an orthosteric site, but with different affinity (K_d) compared to the compound evaluated in line 1. The slope of line 2 is different than that of line 1, but with enough radiolabeled ligand added, eventually all receptors will be displaced and the line will cross the *x*-axis at B_{max}. Line 3 is characteristic of a drug (a PAM or NAM) binding at an allosteric site. In this example, adding a radiolabeled ligand will not displace the drug, but a conformational change makes fewer binding sites available for the radiolabeled ligand, and B_{max} (*x*-intercept) decreases while the slope remains the same as the line 1.

An alternative plot compares the percent of bound known ligand vs. the concentration of drug that decreases binding of that ligand (**Figure 6.9c**). The amount of drug required to decrease binding of a standard by 50% is known as **IC_{50} (inhibitory concentration)**. A drug with higher binding affinity (drug B) will have a lower IC_{50} than a drug with lower binding affinity (drug A). An alternative value is pIC_{50}, which is the negative log of the IC_{50}, which can be readily derived from the graph in **Figure 6.9c**.

6.2.3 Drug Activity Is Classified by Examining Biological Response as Compared to Dose

The effect of a drug relies on efficacy as well as binding affinity, and in order to measure the biological effects of a drug, dose–response curves, such as those in **Figure 6.10**, are generated, where the physiological effect, such as changes in cAMP or Ca^{2+} levels or changes in receptor internalization mediated by β-arrestin, is measured as a function of compound concentration. The **EC_{50} (effective concentration)** of the drug is the concentration that produces 50% of maximal response and is derived from the graph. In practice, this value may be expressed as the pEC_{50} (known in earlier literature as pD_2), which, like pIC_{50}, is the negative logarithm of EC_{50} and expresses the potency of a drug. As in **Figure 6.7**, if the agonist generates the same maximal response (E_{obs}) as the natural ligand or a defined full agonist (E_{max}), it is considered 100% efficacious.

By use of these types of dose–response curves, the activity profile of a drug or small-molecule probe can be determined, with different shapes of curves characteristic of agonists, partial agonists, neutral antagonists, and inverse agonists. A full agonist

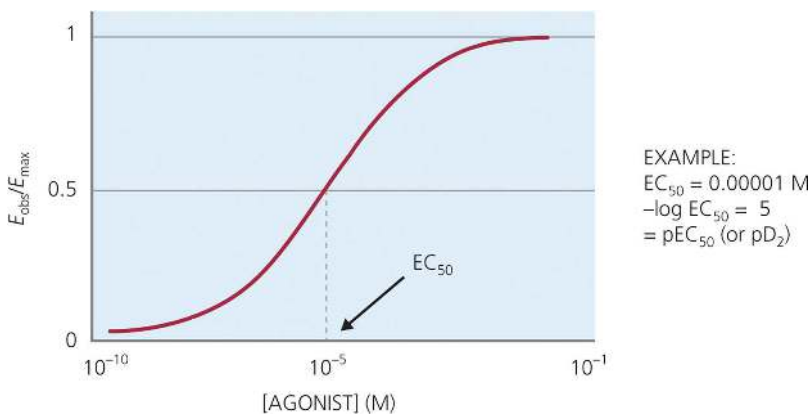


Figure 6.10 Determination of EC_{50} of a full agonist. Concentration of the compound is plotted vs. defined response (E_{obs}) as a fraction of maximal response (= 1.0; E_{max}). EC_{50} is the dose producing a 50% response. In this example, EC_{50} is 1×10^{-5} M. This may be expressed as –log EC_{50}, which is pEC_{50} (also known as pD_2).

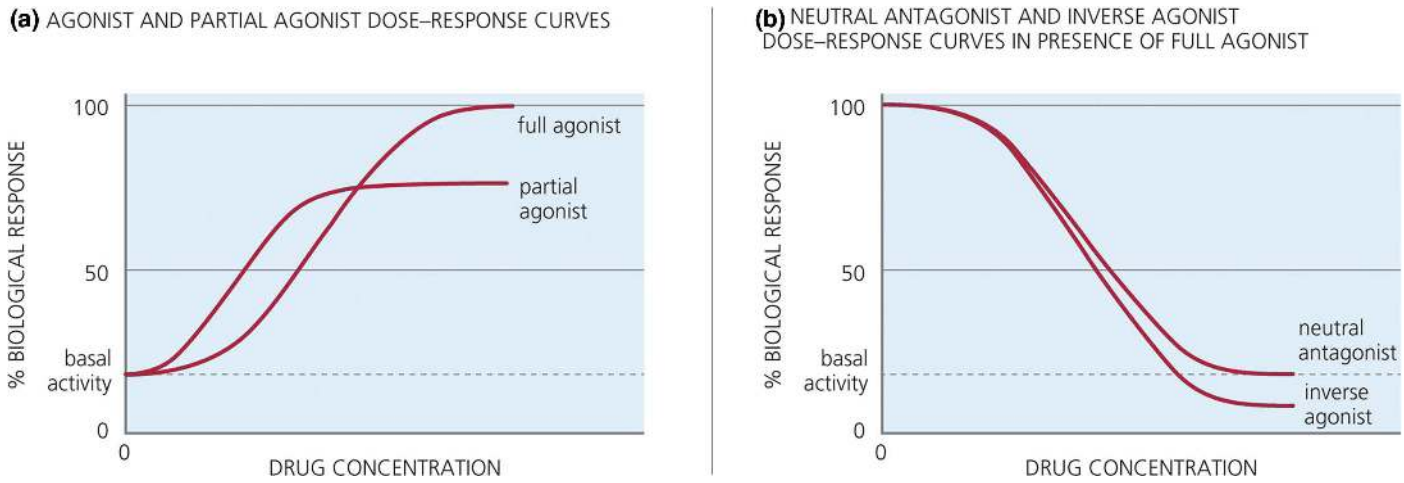


Figure 6.11 Theoretical full agonist, partial agonist, neutral antagonist, and inverse agonist dose–response curves. (a) Full agonists will achieve the maximal response at a given concentration; partial agonists never achieve 100% response regardless of increasing concentration. (b) To measure a dose–response curve, neutral antagonists and inverse agonists are evaluated in the presence of an agonist-activated receptor; therefore, when no drug is present, there is 100% response. A neutral antagonist results in a reduction in response to the basal level with increasing concentration. An inverse agonist results in a lower response than the system when no natural ligand is present and requires that the receptor have a level of constitutive activity.

will reach a 100% response (considered 100% efficacious) as concentration increases. A partial agonist will increase response and plateau at a certain level of efficacy, never reaching 100% (**Figure 6.11a**). A neutral antagonist and an inverse agonist act in the presence of an agonist (such as the natural ligand); therefore, the response when no drug is present is, by definition, 100%. A neutral antagonist decreases the biological response to the baseline in a dose-dependent manner; an inverse agonist will further decrease the constitutive response, as seen in **Figure 6.11b**.

A drug discovery program, depending on the specific receptor target, may be aimed at modulating the activity of a GPCR in a variety of ways. For example, there are several GPCR targets that are involved in controlling serotonin levels, with different effects on the central nervous system. The anti-migraine drugs called triptans are agonists at the 5-$HT_{1B/1D}$ receptors and have the effect of constricting blood vessels in the brain, resulting in headache relief. Examples of antagonists at 5-HT receptors include some antipsychotic drugs, such as risperidone, that act on the 5-HT_{2A} receptor, and others, such as pimavanserin, are inverse agonists at the same (5-HT_{2A}) receptor. The serotonin receptor–binding drugs are just one example of the complexity of receptor subtypes and the biological effects of agonists, partial agonists, antagonists, and inverse agonists.

The effect of a competitive antagonist can be measured by generating a dose–response curve for a known agonist (E_{obs}/E_{max} vs. concentration of agonist) in the presence of different fixed concentrations of the antagonist (**Figure 6.12a**). Since a competitive antagonist directly prevents the natural or known ligand from binding and causing a response, increasing concentrations of the antagonist shift the EC_{50} of the agonist to the right. In **Figure 6.12a**, the EC_{50} of the agonist alone is estimated to be 10^{-5}; at each higher concentration of antagonist (lines 2–4), it consistently shifts to larger (less potent) numbers.

A noncompetitive antagonist occupying an allosteric site may reduce or prevent the response caused by an agonist but may not prevent its binding. **Figure 6.12b** shows a dose–response curve for an agonist in the presence of different fixed concentrations (curves 2, low, and 3, medium) of a noncompetitive antagonist. As in the case of a competitive antagonist, the EC_{50} of the agonist shifts in the presence of a fixed concentration of noncompetitive antagonist. However, in this case, since the noncompetitive antagonist does not compete with the agonist, maximal effect is never reached.

The ratio of the dose of agonist required to give an effect at a defined antagonist concentration (D_2) compared to the dose of agonist required to give the same effect without a competitive antagonist present (D_1) is the **dose ratio**, $r = D_2/D_1$. Using the dose ratio obtained from the first plot (**Figure 6.12a**), log ($r - 1$) vs. log [antagonist] is graphed as a Schild plot (**Figure 6.12c**), based on the Schild equation (6.1), which gives an estimate of the antagonist binding constant K_b, often expressed as $pA_2 = -\log K_b$.

$$\log(r-1) = \log[\text{antagonist}] - \log K_b \,(\text{or} + pA_2) \qquad (6.1)$$

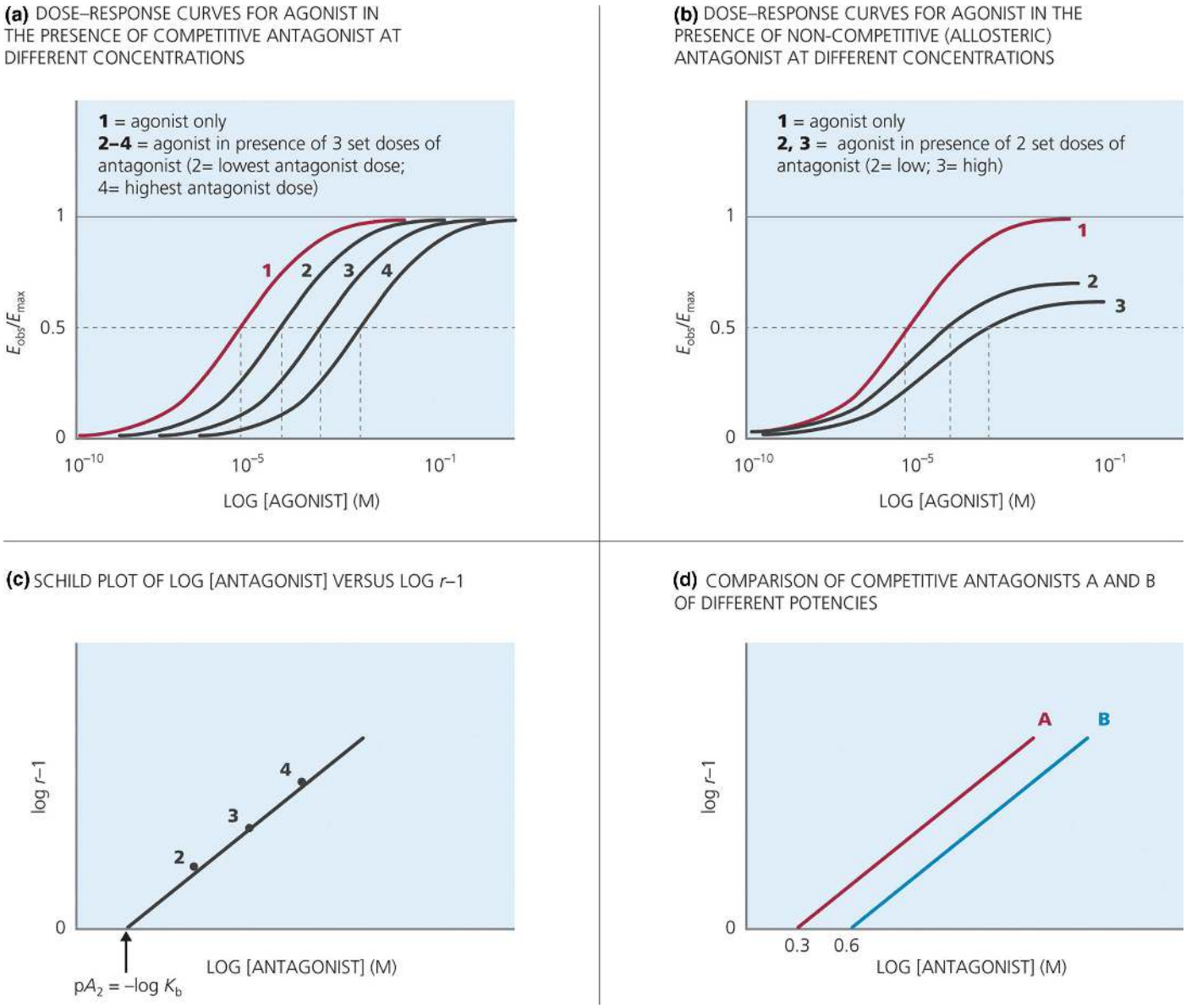


Figure 6.12 Characterization of antagonists based on dose–response curves.
(a) Dose–response curves relating the biological effect E_{obs}/E_{max} to increasing concentrations of agonist in the presence of different fixed concentrations of a competitive antagonist. Curve 1 (red) shows the effect of agonist only; curves 2–4 show the effects of three different increasing antagonist concentrations on the agonist dose–response curve. Dashed lines to the *x*-axis show agonist EC_{50} for each concentration of antagonist present. The right shift of the dose–response curves as the concentration of antagonist increases is characteristic of a competitive antagonist. (b) Dose–response curves for agonist in the presence of different fixed concentrations (curves 2 and 3) of noncompetitive antagonist; increasing agonist concentration will not achieve full biological effect. (c) Schild plot of antagonist concentration vs. log (dose ratio − 1), generated from agonist dose–response curves 2–4 in graph A. The *x*-intercept = pA_2, which is the concentration of antagonist, where $r = 2$ (pA_2 may be used to estimate binding constant K_b, where $pA_2 = -\log K_b$). (d) Comparison of Schild plots of two competitive antagonists of different potencies, where the higher pA_2 (*x*-intercept) corresponds to higher potency.

Solving the equation for when the dose of agonist must be doubled to reach full effect when $r = 2$ and $\log (r - 1) = 0$, then the *x*-intercept equals pA_2 for the antagonist. Because the value is on a log scale, a more potent antagonist will correspond to a higher pA_2 value. In Figure 6.12d, compound A has $pA_2 = 0.3$ M and compound B has $pA_2 = 0.6$ M. Comparing the pA_2 values, twice as much of compound A was required to achieve a dose ratio of 2 as compound B; therefore, compound B is the more potent antagonist. pA_2 may be used as an estimate of the binding constant, $-\log K_b$, based on certain assumptions, which include competitive antagonism, reversible binding of

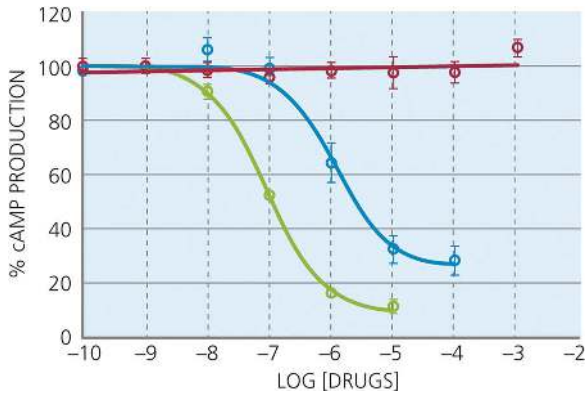


Figure 6.13 Dose–response curves showing effect of drugs acting at the H_2 receptor. Plots show percent of cAMP (the basal level set at 100%) vs. log concentration of compound. Ranitidine (shown in green) and cimetidine (shown in blue) are both inverse agonists. Burimamide (shown in red) does not change the basal level of cAMP and therefore is a neutral antagonist.

Source: Adapted from Smit, M. J.; Leurs, R.; Alewijnse, A. E.; et al. (1996). *Proc Natl Acad Sci USA* 93:6,802–6,807. With permission from the National Academy of Sciences.

agonist and antagonist, a system at equilibrium, and a biological response dependent upon occupancy and activation of the receptor.

It can be difficult to differentiate a competitive antagonist from an inverse agonist on the basis of a physiological response, as most often an agonist is present, and both reduce the signaling event. In order to prove neutral antagonism, a system must be defined by an inverse agonist, since it is necessary to know if the receptor's constitutive level of activity can be reduced (inverse agonist) or will not change from the basal level (neutral antagonist). For this reason, many compounds originally classified as antagonists are being reclassified with further work. In fact, a 2004 study of 380 compounds originally described as competitive antagonists found that newly reported data showed that 85% of those have been reclassified as inverse agonists. For example, in a study done in 1996, it was shown that both cimetidine and ranitidine, drugs originally classified as neutral antagonists of the H_2 receptor, are inverse agonists. In this case, new data showed that the H_2 receptor exhibits basal activity, so a reduction in cAMP levels when no agonist is present can be observed (**Figure 6.13**). However, an early clinical candidate that also acts at the H_2 receptor, burimamide, is a neutral antagonist.

There are several ways to assay for biased ligands. After controlling for multiple factors that can influence results, dose–response curves are obtained for both G protein and β-arrestin responses. In cases where the receptor and system are unbiased, a balanced agonist would show equivalent potency for the two different pathways (**Figure 6.14a**). In contrast, a β-arrestin-biased agonist would produce a leftward shift in the β-arrestin curve vs. the G protein curve, reflecting that the ligand is more potent at β-arrestin signaling than G protein signaling (**Figure 6.14b**). Similarly, a G

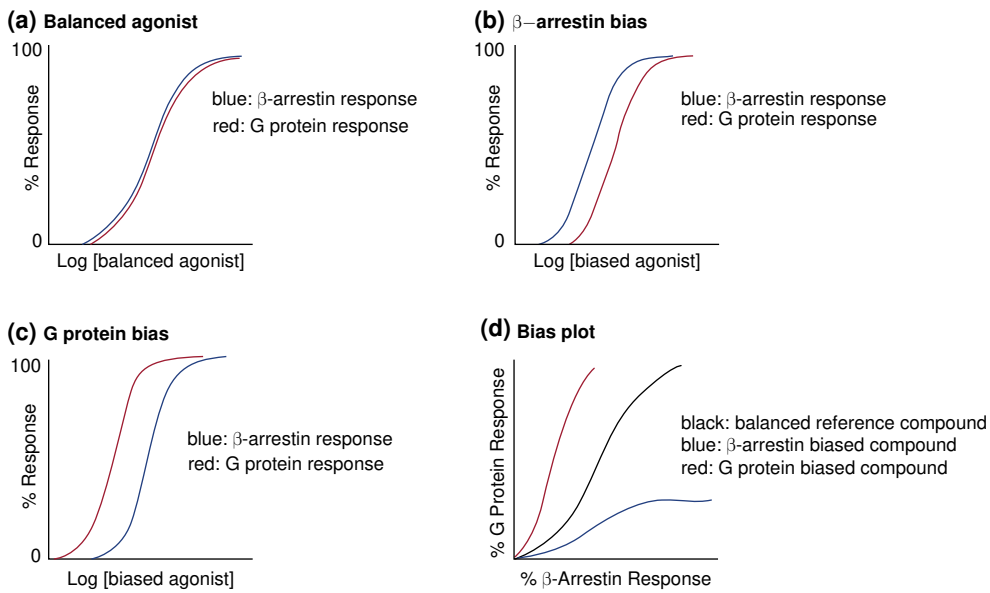


Figure 6.14 Plots for balanced and biased agonists: β-arrestin response (blue) and G protein response (red). (a) Dose–response curves for both β-arrestin response and G protein response for a balanced ligand. (b) Dose–response curves for both β-arrestin response and G protein response for a β-arrestin-biased ligand. (c) Dose–response curves for both β-arrestin response and G protein response for G protein–biased ligand. (d) Bias plot of the efficacy at G protein vs. β-arrestin pathways at identical concentrations showing a drug (blue) with a bias toward β-arrestin and a drug (red) with a bias toward G protein, both vs. a reference balanced compound (black).

protein–biased agonist would show a leftward shift (**Figure 6.14c**). A bias plot can then be generated based on the percent of G protein vs. β-arrestin response at identical ligand concentrations in comparison to a balanced ligand (**Figure 6.14d**). A biased ligand will show deviations from the balanced ligand depending on which response it is biased toward. Alternatively, for biased ligands, a **bias factor β** can be calculated based on plotting dose–response curves for each different signaling pathway against each other to evaluate toward which pathway the ligand is biased and by what degree. This factor is equal to the log of the relative activity of one path (A) divided by the relative activity of the other path (B), so that a bias factor of 1 indicates 10-fold higher signaling by way of path A, and a bias factor of 0 indicates a balanced ligand.

6.2.4 Theories Have Been Developed to Explain How Drugs Binding at the Same Receptor Site May Have Different Types of Activity

Theories of how ligand–receptor and drug–receptor interactions affect bioactivity have evolved over the past century, with the goal of explaining how different drugs may bind to the same receptor but have different levels of activity and even different downstream signaling effects. In 1937, A. J. Clark was the first to apply chemical laws to quantify drug–receptor activity, proposing that the fraction of receptors occupied was directly proportional to the biological response. Modification of this theory by Everhardus Ariens and R. P. Stephenson, and others, led to the current version of the **occupation theory**, which explains how partial agonists can occupy all receptors but not produce a full response. A key concept is that of **efficacy** or **intrinsic activity**: the degree to which a drug, by binding to a receptor, has the capacity to result in a biological response. In this theory, the biological activity of a molecule is not only directly proportional to the fraction of occupied receptors but is also dependent on the molecule's intrinsic activity. The **rate theory**, developed by W. D. M. Paton and H. P. Rang, proposed that activity is a function of the rate of association and dissociation and not just the number of occupied receptors.

Multiple models of GPCR action have been developed and adopted, and each has been refined over time. In 1993, the **extended ternary complex** (ETC) model was first proposed by Robert Lefkowitz, who, together with Brian Kobilka, was awarded the Nobel Prize in Chemistry 2012 "for studies of G protein–coupled receptors." In this model, without a ligand (A) bound, three forms of the receptor exist: inactive (R_i), active (R_a), and active/bound to a transducer (R_a-T), *T* being either a G protein or a β-arrestin, with the signal being transduced from the R_a or R_a-T state. In the absence of a ligand, the equilibrium between active and inactive receptors lies far toward the inactive state (R_i), and little, if any, signal is transduced. In the case of constitutively active receptors, a small percentage of receptors are in the active states, R_a or R_a-T, accounting for the modest amount of signal transduction. Binding of a ligand (A) to any one of the three forms changes the resting equilibrium (based on the equilibrium constants K_a and K_T) and leads to changes in signal transduction. This model is outlined in **Figure 6.15**.

This theory provides a framework for how molecules that bind to the same putative site can exert vastly different effects, although it is a simplification and it is now understood that there are multiple reception conformations, not simply one active and one inactive conformation. In the ETC model, an agonist has high affinity for the

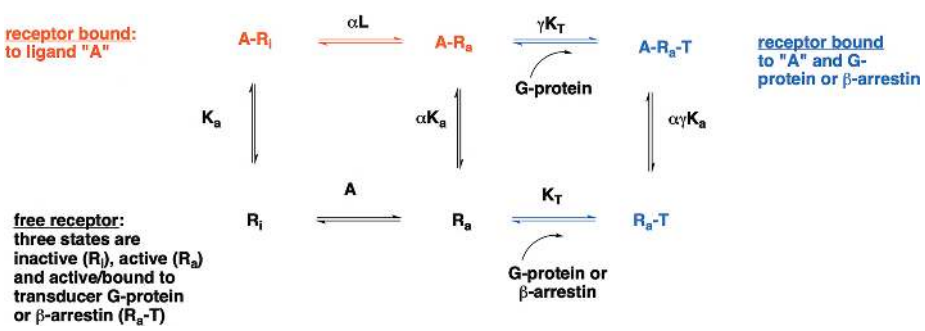


Figure 6.15 Extended ternary complex model for GPCRs. The three components are *R* = receptor; *A* = ligand; and *T* = transducer (G protein or β-arrestin). Three forms of free receptor (R_i, R_a, and R_a-T), shown at the bottom, are in equilibrium with one another; when ligand A is present, there are three additional forms bound to ligand A (A-R_i, A-R_a, and A-R_a-T), shown at the top. Equilibrium constants K_a, αK_a, and $\alpha\gamma K_a$ characterize the binding of ligand A to each form of the receptor. K_T and γK_T are characteristic of binding of the G protein or β-arrestin to the free active form (R_a) and the bound active form (A-R_a), respectively. In this model, the terms α and γ are parameters characteristic of the specific ligand A and reflect affinity and efficacy, and *L* is an allosteric constant that describes the natural propensity of the inactive receptor R_i to adopt the active state R_a.

active R_a state and, upon binding, stabilizes the active form, shifts the equilibrium away from R_i to A-R_a and R_a-T when T = the G protein, and increases the amount of signal transduced. In the case of a partial agonist, the equilibrium also shifts toward A-R_a and R_a-T, but not to the same extent as a full agonist. An inverse agonist has high affinity for the R_i state and stabilizes the inactive form, shifting the equilibrium even further away from R_a and R_a-T, and thereby reducing any constitutive activity. A neutral antagonist binds but does not change the equilibrium. A β-arrestin-biased ligand, like an agonist, also shifts the equilibrium away from the inactive state but stabilizes the A-R_a and R_a-T when T = β-arrestin. Differences in activity, as evidenced by dose–response curves, result from affinity for different forms of the receptor, and the parameters α and γ in the ETC are characteristic of the specific ligand and reflect activity and efficacy. Equilibrium constants K_a, αK_a, and $\alpha\gamma K_a$ characterize binding of ligand A to each form of the receptor.

The ETC model relies on three defined (unoccupied) receptor species as the resting population. A different way to explain ligand–receptor behavior is the use of a **probability model** or a model based on ensembles of conformations. In this model, there is a large distribution (rather than just three) of resting receptor states or micro-conformations, and binding of a ligand shifts the overall distribution, increasing some states and decreasing others.

6.2.5 Drugs Acting at GPCRs Include Drugs to Treat Allergy, Cardiovascular Disease, Asthma, and Ulcers

Many early drugs that act at class A GPCRs were discovered serendipitously or by using whole-animal models of diseases rather than by target-based drug design. Their mechanisms and specific targets were only determined well after the drugs were approved and widely used, and as described earlier, in many cases, the characterization of these molecules as an agonist or antagonist has been revised. More recently, drug discovery for GPCRs has relied on design based on the natural ligand when the natural ligand is a small molecule. With the availability of technology that allows for the efficient screening of GPCRs for functional activity, focused screening of specialty libraries is also common. The forefront of GPCR drug discovery takes advantages of the very recent advances in structural biology and the availability of high-resolution structures of small molecules bound to GPCRs.

The early antihistamine drugs, such as diphenhydramine (**Figure 6.16**), act as antagonists of the natural ligand histamine primarily on the H_1 subtype, with activity in the periphery against the itching and the inflammation associated with allergies. However, they also act on H_1 receptors in the central nervous system (CNS), leading to sedation, and they have effects as an antagonist at another GPCR, the muscarinic receptor. Newer drugs, such as cetirizine and loratadine, do not penetrate the CNS and thus exert their activity only on peripheral receptors, so patients experience much

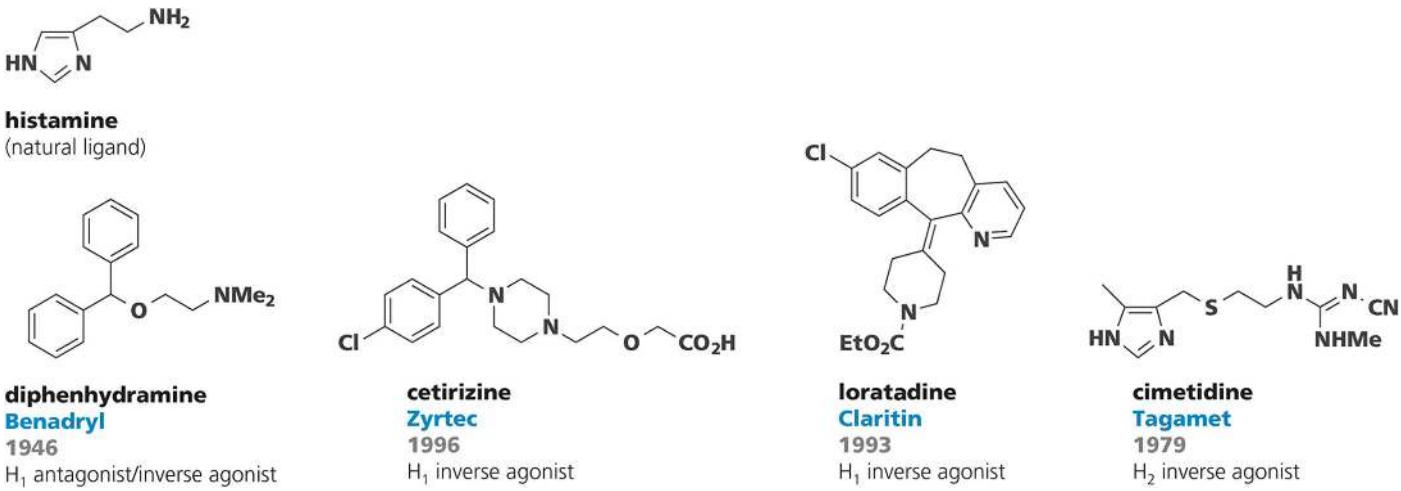


Figure 6.16 Examples of drugs acting at histamine receptors and natural ligand histamine. Diphenhydramine, cetirizine, and loratadine are all antagonists or inverse agonists at the H_1 receptor. Diphenhydramine, cetirizine, and loratadine act at the H_1 subtype and are used to treat allergies, while cimetidine acts at the H_2 subtype and is used to treat gastric ulcers.

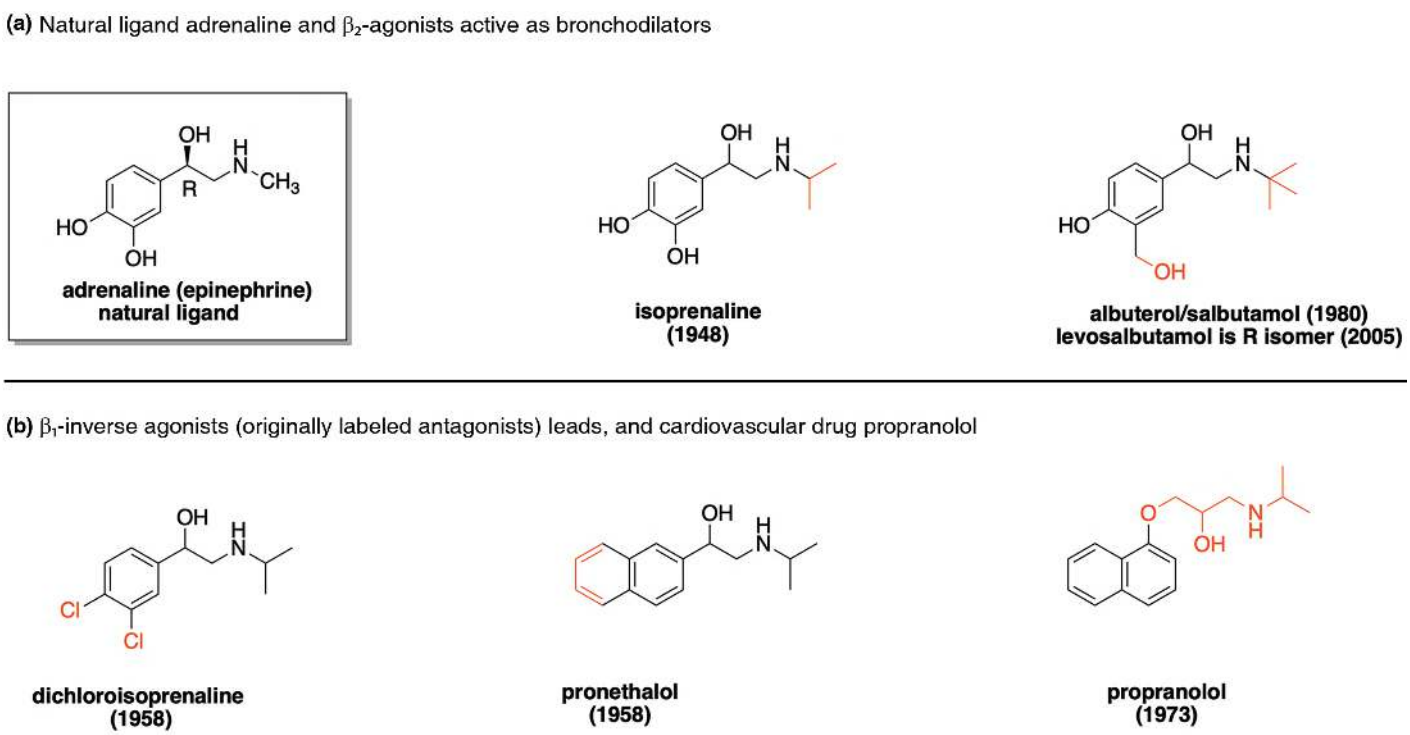
lower levels of sedation. Although long categorized as antagonists, loratadine and diphenhydramine have recently been reclassified as inverse agonists. These compounds are also examples of drugs that were originally available only by prescription, and then, on the basis of their safety record, became available over the counter.

Drugs that are selective for the H_2 subtype act to prevent the effects of histamine on cells in the gut that secrete gastric acid, leading to their use as ulcer medications. One of the first drugs developed using the concept of modern drug design was cimetidine (**Figure 6.16**). Robin Ganellin, Sir James Black, and others at Smith, Kline, & French designed a series of compounds based on the structure of the natural ligand histamine (see ancillary materials for Chapter 3).

Drugs acting at the adrenaline receptor are used to treat a variety of conditions, including asthma, heart failure, arrhythmia, and hypertension. These drugs are categorized by the receptor subtype that they act upon, as well as their agonist or antagonist activity. In 1948, two major subtypes of the adrenergic receptor, α and β, were recognized. These have been delineated further into α_1 and α_2 subtypes and β_1, β_2, and β_3 subtypes. Two of the most clinically useful categories of drugs are β_2 agonists and β_1 inverse agonists (originally classified as antagonists). The stimulation of β_2 receptors in the lung results in bronchodilation, and a number of agonists based on the phenylethanolamine structure of the natural ligand adrenaline have been developed to treat asthma (**Figure 6.17**). Increasing the size of the alkyl substituent on the nitrogen atom leads to efficacy at the β_2 receptors, with varying levels of selectivity over other adrenaline receptors. Isoproterenol (isoprenaline), containing an isopropyl group in place of the methyl group in adrenaline, is an early example of an agonist at the β_2 receptor, and use of it results in bronchodilation in patients. However, the catechol is rapidly metabolized to an inactive metabolite. Homologation of one of the catechol hydroxyl groups to a benzylic alcohol increased the duration of action, and this modification, combined with further incorporation of steric bulk, led to *rac*-albuterol (salbutamol), a drug commonly used to treat asthma. Treatment with the optically pure *R*-isomer levosalbutamol avoids the inflammatory side effects of the *S*-isomer.

Removing the hydroxyl groups of the aromatic ring in adrenaline leads to activity as an inverse agonist rather than agonist activity. Dichloroisoprenaline, which blocks the effects of adrenaline on bronchodilation and cardiac stimulation, as well as on uterine relaxation in animal models, is a representative compound. Replacing the dichloro phenyl scaffold with a naphthalene led to pronethalol, the first β_1-inverse

Figure 6.17 Examples of drugs acting at adrenergic receptors and the natural ligand adrenaline. (a) β_2-receptor agonists isoprenaline and albuterol, both containing the phenylethanolamine moiety found in adrenaline, are effective as bronchodilators (modifications are shown in red). Increasing steric bulk at the nitrogen atom increases potency; extension of alcohol by a carbon atom improves duration of action. Levosalbutamol is the R-enantiomer of albuterol B: β_1-receptor inverse agonist led to dichloroisoprenaline, along with pronethalol and propranolol, which is used as a cardiovascular drug (modifications are shown in red). Propranolol, containing an aryloxypropanolamine scaffold, is the prototypical β-blocker.



agonist marketed. However, it was found to cause thymic tumors and was withdrawn from the market. Further modification, by reorientation and homologation of the side chain through the insertion of an -OCH_2 group, led to the first example of an important class of drugs characterized by the aryloxypropanolamine scaffold, commonly known as β-blockers. Propranolol is a prototypical β-blocker that is widely used to treat both cardiac arrhythmia and hypertension. Sir James Black was awarded the Nobel Prize in Physiology or Medicine in 1988 for his discovery of this drug as well as others.

Drugs acting at GPCRs found in the CNS are numerous and are important analgesics and antipsychotic medications. These include opiates, such as morphine, codeine, and fentanyl, to treat pain, and antipsychotic drugs acting at dopamine receptors, such as Thorazine, haloperidol, and clozapine. All these drugs act at GPCRs and are discussed in detail in Chapter 13.

The design of the anti-migraine serotonin receptor agonists, such as sumatriptan and zolmitriptan (**Figure 6.18a**), was based on the vasoconstriction effects of the natural ligand serotonin and of the natural product ergotamine, which was approved to treat migraine. The seven approved triptans, the natural ligand, and ergotamine all contain

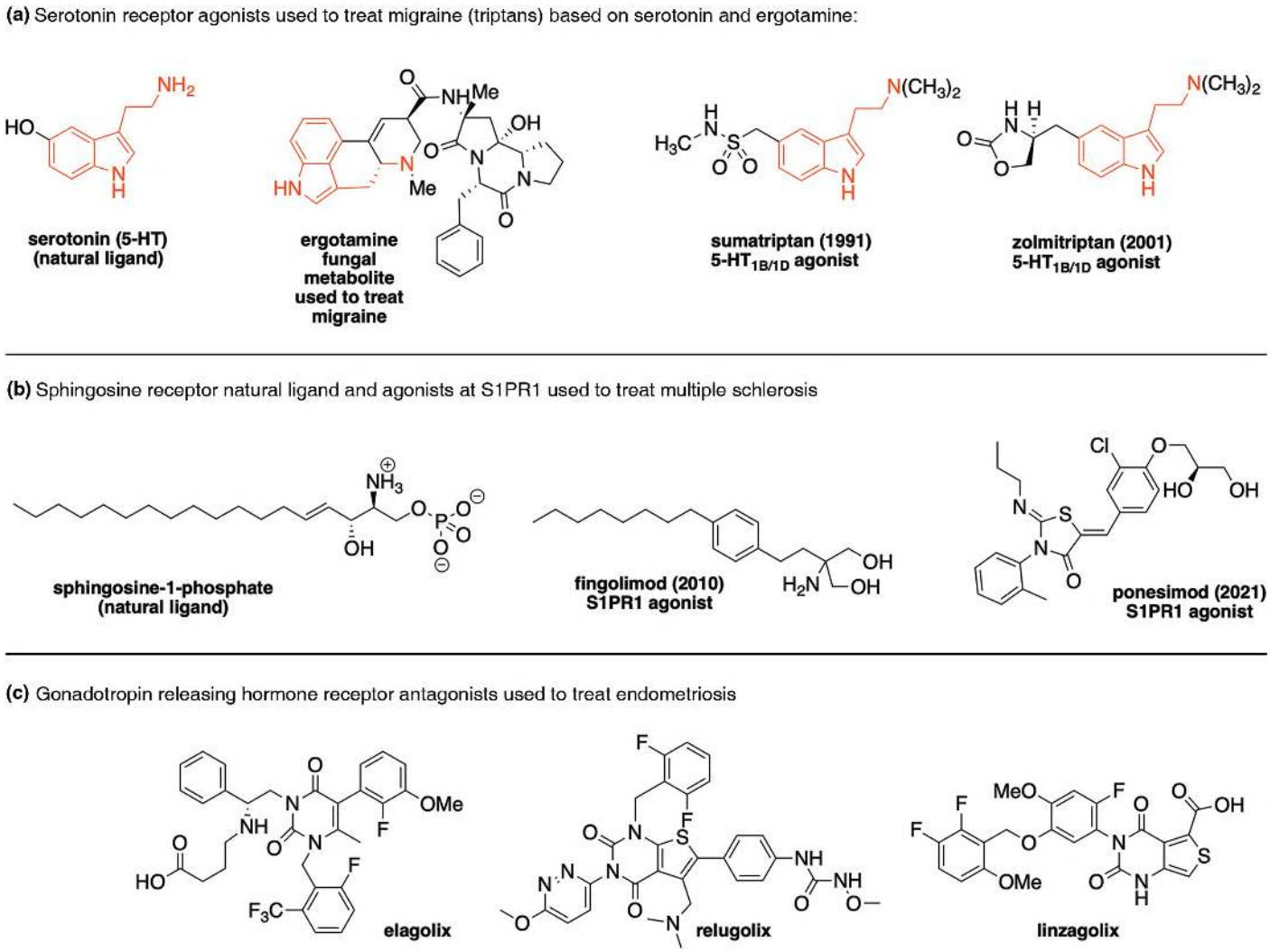


Figure 6.18 Examples of drugs acting at GPCRs with diverse bioactivities.
(a) 5-HT receptor natural ligand, serotonin (5-HT); natural product, ergotamine, which causes vasoconstriction and is used to treat migraine; and two triptans, sumatriptan and zolmitriptan, which are serotonin 5-HT1B/1D agonists (common structural features are shown in red). (b) S1PR1 receptor natural ligand, sphingosine-1-phosphate, and fingolimod and ponesimod, used to treat multiple sclerosis. Phosphorylation of fingolimod generates the phosphate, which is an agonist at the S1P1 receptor; ponesimod is an example of selective S1PR1 agonist that does not require phosphorylation. (c) GnRH-R antagonists elagolix, relugolix, and linzagolix are used in the treatment of endometriosis and/or uterine fibroids, among other indications.

an indole ring and a pendant amine function. These triptans are selective for the 5-$HT_{1B/1D}$ subtypes of the serotonin receptors and constrict blood vessels in the brain.

Fingolimod (**Figure 6.18b**), the first disease-modifying drug approved to treat multiple sclerosis (MS), also acts at a GPCR. It is phosphorylated, and that product, a monophosphate, is an agonist at sphingosine 1-phosphate receptor (S1PR1), a lipid-activated GPCR involved in the immune response, and it is this effect that is thought to contribute to its efficacy. Fingolimod also modulates other S1P receptors, which contributes to some of its side effects. Newer approved compounds include siponimod, ozanimod, and ponesimod. These compounds not only are more selective for the S1P1 receptor but also do not require phosphorylation *in vivo*, as fingolimod does. In addition to MS, ozanimod and ponesimod are also approved to treat additional autoimmune disorders.

One active area of research is the development of gonadotropin-releasing hormone receptor (GnRH-R) antagonists for the treatment of endometriosis, uterine fibroids, and other disorders. The natural ligand for the receptor is the decapeptide gonadotropin-releasing hormone, also known as luteinizing hormone–releasing hormone, and first-generation compounds were peptides based on this natural ligand peptide. More recently, orally available small-molecule antagonists have been approved. The original leads were based on library screening and are not close analogs of the natural peptide ligand (**Figure 6.18c**). The first orally active drug in this class, elagolix, was approved in 2018, and relugolix (2019) and linzagolix (approved in the EU and UK; 2022) were approved shortly thereafter.

6.3 ION CHANNELS (IONOTROPIC RECEPTORS) AS DRUG TARGETS

Ion channels are proteins that span the cell membrane and regulate the highly selective transport of ions from one side of the cell to the other. Their function is critical to physiological processes, such as muscle contraction and neurotransmission. There are hundreds of genes that code for ion channels, and many conditions are related to the malfunction of those channels. Examples include cardiovascular diseases related to malfunction of calcium channels; depression and anxiety, which can be caused by a dysfunction of γ-aminobutyric acid (GABA) chloride ion channels; and pain, which can be affected by sodium channels. Ion channels are made up of multiple subunits that combine to form a cylindrical structure with a central pore. The pore is lined with hydrophilic residues that are arranged in a specific geometry (**Figure 6.19a**). This geometry serves to select, usually by size, which specific ions are transported and which are excluded. Ion selectivity is regulated by an ion filter (for charge and/or size) and a gate that allows the selected ions to pass into or out of the cell. Ion channels are classified as either voltage-gated ion channels (VGIC) or ligand-gated ion channels (LGIC), depending on whether a change in membrane potential or binding of a ligand initiates their opening or closing. LGICs are also referred to as ionotropic receptors. In addition, ion channels can be classified according to the ions they transport.

Under normal physiological conditions, an ion channel exhibits a range of resting (open), activated (open), and inactivated (closed) states over a given period of time (**Figure 6.19b**). The equilibrium between these states is affected by ligands in the case of LGICs and by changes in membrane potential in the case of VGICs.

Natural products have been an excellent source of small molecules that act at ion channels. These structures, sometimes themselves drugs, are often used as starting points for drug discovery and medicinal chemistry activities. In addition, high-throughput screening and ligand-based screening are also common starting points. The first X-ray structure of an ion channel (a bacterial K^+ channel) was solved in 1998, and like the GPCRs, this field is evolving as new structural information becomes available. The recent successful solution of X-ray and cryo-EM structures illustrating ion channels in various points of opening has further supported the use of structure-based design.

Assaying for activity at ion channels is simpler than for GPCRs, since the downstream event is an immediate change in membrane potential due to the release of ions. A widely used assay is the patch clamp assay, developed in the late 1970s by

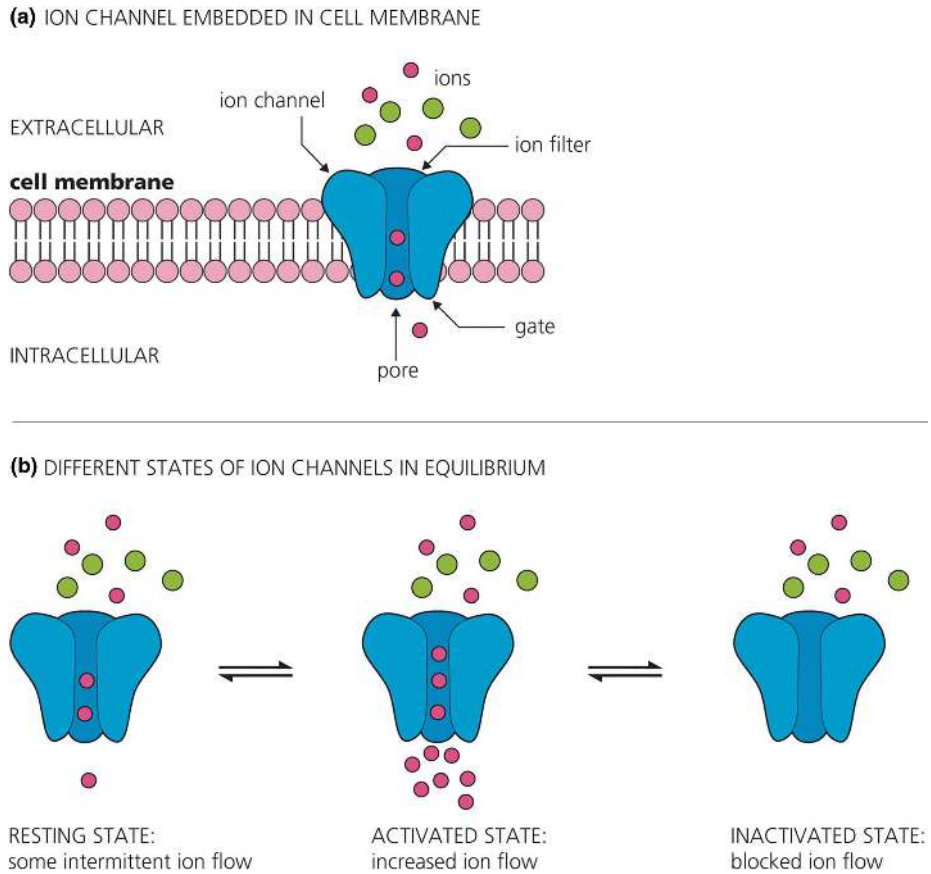


Figure 6.19 Ion channels. (a) General structure of an ion channel (shown in blue) embedded in a cell membrane. Ions of various sizes (shown in green and pink) on the outside of a cell pass through an ion selectivity filter into the hydrophilic pore. Only those with the appropriate size and charge (here shown in pink) proceed through a gate into the cell. (b) Different activation states of ion channels. At resting state, there is some basal level of ion flow; in the activated state, there is increased ion flow; and in the inactivated state, ion flow is blocked or decreased from the basal level.

Erwin Neher and Bert Sakmann, for which they won the Nobel Prize in Physiology or Medicine in 1991. In this assay, a patch of membrane containing as few as one ion channel is used, and the voltage change can be measured with a microelectrode over millisecond time periods, indicating whether the ion channel is open or closed. As shown in Figure 6.20, an *agonist* or *opener* will cause the channel to open for a given

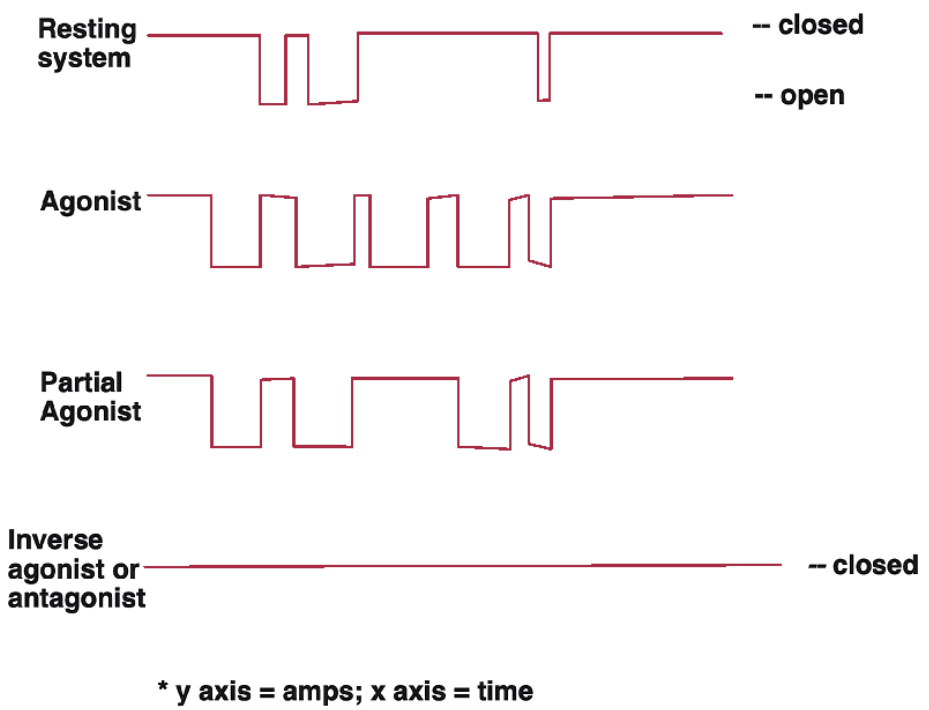


Figure 6.20 Patch clamp assays. Time is plotted vs. conductance (measured in amperes), reflecting an open or closed channel. Top trace shows the changes in conductance at resting state. The effect of an agonist is increased ion flow through the open channel, which may open more often and/or for longer periods of time. The effect of a partial agonist is increased flow, but less than for a full agonist. The effect of an inverse agonist or antagonist (blockers) is decreased or no opening of the channel.

period of time or cause it to stay open longer or to open more frequently than it does in its resting system. A *partial agonist* will open the channel, but for a shorter time period and/or less frequently than a full agonist. *Antagonists* and *inverse agonists* (also called channel blockers) prevent the channel from opening. Patch clamp assays have been modified and automated for use in HTS, as have other assay techniques. These include the use of flash luminescence assays for calcium channels, which rely on Ca^{2+}-sensitive photoproteins that emit a burst of light upon calcium binding, as well as fluorescence-based assays that respond to changes in ion concentration.

6.3.1 Ligand-Gated Ion Channels Open or Close in Response to Binding of a Drug or Natural Ligand

Ligand-gated ion channels all have a cylindrical shape composed of protein subunits with a hydrophilic internal surface that, when open, allows the passage of ions through the pore. When closed, the pore is locked until an agonist or natural ligand binds. Binding causes a conformation change that opens the gate and allows the ions to pass through. In addition to being classified according to the type of ion allowed to pass through, LGICs are also characterized by the ligand or agonist that binds and according to the arrangement of their transmembrane regions. LGICs may open or close in response to neurotransmitters or other small molecules or to the presence of calcium ions or cAMP as second messengers.

The superfamilies of LGICs include the pentameric Cys-loop family, the tetrameric ionotropic glutamate receptors (iGluR), and the trimeric purinergic ATP-gated channels (P2X). **Figure 6.21** highlights the general structural differences between the families of LGICs, based on the number and arrangement of subunits. A top view of each superfamily is shown. The regions colored red represent hydrophilic portions that align on the inner core of the Cys-loop and ionotropic glutamate channels, while blue indicates hydrophobic regions that align with the cell membrane.



Figure 6.21 General schematic of families of ligand-gated ion channels. Top views are shown. Hydrophobic (blue) and hydrophilic (red) portions in Cys-loop and ionotropic glutamate channels are grouped in the transmembrane-spanning regions that make up each subunit (yellow circles); P2X channels contain hydrophobic patches. Although not designated in this general view, each subunit may consist of separate units, called α, β, γ, or δ. The Cys-loop channel consists of five subunits, the ionotropic glutamate ion channel has four subunits, and the P2X ion channel consists of three subunits.

The Cys-loop receptors are characterized by five subunits, usually two α and three among β, γ, and δ, which have different sequence homology. Near the N-terminal extracellular region, there is a disulfide linkage between two cysteine residues that forms a loop. The ionotropic glutamate receptors have four subunits, and the P2X receptors have three. **Table 6.2** illustrates examples of LGICs and the ions they transport, their natural ligands, the drugs that target each, and the approved indication.

One of the most well-studied LGICs is the nicotinic acetylcholine receptor (nAChR). In 1903, John Langley proposed the existence of specific receptive substances for nicotine and curare, which we know today as the nAChR. He noted that curare, a muscle relaxant that is the ingredient in a South American blow dart poison, antagonized the action of nicotine, a muscle contractant. The N2 subtype of the nAChR (N2 nAChR), also known as the neuronal receptor subtype, is a pentameric Cys-loop channel composed of two α_2 and one each β, γ, and δ subunits that allows passage of Na^+ ions as well as K^+ and Ca^{2+} ions. The N2 subtype receptors are located on the surface of muscle cells, near the synapse between neurons and muscle cells. In the normal course of action, two molecules of the natural ligand acetylcholine bind to each of the α subunits, resulting in channel opening and the passage of Na^+ ions. This opening causes the depolarization of the postsynaptic membrane. As the potential spreads along the muscle, the muscle contracts. **Figure 6.22** shows a diagram of the N2 nAChR based on an X-ray structure.

Table 6.2 Examples of Ligand-Gated Ion Channel Receptors, Their Ligands, the Ions Transported, the Drugs Targeting It, and the Approved Indication

Natural Ligand	Receptor	Ion(s) Transported	Drug or Ligand Examples	Indication
Cys-Loop Receptors				
Acetylcholine	Nicotinic AChR (neuromuscular)	Na^+, K^+	Vecuronium, pancuronium	Paralytics
	Nicotinic AChR (neuronal)	Na^+, K^+	Nicotine, varenicline	Partial agonist, smoking cessation
Glycine	GlyR	Cl^-	Strychnine[a], brucine[a]	
γ-aminobutyric acid	$GABA_A$	Cl^-	Benzodiazepines	Sedative
			Barbiturates, ethanol, inhaled anesthetics	Muscle relaxants, sedatives
			Zuranolone	Postpartum depression
Serotonin (5-HT)	$5\text{-}HT_3$	Na^+, K^+	Ondansetron	Nausea
Ionotropic Glutamate Receptors				
α-amino-3-hydroxy-5-methyl-4-isoxazole-propionic acid	AMPA	Na^+, K^+, Ca^{2+}	Perampanel, quisqualate[a], 5-methyl-4-aniracetam[a]	Anticonvulsant
N-methyl-D-aspartate	NMDA	Na^+, K^+	Ketamine	Anesthetic
			Phencyclidine (PCP)[a]	
Kainate, glutamate	KAR, $GluR_{5-7}$	Na^+, K^+	Tezampanel[a]	
P2X Receptors				
ATP	$P2X_{1-7}$	Na^+, Ca^{2+}	Suramin[a]	

[a]Ligand is not used clinically and either is a probe molecule or is in development or has been removed from the market.

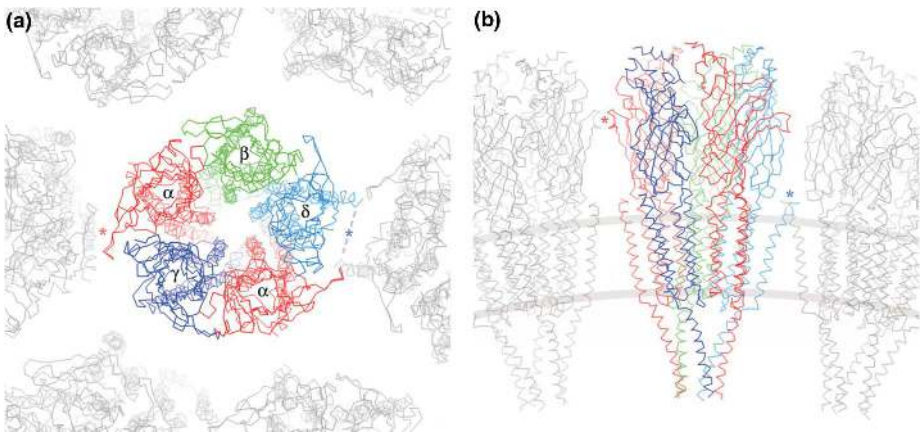


Figure 6.22 Structure of the N2 nicotinic acetylcholine receptor based on X-ray. Views from the top (*left*) and the side (*right*) relative to the membrane are shown. Subunits α, β, γ, and δ are colored and labeled in the left image.

Source: From Unwin, N. (2005). *J Mol Biol* 346:967–989. With permission from Elsevier.

6.3.2 Neuromuscular Blockers, Sedatives, and Anti-Nausea Agents Are Examples of Drugs Acting at LGICs

Drugs that act as antagonists at all subtypes of the nAChR are used as neuromuscular blocking agents in anesthesia. One of the simplest examples is succinylcholine, which can be thought of as two units of the natural ligand (acetylcholine)

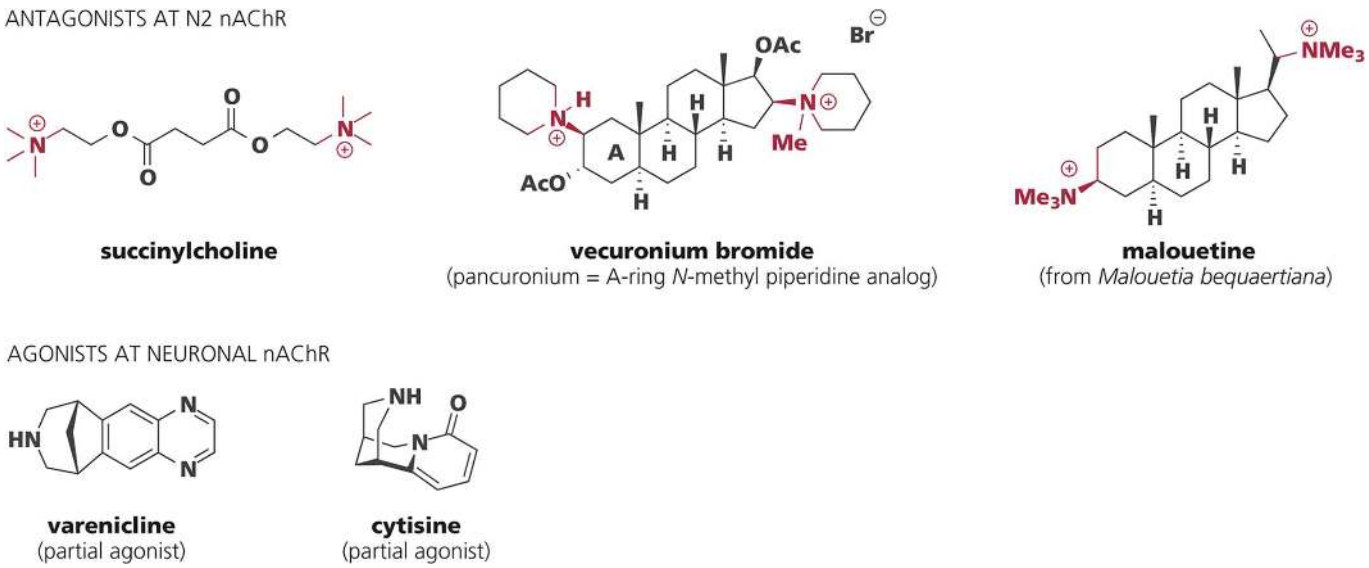


Figure 6.23 Examples of drugs acting at various nicotinic acetylcholine receptors. Antagonists such as succinylcholine, vecuronium, based on malouetine, are used as a neuromuscular blocker (common quaternary amines are shown in red). The partial agonist varenicline is used for smoking cessation and was based on the natural product cytisine.

linked together. Pancuronium and vecuronium are steroids based on the natural product malouetine (Figure 6.23). One of the natural products first found to act on nAChRs was tubocurarine, described in Chapter 1. The effects of these compounds can be profound, with pancuronium being one of several drugs used for lethal injections. The common structural feature of these compounds is the presence of two quaternary amines positioned to interact with two separate α subunits of the ion channel.

The partial agonist varenicline was developed at Pfizer and is used to treat nicotine addiction. Dependence on nicotine is related to its activation of an $\alpha_4\beta_2$ subtype of the ion channel, which results in increased dopamine levels. Varenicline is selective for neuronal $\alpha_4\beta_2$ ion channels. Its structure is based on cytisine, a natural product found in a number of plants, including *Cytisus laburnum*, the leaves of which have been used as a tobacco substitute and as a natural smoking cessation aid.

Benzodiazepines, such as diazepam, act at the Cys-loop $GABA_A$ chloride ion channel and bind to an allosteric binding site, now termed the benzodiazepine (BZ) site, located between the α and γ subunits. The natural ligand GABA binds between the α and β subunits and is an inhibitory neurotransmitter—meaning, it decreases the likelihood that the postsynaptic neuron will generate an action potential. Other allosteric agonists (PAMs) include barbiturate sedatives, ethanol, and inhaled anesthetics, such as isoflurane. Zuranolone, also a PAM, is a neurosteroid that acts at a distinct allosteric site and modulates both synaptic and extrasynaptic $GABA_A$ conductance.

The benzodiazepines are an example of a family of drugs where small structural modifications lead to significant differences in intrinsic activity. Early benzodiazepines, such as diazepam, act as positive allosteric modulators and amplify the effects of GABA, resulting in the opening of the channel. These benzodiazepines are widely used as anti-anxiety drugs and as anticonvulsants, but they also impair memory. Minor changes in the structure lead to differences in efficacy, although they all bind at the benzodiazepine site. For example, while diazepam is a full PAM, the analog bretazenil is a partial PAM (Figure 6.24a) that maintains anti-anxiety activity, although it is not marketed. The structurally similar flumazenil is a negative allosteric modulator (NAM) and is used to treat benzodiazepine overdose. The NAM Ro15-4513, an experimental agent, is identical to flumazenil except for the replacement of the fluorine atom with an azide group. While not as effective as flumazenil against benzodiazepine overdose, Ro15-4513 was effective as an antidote to alcohol poisoning in animal models, although it also caused seizures. One physiological

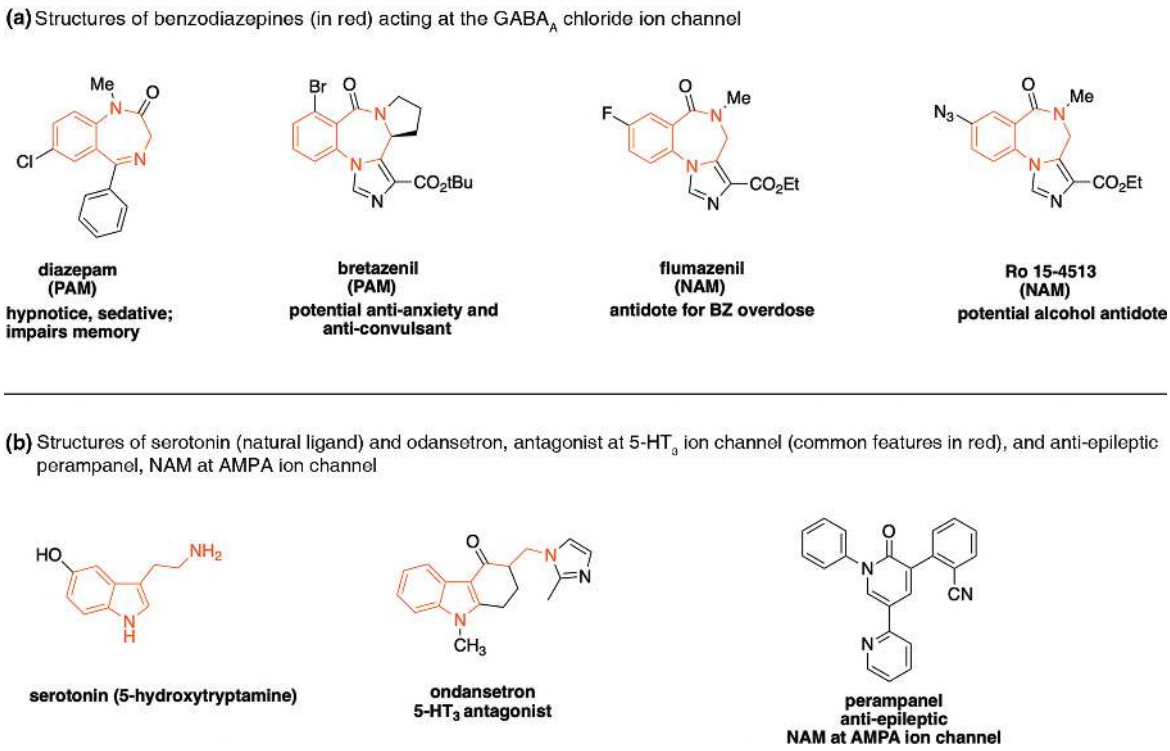


Figure 6.24 Examples of drugs acting at GABA$_A$, 5-HT$_3$, and AMPA ion channels. (a) Benzodiazepines acting at the GABA$_A$ chloride ion channel, with common benzodiazepine structure shown in red. PAMs, such as diazepam and bretazenil, are anti-anxiety drugs and can be anticonvulsants. The NAM flumazenil is an antidote for benzodiazepine overdose, and the experimental NAM Ro15-4513 was a potential alcohol antidote. (b) Structures of serotonin (natural ligand) and ondansetron, an antagonist at the 5-HT$_3$ ion channel, used to treat nausea (common features in red), and perampanel, a NAM at the AMPA ion channel, used to treat epilepsy.

effect of NAMs is an enhancement in memory; however, this is offset by anxiogenic and convulsant properties.

Although the other subtypes of serotonin receptors are GPCRs, the 5-HT$_3$ receptor is a sodium/potassium ion channel found in several regions of the nervous system that responds to 5-hydroxytryptamine (serotonin). Antagonists such as ondansetron, structurally similar to the natural ligand serotonin, are used primarily to treat nausea associated with chemotherapy in cancer patients (Figure 6.24b). The AMPA ion channel is involved in glutamate-mediated excitatory neurotransmission and is a target for anticonvulsant medications. Scientists at Eisai discovered the negative allosteric modulator perampanel through medicinal chemistry optimization of a high-throughput screening hit identified in an assay for compounds that protected cells from AMPA-induced neuronal cell death. Approved in 2012, perampanel is used to treat epilepsy.

6.3.3 Voltage-Gated Ion Channels Open or Close in Response to Changes in Membrane Potential

Voltage-gated ion channels (VGICs) open or close in response to a change in membrane potential and may be selective for Na^+, K^+, Ca^{2+}, or Cl^- ions. Like LGICs, they exist in equilibrium between three states: resting, open, and inactivated. While the states vary depending on membrane potential, drugs that act at these ion channels have affinity for one form over another and hence affect their equilibrium. Voltage-gated channels typically contain four transmembrane domains with a selectivity filter at the top surrounded by four voltage sensors. The voltage-sensing domain responds to membrane depolarization, causing the opening of the channel. Several models for the function of VGICs have been proposed, including a piston-type model, a flap model, and a transporter-type model (Figure 6.25).

In the piston model, the voltage-sensing domains move up or down, as would a piston, resulting in a conformational change that allows ions through the pore. On the basis of an X-ray structure of a bacterial K^+ voltage-gated channel, this model has been disfavored, and ones where the voltage-sensing domains are more like paddles

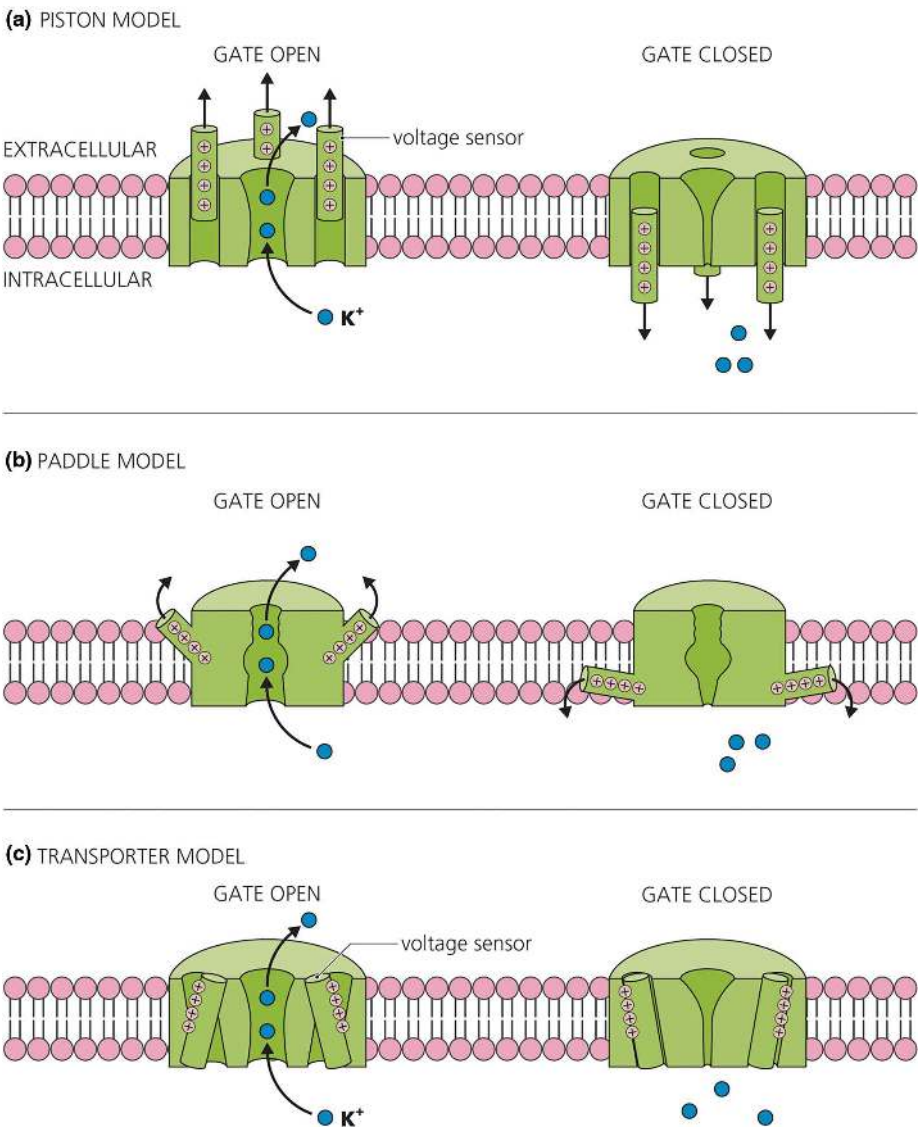


Figure 6.25 Several models proposed for voltage-gated ion channel opening and closing. (a) Piston model, showing the movement of charged pistons (cylinders) as ion concentration changes. (b) Paddle model, showing charged paddles that move as levers to open and close the ion channel. (c) Transporter model, where charged subunits do not move through the membrane but translocate within the receptor to affect the opening and the closing of the channel (membrane in pink; VGIC in green, with charges on pistons in pink, ions in blue).

Source: Adapted from Ashcroft, F.; Benos, D.; Bezanilla, F.; et al. (2004). *Nat Rev Drug Discovery* 3:239–278. With permission from Macmillan Publishers Ltd.

that either move through the membrane, as in the paddle model, or simply pivot along an axis, as in the transporter model, are now preferred. As more and more structural information on VGICs becomes available, it is likely that more accurate models of their mechanism will result.

Unlike GPCRs and LGICs, VGICs do not respond to binding by a natural ligand for their activation (or opening and closing); therefore, hit identification cannot rely on knowledge of the structure of a natural ligand. As such, drug discovery targeting VGICs has not been as active an area as it has been for GPCRs and LGICs. Most drugs in this category were discovered serendipitously, without knowledge of their mechanism of action. Many plant and animal toxins have activity at VGICs, including tetrodotoxin from pufferfish and saxitoxin from shellfish, as well as batrachotoxin, found in frogs. More recent examples of drugs have been based on pharmacophore models developed from such discovered ligands. Compounds that act to open the channel are known as **channel activators**, and those that keep the channel closed are termed **channel blockers**. Significant advances in miniaturization and standardization of techniques to measure ion flux have been developed, and these techniques now allow high-throughput screening as a viable approach to hit identification and optimization. Like other receptors, selectivity for one LGIC in the presence of many others that transport the same ion is often a significant hurdle. Some therapeutically important VGICs are listed in Table 6.3.

Table 6.3 Examples of Voltage-Gated Ion Channel Receptors, the Ions Transported, the Ligands or Drugs Targeting It, Their Mechanism of Action, and Their Approved Indication

Channel	Ion(s) Transported	Drug or Ligand Examples	Mechanism	Indication
Na_v	Na^+, K^+	Lidocaine Carbamazepine, phenytoin	Blocker Blocker	Anesthetic Anticonvulsants
K_v7	K^+	Retigabine, flupirtine	Activator	Anticonvulsant, analgesic
Ca_v ($\alpha_2\delta$)	Ca^{2+}	Gabapentin, pregabalin	Activator	Anticonvulsant, analgesic
L-type Ca_v	Ca^{2+}	Verapamil, diltiazem, nifedipine	Blocker	Cardiovascular disease
$Ca_v2.2$	Ca^{2+}	Ziconotide	Blocker	Analgesic
influenza M2	H^+	Amantadine	Blocker	Influenza

6.3.4 Blockers of Sodium, Calcium, and Potassium VGICs Include Local Anesthetics, Analgesics, Anticonvulsants, and Anti-Arrhythmia Drugs

Drugs typically have affinity for either the activated (open) or inactivated (closed) state of the VGIC, binding either directly to the ion channel subunits or to auxiliary subunits. Historically, important drugs that act through voltage-gated sodium channels include the "caine" anesthetics. Although illegal today, the natural product cocaine was used as a local anesthetic in the late 1880s, and that bioactivity was later attributed to action at a voltage-gated sodium channel. However, cocaine has multiple effects, and today a number of much more selective caine anesthetics are available and are used as local anesthetics. The sodium channel involved initiates and propagates the action potentials in neurons. These molecules bind to the open and inactivated states with high affinity. Early analogs of cocaine included amylocaine (Stovaine) and the more commonly used procaine (Novocaine), which was introduced in 1905. Problems with procaine include its short duration of action and instability. The search for more stable analogs led to the amides lidocaine (Xylocaine) and prilocaine (Figure 6.26). Lidocaine was originally derived from a synthetic intermediate prepared during a synthesis of isogramine, a natural product found in a barley plant that, like cocaine, has activity as a local anesthetic. Isogramine itself is too irritating to be a useful anesthetic, and the chemists Nils Löfgren and Holger Erdtman prepared many analogs of isogramine, as well as its synthetic precursor, before identifying lidocaine.

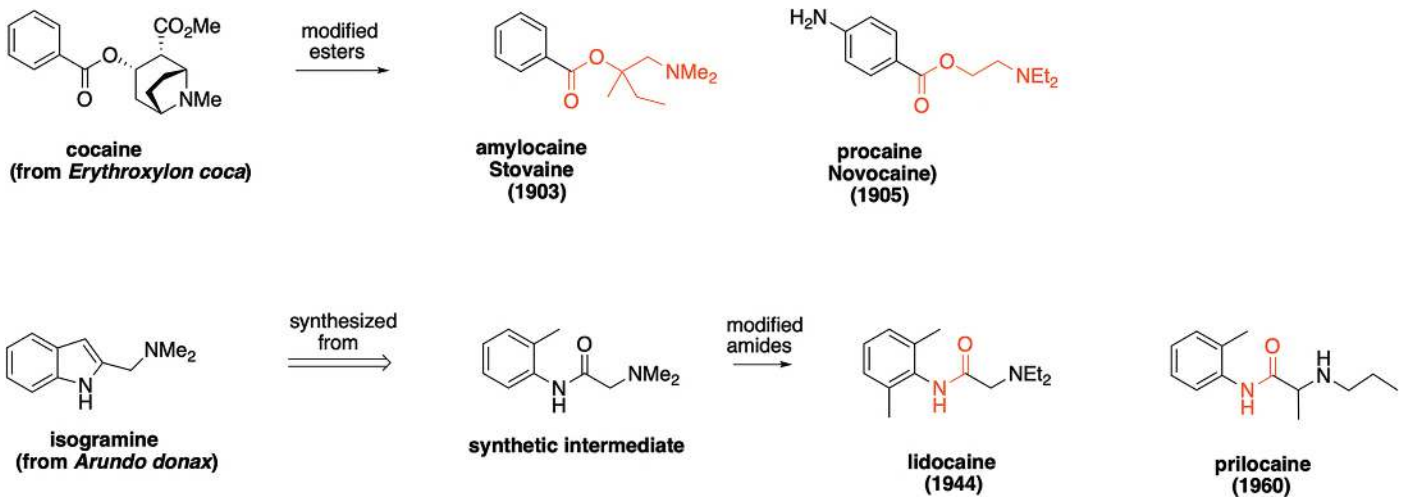


Figure 6.26 Examples of caine anesthetics. *Top*, cocaine was the basis of analogs amylocaine and procaine, commonly used anesthetics (modifications shown in red). *Bottom*, lidocaine and procaine (amide groups shown in red) were derived from a synthetic intermediate used to make the natural product isogramine. Both are frequently used as local anesthetics.

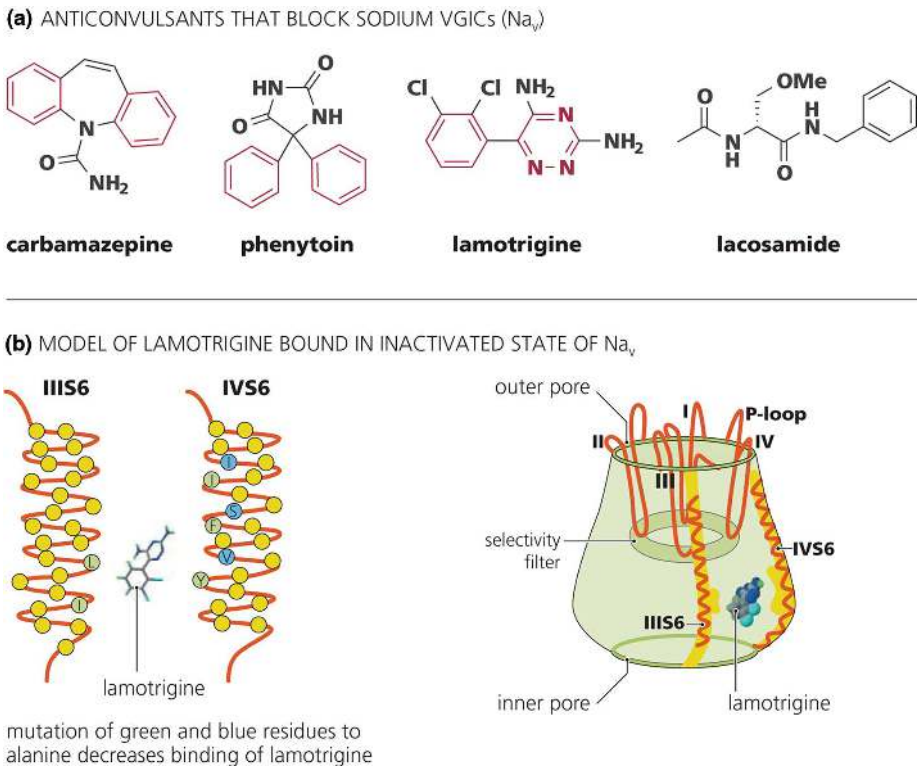


Figure 6.27 Anticonvulsant drugs acting at voltage-gated ion channels. (a) Carbamazepine, phenytoin, lamotrigine, and lacosamide all block sodium VGICs. Common di-aryl structural features of the first three are shown in red. (b) Model of lamotrigine localized at its binding site in the brain Na_v. *Left*, close-up view of binding site: IIIS6 and IVS6 are subunits in the pore region, and lamotrigine binds between them. Amino acids mutated (green and blue) to identify binding site. *Right*, model of the Na_v pore: I–IV (red lines) each represents an α-subunit, each of which has six segments (S1–S6); the wall of the ion channel is formed by S5 and S6 α-helices, with the IIIS6 and IVS6 that contain the binding site highlighted. The outer pore and selectivity filter are formed by P loops that connect the subunits.

Source: Adapted from Rogawski, M. A., and Löscher, W. (2004). *Nat Rev Neurosci* 5:553–564. With permission from Macmillan Publishers Ltd.

A number of drugs that are used to treat convulsions act at the Na_v or sodium VGICs. Convulsions can occur when neurons are depolarized and active for long periods of time. Most anticonvulsant drugs were discovered through phenotypic assays rather than by design, and those that act at sodium VGICs include diverse structures, such as carbamazepine, phenytoin, lamotrigine, and most recently, lacosamide (**Figure 6.27a**). Their overall mechanism is stabilization of the inactivated state of the receptor, preventing channel opening and ending the action potential that leads to convulsions. Carbamazepine was discovered in the 1950s and first marketed in the 1960s as an anticonvulsant. While it stabilizes the inactivated state of the ion channel, it is not selective and also acts as a GABA agonist. Phenytoin, approved in 1953, was discovered in the early 1900s as a non-sedating analog of the barbiturates. It also binds to the inactivated state of the ion channel. Lamotrigine, approved in 1994 as an anticonvulsant, acts by blocking the inactivated form of Na_v. While these three compounds are from distinct scaffolds, they bind at a common site and contain two pendant aryl rings as a common structural feature. On the basis of site-directed mutagenesis studies of the brain type II Na_v ion channel, lamotrigine does not physically block the pore but is proposed to bind to hydrophobic amino acids at two particular α-helices in the pore region of the ion channel (**Figure 6.27b**). Structurally, this channel is composed of four homologous α-subunits (I–IV), each composed of six α-helical segments (S1–S6).

Lacosamide is an anticonvulsant that acts at a voltage-gated sodium channel. It is also approved to treat neuropathic pain. Although its discovery was based on activity in *in vivo* assays, it has been found to stabilize the inactivated form of the ion channel.

A number of drugs to treat cardiovascular conditions, such as high blood pressure, arrhythmia, angina, coronary artery disease, and heart failure, act at ion channels. These include drugs such as nifedipine, diltiazem, and verapamil, which act as blockers at the L-type potential-dependent calcium channels in the heart. Nifedipine, the first of the 1,4-dihydropyridine structural class approved, has been used since the 1970s (**Figure 6.28**), with others in this family, such as amlodipine, approved in the late 1980s. Despite a common mechanism of action as blockers, these structurally diverse compounds (1,4-dihydropyridines, benzothiazepine, and phenylbutyl amine) all bind at different sites on the calcium channel. One cardiac channel that is particularly important to avoid interacting with is the hERG potassium channel (see

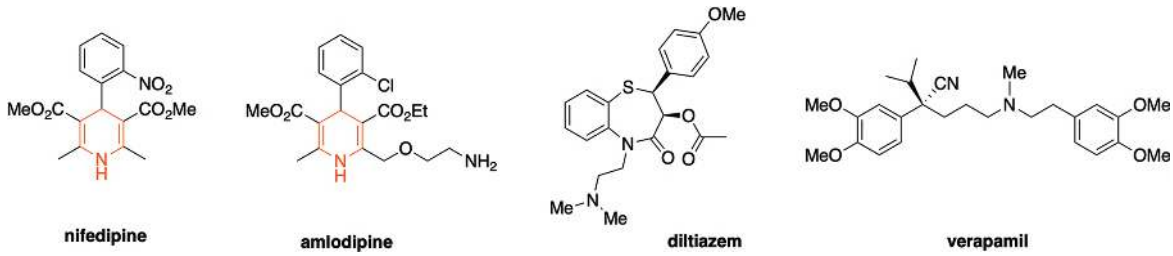


Figure 6.28 Cardiovascular drugs that target L-type Ca^{2+} channels. Examples are structurally diverse and include 1,4-dihydropyridines (in red) nifedipine and amlodipine, as well as diltiazem (a benzothiazepine), and verapamil (a phenylbutylamine). Each is an L-type calcium channel blocker, but they bind at distinct sites.

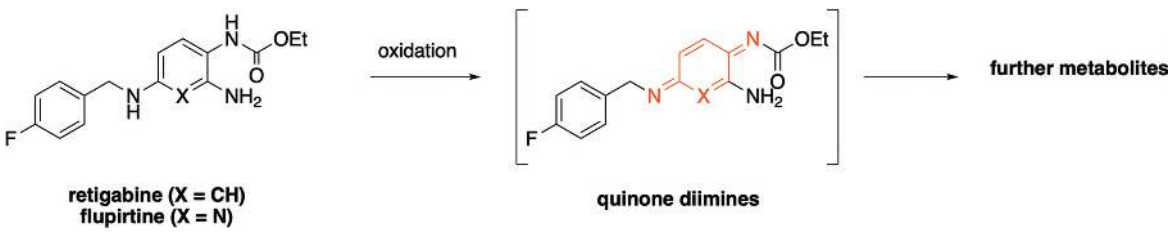


Figure 6.29 Structures of neuronal Kv7 channel activators. Retigabine and flupirtine are oxidized to quinone diimines, which then go on to further toxic metabolites.

Chapter 5). This channel affects the electrical current in the heart, and inhibition leads to a potentially fatal condition called long QT syndrome.

One area of active research is the identification of activators of the neuronal potassium channels K_v7. The close analogs flupirtine and retigabine were approved in the early 2010s (**Figure 6.29**) for the treatment of epilepsy, as well as neuropathic pain. However, both have been withdrawn due to the formation of reactive metabolites, leading to side effects such as skin discoloration and retinal pigmentation. Continuing research is focused on developing more metabolically stable analogs.

6.4 NUCLEAR RECEPTORS AS DRUG TARGETS

Nuclear receptors, like GPCRs, are proteins that transduce a signal based on the binding of a ligand. Unlike GPCRs and ion channels, however, they do not span the cell membrane but instead are intracellular receptors. After binding of a ligand, they function as transcription factors that regulate gene expression. The processes affected include reproduction, energy homeostasis, immunity, circadian rhythm, and metabolic processes, among many others. The natural ligands of nuclear receptors are often lipophilic small molecules, such as estradiol, testosterone, calcitriol (metabolite of vitamin D), and retinoic acid (vitamin A). However, there are also a number of orphan nuclear receptors, for which the natural ligand is unknown. Nuclear receptors and their signaling mechanisms are highly complex. The same receptors may be located in multiple tissues but may have very distinct responses to the same ligand binding. Their role in a variety of physiological and pathophysiological processes, as well as the small-molecule nature of their ligands, makes them a popular target for drug discovery. In fact, approximately 16% of drugs target nuclear receptors.

Structurally, the nuclear receptors share a common N-terminus, which is composed of an activation function (AF-1), a highly conserved DNA-binding domain (DBD), a linking domain, and a ligand-binding domain (LBD), which is part of the second activation function (AF-2) (**Figure 6.30a**). The DBD contains zinc finger motifs (typically two) that bind to Zn atoms and to specific sequences of DNA. In most cases, after binding of the ligand at the LBD, the receptor undergoes a conformational change, which facilitates dimerization with other proteins (termed *co-regulator proteins*) and binding to a specific DNA sequence called the DNA response element, leading to subsequent changes in gene expression. **Figure 6.30b** is an X-ray structure of a truncated form of the nuclear receptor peroxisome proliferator–activated receptor

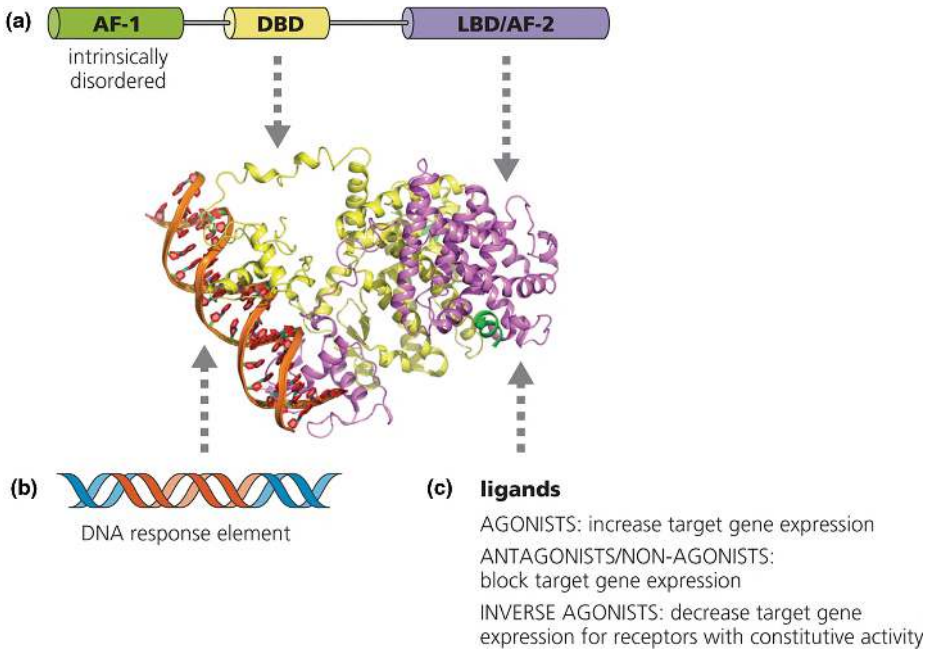


Figure 6.30 Structure of nuclear receptors and actions of ligands. X-ray structure of a complex of a truncated nuclear receptor PPARγ (yellow) bound to its dimerization partner RXRα (purple) and to DNA (orange: DNA response element) and sites for ligands and regulator peptides (green). The AF-1 region is intrinsically disordered and, therefore, not visible in X-ray. (a) Schematic of nuclear receptor: N-terminal domain with activation function 1 (AF-1) linked to the DNA-binding domain (DBD) and the ligand-binding domain (LBD) and activation function 2 (AF-2). (b) DNA response element, which controls gene expression. (c) Ligand types and results: Agonists increase target gene expression, antagonists block gene expression, and inverse agonists decrease gene expression.

Source: Adapted from Kojetin, D. J., and Burris, T. P. (2013). *Mol Pharmacol* 83:1–8. With permission from the American Society for Pharmacology and Experimental Therapeutics.

(PPARγ) dimerized to the retinoid X receptor (RXR) and bound to a fragment of DNA. The natural ligand or an agonist results in an increase in target gene expression; antagonists block gene expression in the presence of the natural ligand, and inverse agonists decrease target gene expression below any basal level.

Nuclear receptors are characterized as class 1 (steroid receptors), class 2, or class 3. Table 6.4 gives examples of common classes 1 and 2 nuclear receptors that are drug targets, their ligands, the drugs that target each, and the therapeutic indication. **Steroid receptors** include the estrogen receptor (ER), androgen receptor (AR),

Table 6.4 Examples of Nuclear Receptors, Their Ligands, the Drugs Targeting It, and the Approved Indication

Natural Ligand	Receptor	Drug or Ligand[a] Examples	Indication
Steroid Receptors (Class 1)			
Estrogen	Estrogen (ER)	Tamoxifen, raloxifene	Breast cancer, osteoporosis
Testosterone, dihydrotestosterone	Androgen (AR)	Flutamide, bicalutamide	Anti-androgens, prostate cancer, hirsutism
Progesterone	Progesterone (PR)	Mifepristone	Abortifacient
Cortisol	Glucocorticoid (GR)	Dexamethasone	Anti-inflammatory
Aldosterone	Mineralocorticoid (MR)	Fludrocortisone	Salt imbalance
Class 2			
Thyroxine	Thyroid	Thyroxine[a]	Hypothyroidism
Vitamin D	Calcitriol, or vitamin D (VDR)	Calcitriol	Osteoporosis
Retinoic acid	Retinoic acid	Isotretinoin, tretinoin	Acne
Fatty acids	Peroxisome proliferator–activated receptors (PPAR)	Rosiglitazone, pioglitazone	Diabetes

[a]Natural ligand, also used as a drug.

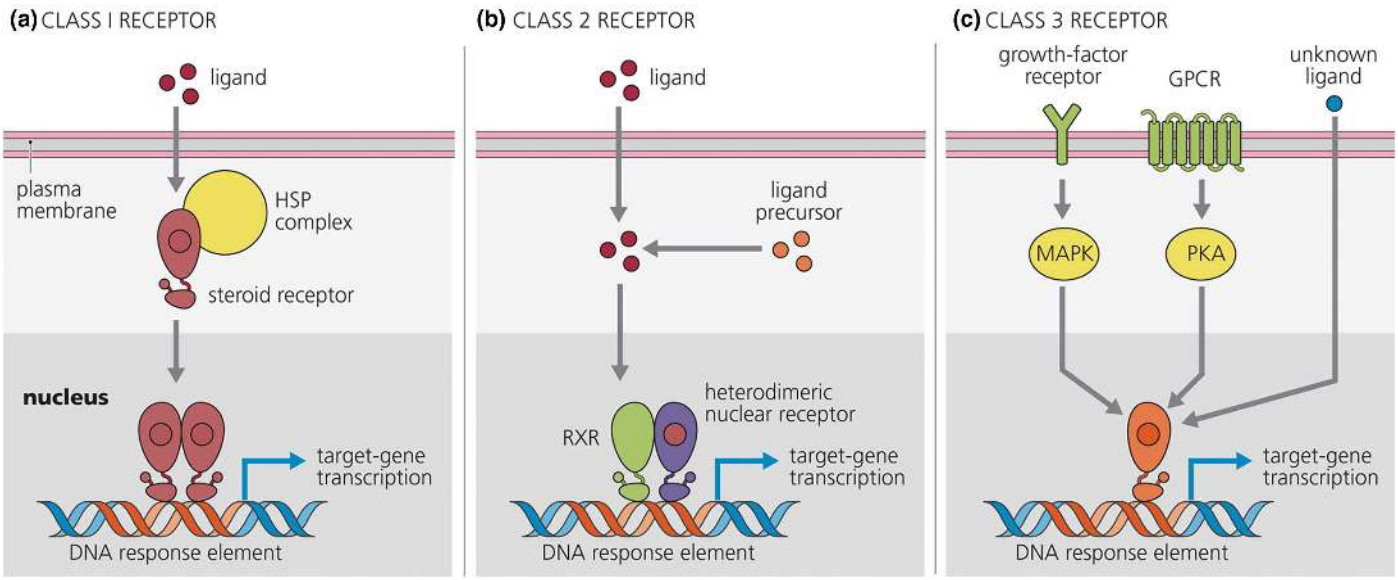


Figure 6.31 Mechanisms of nuclear receptors. (a) Steroid receptor (class 1) mechanism: Ligand binding causes the release of HSP from the receptor–HSP complex and is followed by the translocation of the nuclear receptor into the nucleus and binding as a homodimer to DNA response element, resulting in gene transcription. (b) Class 2 receptor mechanism: The diffusion of the ligand into the nucleus and binding to a pre-associated receptor–RXR heterodimer causes the complex to bind to the DNA response element and results in gene transcription. The ligand may be extracellular or may be produced from a precursor in the cell. (c) Orphan receptor mechanism: Receptors mainly act as monomers (shown), where ligand binding causes binding to DNA response element, resulting in gene transcription, which may be the result of phosphorylation by kinases, such as mitogen-activated protein kinase (MAPK) or protein kinase A (PKA).

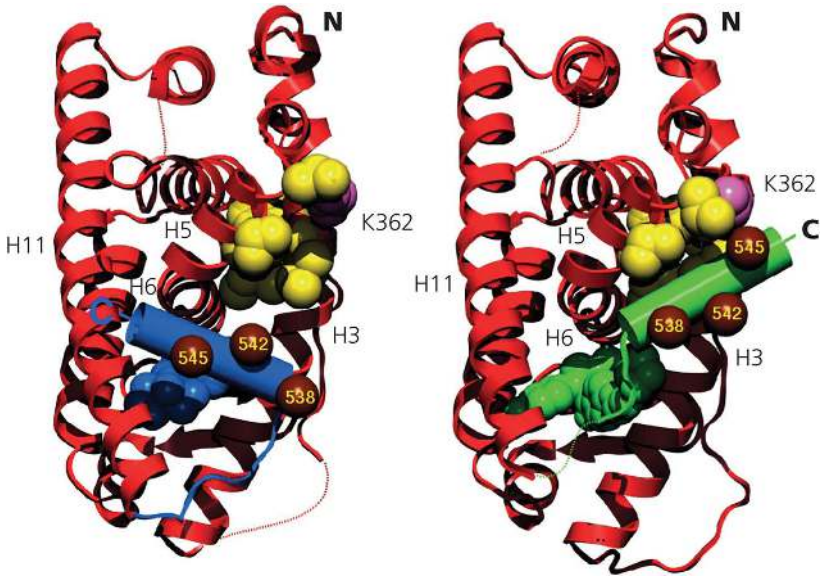
Source: Adapted from Glass, C. K., and Ogawa, S. (2006). *Nat Rev Immunol* 6:44–55. With permission from Macmillan Publishers Ltd.

progesterone receptor (PR), glucocorticoid receptor (GR), and mineralocorticoid receptor (MR), all of which are targets of important drugs. Class 2 receptors include thyroid, vitamin D, retinoic acid, and peroxisome proliferator–activated receptors, also targets of a number of drugs. Class 3 comprises orphan receptors, which have been identified by sequence similarity to known nuclear receptors but, as yet, have no identified endogenous ligands.

Although the nuclear receptors are structurally related, the mechanisms of each class differ. Steroid receptors (class 1) are bound to heat shock proteins (HSPs) in the cytoplasm. When their ligand diffuses into the cell, it binds to the receptor, which then releases the HSP, translocates into the nucleus, and binds to DNA hormone response elements (HREs) near the target genes. DNA binding results in transcriptional activation. Class 2 receptors are located in the nucleus, bound to DNA as heterodimers with the retinoic X receptor (RXR). Once inside the cell, the ligand binds to the LBD and typically recruits additional co-regulatory proteins prior to gene transcription. Orphan receptors (class 3) are least understood and primarily function as monomers. This class may cause gene transcription as a result of phosphorylation by kinases, such as MAPK or PKA (**Figure 6.31**).

As with GPCRs and LGICs, the activity of nuclear receptors can be modulated by small-molecule agonists, partial agonists, antagonists, and inverse agonists. Methods for measuring drug–nuclear receptor affinity and efficacy include detection of direct binding to the nuclear receptor by use of biophysical methods and reporter assays that measure levels of gene transcription, respectively. Hit identification methods include those based on the structure of the natural ligand, high-throughput screening, and structure-based drug design. The observation that some modulators of nuclear receptors have distinct effects depending on where the receptor is located complicates certain aspects of drug discovery programs aimed at these targets. The existence of multiple subtypes of specific receptors that respond to the same natural ligand opens up the possibility of designing selective drugs that modulate some of the effects of the natural ligand but not all.

The X-ray crystal structures of domains of multiple steroid receptors have been solved. One of the earliest examples (1997) of the structure of the ligand-binding domain (LBD) of the estrogen receptor illustrated the drastically different conformer



of the receptor when an agonist vs. an antagonist is bound. As seen in Figure 6.32, when estradiol, an agonist, is bound (left structure), helix 12 (shown in blue) is in one conformation, but when the antagonist raloxifene is bound (right structure), helix 12 (shown in green) rotates by 130°, while the rest of the receptor (shown in red) undergoes minimal change. This is a nice example of stabilization of different receptor conformers by ligand binding, as explained by receptor theory.

6.4.1 Drugs That Target Nuclear Receptors Are Used to Treat Cancer, Inflammation, and Diabetes and to Regulate the Physiological Effects of Sex Hormones

The structures of drugs acting at steroid receptors usually bear resemblance to the natural steroid ligand, and they often have broad effects since steroid receptors are found throughout the body. For example, the wide use of estrogens to treat symptoms of menopause (estrogen replacement therapy) was restricted when it was understood that they may also promote certain estrogen-dependent cancers.

A stilbene scaffold was found to be an antagonist of the ER, thereby blocking the effects of estrogen and having potential use as a contraceptive. In addition, the idea of antagonizing the ER to treat breast tumors emerged in the 1970s, based on the finding that certain breast tumors were dependent on estrogen for their growth. Tamoxifen, which contains the stilbene scaffold (red), was unsuccessful as a contraceptive but found wide use to treat breast tumors that are dependent on estrogen. The chromane ormeloxifene (centchroman), which can be considered a constrained tamoxifen analog, is used as a contraceptive in some countries. Raloxifene, which bears some structural similarity to ormeloxifene and tamoxifen, was ineffective as an anti-cancer agent but is effective in preventing and treating osteoporosis. The effects of tamoxifen and raloxifene are complex. They are termed **selective estrogen receptor modulators** (SERMs) because their activity on the ER varies depending on the location of the receptor. For example, tamoxifen and toremifine block breast tissue ER, but not uterine or endometrial tissue ER, and are used to treat certain breast cancers. Raloxifene mimics estrogen's effects (acts as an agonist) in skeletal tissues (strengthening bones) but blocks estrogen effects (antagonist) in breast and endometrial tissue (proliferation).

In addition to being selective for specific locations of the estrogen receptor (ER), compounds may be selective for specific forms of it. There are two forms of the ER, α and β, and the activated receptor dimers can be either homodimeric or heterodimeric. Selective ERβ agonists have the potential to be used as anti-inflammatory drugs and as neuroprotectants, while selective ERα agonists may be useful to promote cardiovascular health. Although no drugs in this category have yet been approved, ERB 041 advanced to phase II clinical trials as a treatment for rheumatoid arthritis. Designing

Figure 6.32 X-ray structures of ligands bound to the ligand-binding domain of the estrogen receptor in two conformations. The receptor is shown in red, with helix 12 highlighted. *Left*, X-ray structure of ER LBD bound to estradiol, an agonist; helix 12 is shown in blue. *Right*, X-ray structure of ER LBD bound to raloxifene, an antagonist; helix 12, shown in green, rotates by 130° compared to the agonist bound structure. The hydrophobic residues in the groove between helix 5 and helix 3 are shown as space-filling spheres and highlighted in yellow. Residue K362 (shown in pink) is a lysine required for receptor activation, which is partially buried by the proximity of helix 12 when an antagonist binds.

Source: From Brzozowski, A. M.; Pike, A. C. W.; Dauter, Z.; et al. (1997). *Nature* 389:753–758. With permission from Macmillan Publishers Ltd.

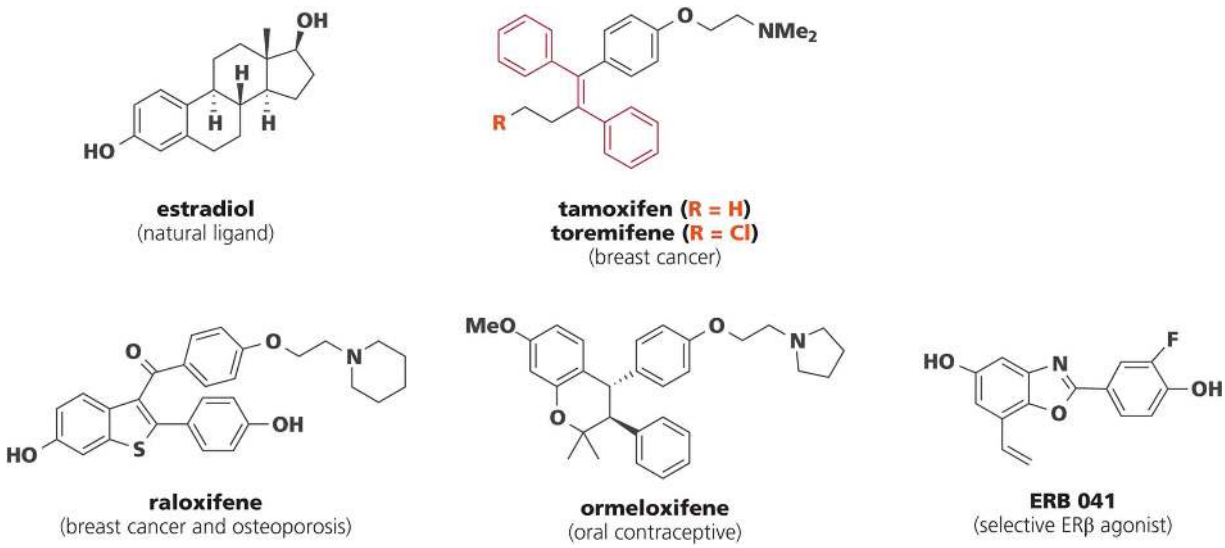


Figure 6.33 Examples of compounds acting at estrogen receptors. The natural ligand estradiol, styrenes tamoxifen and toremifene (styrene in red), and related analogs raloxifene and ormeloxifene. The last four are considered SERMs and are approved for various uses, including breast cancer, contraception, and osteoporosis. ERB 041 is an experimental compound that only interacts with one form of the ER.

drugs to target the ER is further complicated by the involvement of co-activators or co-repressors, proteins that interact with the ER and either increase or decrease gene expression. For example, tamoxifen's differential activity (antagonist in breast tissue, agonist in bone) is believed to be due to different levels of specific co-activator proteins in breast and bone tissue. Further complicating the understanding of this system is the discovery of other receptors, such as the G protein–coupled receptor GPR30, which was recently found to be activated by estrogen. Structures of the natural ligand estradiol, as well as synthetic ER modulators, are shown in **Figure 6.33**.

A more recent approach is the development of selective estrogen receptor degraders (SERDs), shown in **Figure 6.34**. Fulvestrant, given intramuscularly, was approved in 2002 and is an ER antagonist, but it also impairs ER dimerization. The fulvestrant–ER complex is unstable, and this leads to the degradation of the ER protein by the proteasome. An orally active drug with a similar mechanism, elacestrant, was approved in 2023. There are a number of other drugs in clinical trials that act via degrading steroid receptors; these will be discussed further in Chapters 8 and 10.

Drugs acting at androgen receptors, known as selective androgen receptor modulators (SARMs), are useful to treat prostate cancer (antagonists) as well as for their anabolic effects on bone and muscle (agonists). The natural ligands for the androgen receptors are testosterone and its reduction product, 5α-dihydrotestosterone (**Figure 6.35**). Examples of SARMs include the agonist stanozolol, used for its anabolic effects in building muscle mass, and the antagonists flutamide and bicalutamide, used to treat prostate cancer and other disorders involving overactive androgen receptors. Other uses for SARMs include treatment of alopecia (hair loss) and acne. Some side effects that arise with drugs that are analogs of endogenous steroids are due to their activity at other steroid receptors, as well as the fact that they may be substrates for enzymes that act on steroid-like substrates. These may include virilism due to 5α-reductase activity and feminization due to aromatase substrate activity.

Other steroid receptors include the glucocorticoid (GR) and mineralocorticoid (MR) receptors. The glucocorticoid receptor has a role in inflammation as well as in

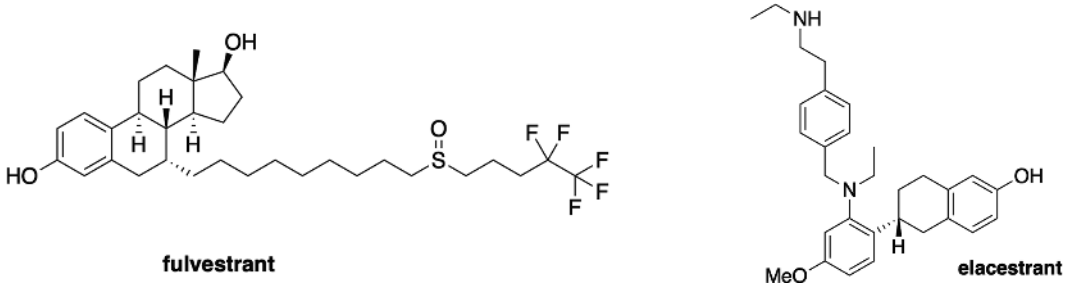


Figure 6.34 Structures of the SERDs fulvestrant and elacestrant.

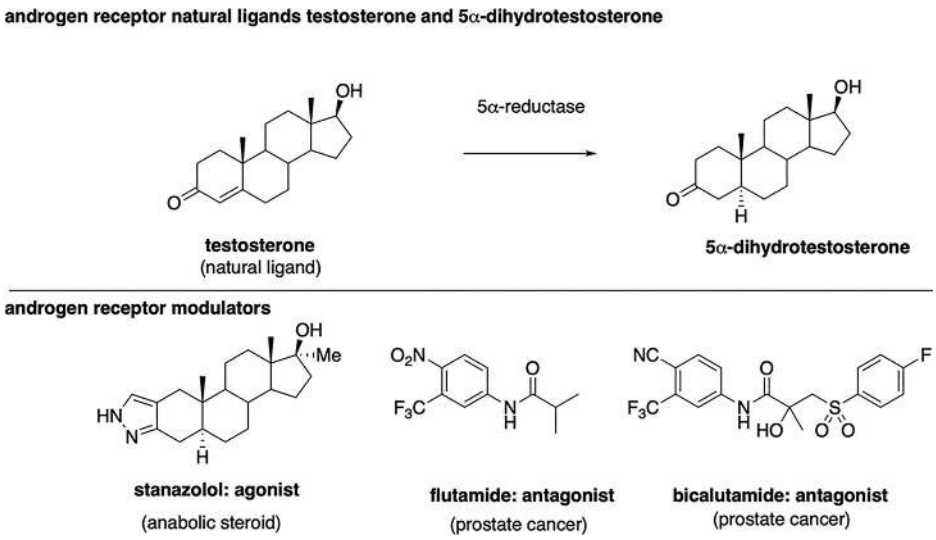


Figure 6.35 Examples of compounds acting at androgen receptors. *Top*, testosterone and its reduction product, 5α-dihydrotestosterone, natural ligands of the AR. *Bottom*, structures of selective androgen receptor agonist stanozolol and antagonists flutamide and bicalutamide.

fat metabolism, glucose regulation, and bone formation. The natural ligand cortisol is used as an anti-inflammatory drug, as is the synthetic analog dexamethasone (Figure 6.36). Cortisol binds at both GR and MR; however, the natural ligand for MR is aldosterone. This steroid is involved in the regulation of plasma Na^+ and K^+ levels and affects blood pressure. The synthetic agonist fludrocortisone is used in conditions such as Addison's disease, where there is a deficiency of aldosterone.

The three **peroxisome proliferator–activated receptors** (PPARs) α, γ, and δ, which are involved in carbohydrate and lipid metabolism and adipose tissue differentiation, are another important nuclear receptor drug target. Their natural ligands are fatty acids and eicosanoids, and binding of those ligands results in the formation of heterodimers with the retinoid X receptor (RXR). The heterodimers then bind to peroxisome proliferator response elements (PPRE), resulting in the transcription of target genes. These receptors are affected by the drug class called fibrates. Clofibrate was reported in the 1960s by ICI Pharmaceuticals and approved as a lipid- and cholesterol-lowering

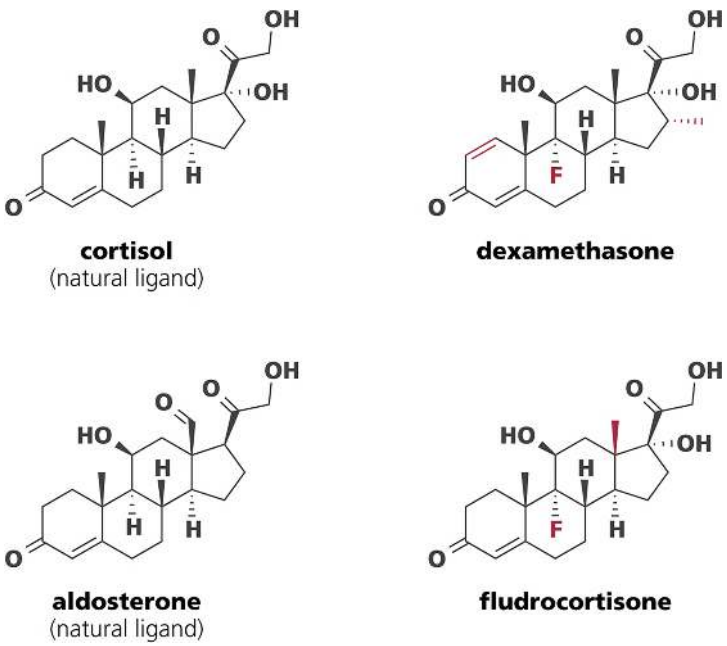


Figure 6.36 Examples of compounds acting at glucocorticoid and mineralocorticoid receptors. Cortisol (GR and MR) and aldosterone (MR) are natural ligands. Both cortisol and dexamethasone, an agonist, are used as anti-inflammatory drugs, and fludrocortisone (MR agonist) is used in diseases where there is a deficiency of aldosterone. Changes in dexamethasone and fludrocortisone, compared to the natural ligand, are in red.

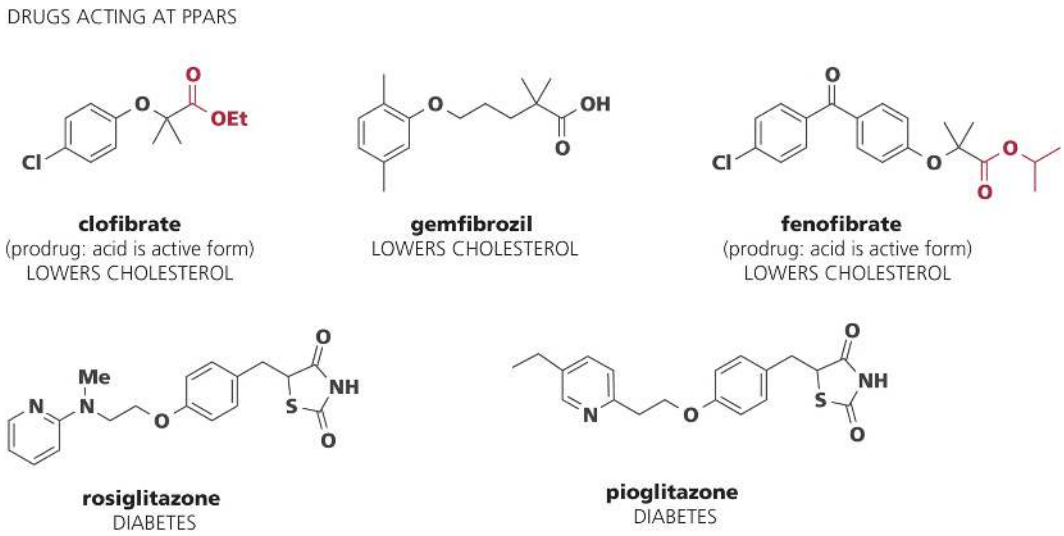


Figure 6.37 Drugs acting at peroxisome proliferator–activated receptors. The fibrates clofibrate, gemfibrozil, and fenofibrate activate PPARα and are used to lower cholesterol levels (esters in prodrugs in red). The thiazolidinediones rosiglitazone and pioglitazone activate PPARγ and are used to treat diabetes.

agent. It is a prodrug that generates the active carboxylic acid *in vivo*. It has since been replaced by safer analogs, such as the agonists gemfibrozil and fenofibrate, which are used to treat hypercholesteremia and hyperlipidemia (**Figure 6.37**). The mechanism of the fibrates, not determined until 1990, involves the activation of the PPARs, with approximately 10-fold selectivity for PPARα as compared to PPARγ.

Thiazolidinediones, exemplified by rosiglitazone and pioglitazone in **Figure 6.37**, are used to treat diabetes and are agonists at PPARγ. Like the fibrates, this class of drugs was identified on the basis of their anti-diabetic activity, and the mechanism of action was not discovered until later. The discovery of rosiglitazone is described in Case Study 2 at the end of this chapter.

6.5 TRANSPORTERS AS DRUG TARGETS

Transporters are membrane-bound proteins that, like receptors, respond to a signal outside the cell. They are also related to ion channels in that they are active or inactive based on the concentration of the ligand they transport. Unlike ion channels, they bind and transport far fewer ligands or ions per cycle. These ligands include amino acids, carbohydrates, and neurotransmitters, as well as ions. Transporters are classified as uniporters, symporters, and antiporters (**Figure 6.38**). Whereas uniporters transport one molecule or ion at a time along a concentration gradient, symporters and antiporters couple the transport of two molecules or ions. A *symporter* couples movement of one molecule against its concentration gradient with movement of a different molecule along its concentration gradient. An *antiporter* transports two different molecules or ions in opposite directions, one with and one against its gradient. Transporters play an important role in maintaining the appropriate concentrations of nutrients, ions, and neurotransmitters in the cell and other environments.

Structural details of transporters have been clarified with the solution of several X-ray structures, such as bacterial LeuT, a sodium-coupled leucine symporter. All transporters have in common a series of structural repeats, each of which has several contiguous transmembrane segments (TM), which form folds. These folds result in a membrane-spanning complex with an outward side and an inward side. In a simplified model for neurotransmitter sodium symporters, binding of Na^+ in the outward open state facilitates binding of the ligand. The outward side closes, followed by the opening of the inward side and the release of Na^+ and ligand.

Assays for transporter activity rely on making synaptosome preparations from the homogenization of tissue rich in these receptors or the use of engineered cells that overexpress the transporter of interest. By tagging the natural ligand (with radioactivity, fluorescence, etc.), the rate of uptake can be measured by adding tagged ligand to the suspension for a specified period of time and detecting and measuring the movement of the tag. Molecules that inhibit transporters will slow or inhibit the protein's

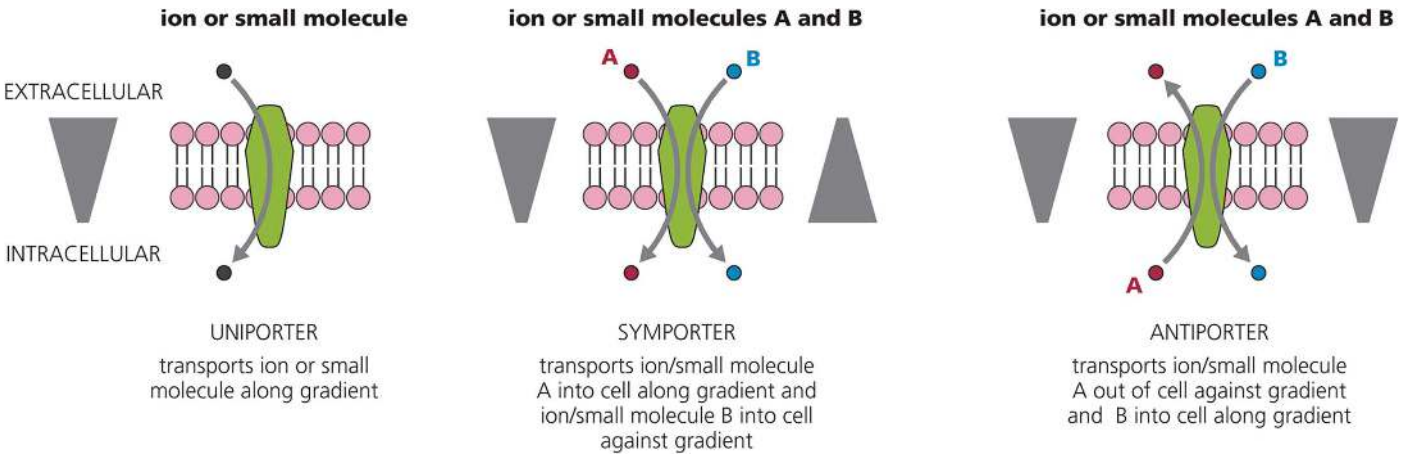


Figure 6.38 General structural features of transporters. Uniporters transport one ion or small molecule (black) along a concentration or electrochemical gradient. A symporter transports two different molecules or ions in the same direction, one with the gradient (A, red) and one (B, blue) against it. An antiporter transports two different molecules or ions (A, red; B, blue) in opposite directions, one with the gradient and one against (transporters in green, membrane in pink, gradient in gray triangles).

ability to take up the tagged ligand as a function of concentration of the drug. To date, only a handful of drug targets are transporters, although the most widely used antidepressants inhibit the serotonin transporter. Most drugs that act on transporters are structurally related to the ligand they transport, and identification often relies on design based on the natural ligand or a known active compound. An important area that has exploited transporters is the development of imaging agents. In these cases, the localization of specific transporters in certain cell types (for example, tumors) can be exploited to transport labeled molecules, for example, those containing a positron-emitting atom, as an imaging agent. Another emerging area is the use of transporters to selectively deliver drugs to certain organs. Table 6.5 lists examples of transporters and the drugs or tools that bind to them.

Table 6.5 Examples of Transporters, Their Ligands and Drugs, and the Approved Indication

Natural Ligand	Transporter	Drug or Ligand Examples	Indication
Serotonin	Serotonin transporter (SERT)	Fluoxetine, paroxetine	Depression
Norepinephrine	Norepinephrine transporter	Methylphenidate	Attention deficit hyperactivity disorder (ADHD)
Na^+/Cl^-	Sodium-chloride co-transporter (NCC)	Hydrochlorothiazide	Diuretics, hypertension
$Na^+/K^+/2Cl^-$	Sodium-potassium-chloride co-transporter (NKCC)	Furosemide	Diuretics, hypertension
Glucose	Sodium-glucose transporter (SGLT)	Canagliflozin, empagliflozin	Diabetes
		^{18}F-fluorodeoxyglucose (FDG)	Tumor imaging agent

6.5.1 Drugs Acting at Transporters Treat a Wide Array of Neuropsychiatric Disorders, Hypertension, and Diabetes

Transporters are important drug targets in CNS diseases such as depression, where neurotransmitter transporters are involved in the regulation of the levels of serotonin, noradrenaline, and dopamine. They are also an important target for the treatment of hypertension, where modulating Na^+/Cl^- and $Na^+/K^+/Cl^-$ symporters in the kidney can result in diuretic effects. A relatively new target is the sodium-glucose transporter, which is involved in the removal of glucose and is an anti-diabetic drug target.

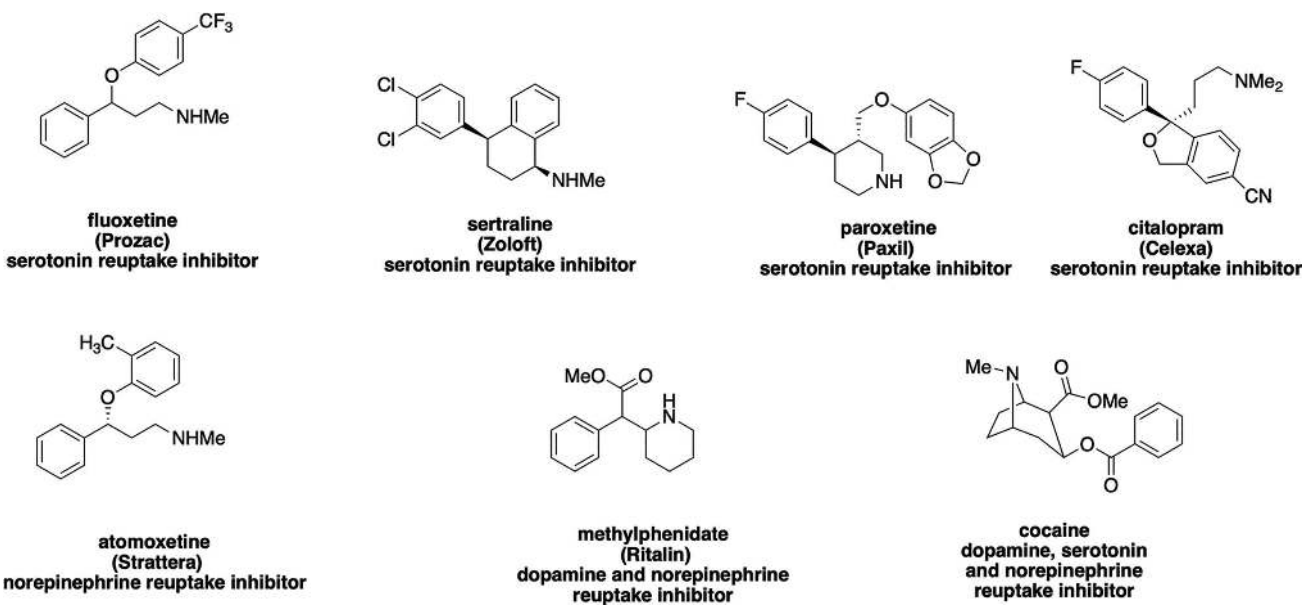


Figure 6.39 Drugs acting at neurotransmitter transporters. The serotonin reuptake inhibitors fluoxetine, sertraline, paroxetine, and citalopram are used as antidepressants. Norepinephrine reuptake inhibitors atomoxetine and methylphenidate are used to treat ADHD. Cocaine, a drug of abuse, has activity at several neurotransmitter transporters.

In the CNS, neurotransmitters are released at the synapse, where they diffuse across and bind to receptors on the postsynaptic cell. Transporters in neurons regulate the concentration of the neurotransmitter at the synapse by transporting it into the cell when the synapse concentration is high. Examples of transporters working in reverse and releasing neurotransmitters into the synapse also exist. Some neurotransmitter transporters co-transport ions, usually Na^+, along with the neurotransmitters and are classified as neurotransmitter sodium symporters (NSS).

Drugs that act at NSS sites include neurotransmitter reuptake inhibitors (**Figure 6.39**), which are widely used to treat anxiety and depression. Selective serotonin reuptake inhibitors (SSRIs), such as fluoxetine, sertraline, paroxetine, and citalopram, block the serotonin transporter, resulting in higher concentrations of serotonin in the synapse. These compounds have revolutionized the treatment of depression. Norepinephrine reuptake inhibitors (NRIs), such as atomoxetine and methylphenidate, are often used to treat attention deficit hyperactivity disorder (ADHD). The success of the SSRIs has led to the development of combined serotonin–norepinephrine reuptake inhibitors (SNRIs) to treat depression and anxiety. Drugs of abuse, including cocaine and amphetamines, also affect dopamine and norepinephrine reuptake mediated by transporters as part of their mechanism of action. Many drugs in this category will be discussed in more detail in Chapter 13.

There are two classes of diuretics that act on co-transporters and are widely used to treat hypertension and cardiovascular disease. One class includes thiazides, such as chlorothiazide, hydrochlorothiazide, and chlorthalidone, which were initially discovered in the 1950s during investigations of sulfonamide antibacterial agents. The sulfonamides in this class are not antibacterial and instead act as diuretics. Long after they were introduced, the mechanism of action was identified as inhibition of co-transporters in the kidney. Chlorothiazide and hydrochlorothiazide block the Na^+/Cl^- co-transporter in the distal convoluted tubule of the kidney. Prevention of Na^+ and Cl^- absorption leads to increased elimination of fluid. The other class of diuretics target the $Na^+/K^+/2Cl^-$ co-transporter (NKCC) in the loop of Henle in the kidney. Diuretics blocking this target include furosemide, bumetanide, and torasemide, with some evidence that they compete for the chloride ion site on the transporter (**Figure 6.40**). Several of these diuretics have been banned from use by athletic organizations due to their ability to mask illicit drugs by increasing their elimination in urine.

Recent interest in sodium-glucose transport proteins (SGLT) has stemmed from efforts to identify new targets to treat diabetes, as these proteins are involved in the reabsorption of glucose. Specifically, subtype 2 (SGLT2) is responsible for glucose reabsorption in the kidney, as opposed to subtype 1 (SGLT1), which operates in the

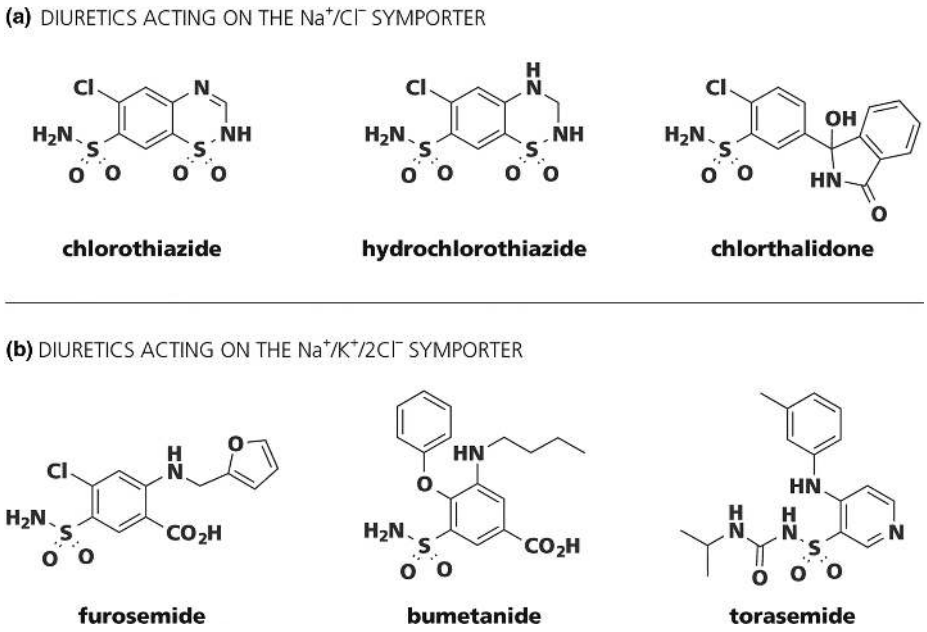


Figure 6.40 Diuretics acting on symporters. (a) The thiazide class (chlorothiazide, hydrochlorothiazide, and chlorthalidone) acts on the Na^+/Cl^- co-transporter in the distal convoluted tubule and is used as diuretics. (b) Drugs acting on the $Na^+/K^+/2Cl^-$ co-transporter in the loop of Henle, furosemide, bumetanide, and torasemide are also used as diuretics.

intestine. Blocking reabsorption of glucose in the kidney leads to the elimination of glucose through urine and lower levels of blood glucose, so selective SGLT2 blockers would be effective anti-diabetic drugs. Phlorizin, found in apple trees and shown to induce excretion of glucose through the urine (glucosuria) in the 1880s, was identified as a nonselective SGLT inhibitor in the early 2000s. Analogs of phlorizin with lipophilic groups replacing the phenols exhibited selectivity for SGLT2, and both O- and C-glycosides were investigated by several research groups. Stability to glucosidases as well as selectivity and potency were achieved with the C-aryl analogs. Drugs in the SGLT2 inhibitor class approved to treat type 2 diabetes include canagliflozin (Invokana, approved in 2013), dapagliflozin (Farxiga), empagliflozin (Jardiance), and bexagliflozin (Benzavvy). (Figure 6.41).

Drugs to treat cystic fibrosis act through a unique type of transporter, called CFTR (**c**ystic **f**ibrosis **t**ransmembrane conductance **r**egulator). While CFTR is a member of the ABC (**A**TP-**b**inding **c**assette) transporter family, like the transporter P-gp described in Chapter 5, it functions as an ion channel to regulate the flow of chloride ions. Mutations in the gene that codes for CFTR result in either the production of a

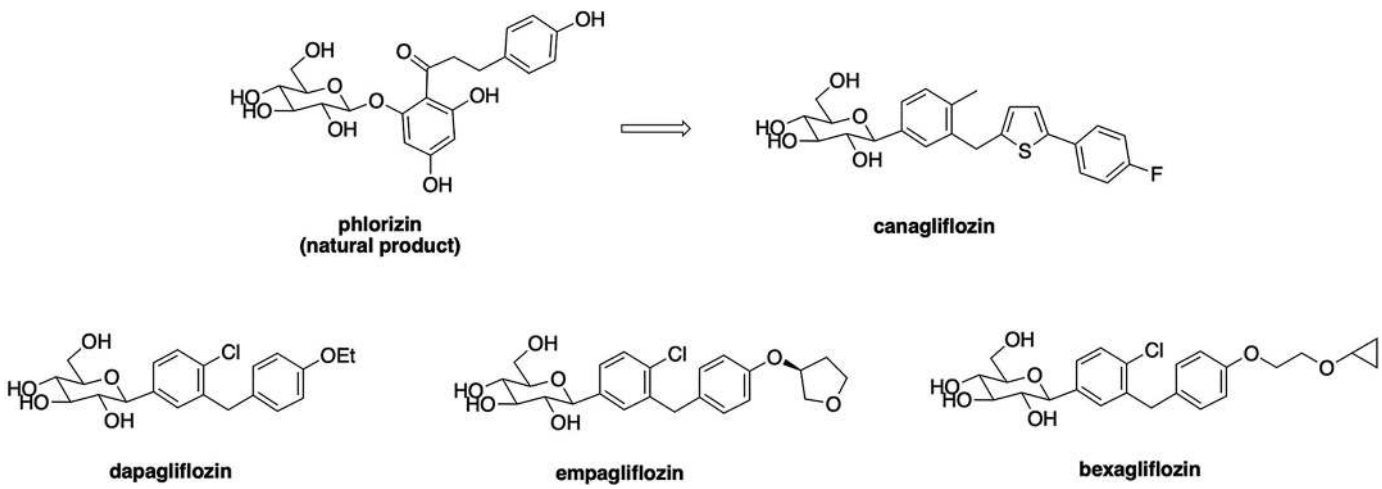


Figure 6.41 Phlorizin and analogs that target SGLT. The natural product phlorizin was found to inhibit SGLT. Analogs where the glycosidic phenol groups were replaced by a C-aryl glycoside are selective SGLT2 and are approved to treat diabetes. This class, called glifozins, includes canaglifozin, dapagliflozin, empagliflozin, and bexagliflozin.

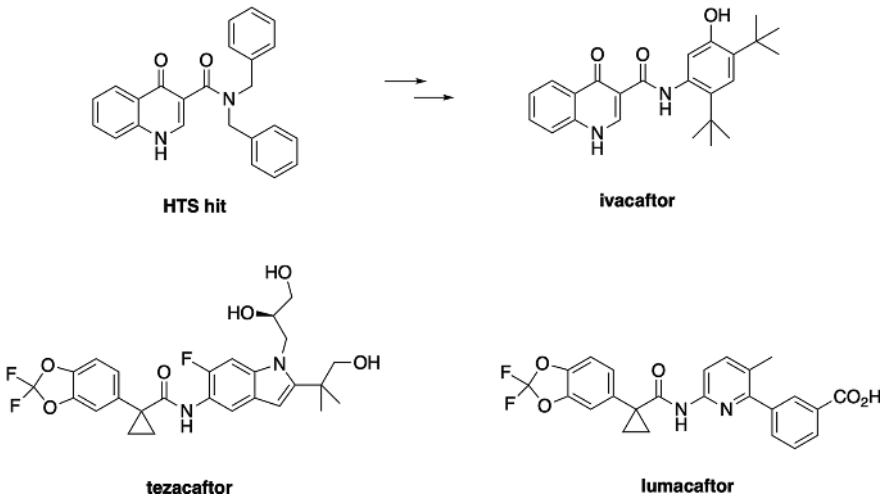


Figure 6.42 Original hit leading to ivacaftor and structures of drugs used to treat cystic fibrosis. Ivacaftor increases chloride ion flow in patients with a specific CTFR mutation. Tezacaftor and lumacaftor allow CTFR to reach the cell surface, thereby allowing ion flow.

faulty protein or the inability of the protein to reach the cell surface. Regardless of the specific mutation, chloride ions are not adequately transported, resulting in a buildup of thick secretions in the lung and other tissues. Ivacaftor, developed from an HTS hit, potentiates CFTR in patients with a specific mutation (G551D), resulting in increased chloride ion flow. Ivacaftor is used in combination with lumacaftor and tezacaftor, which are not potentiators but stabilize the protein in patients with F508del mutation (deletion of Phe508) so that it reaches the cell surface (**Figure 6.42**).

6.6 SUMMARY

Receptors, as a broad class, are the target of more than two-thirds of all marketed drugs. Although diverse in structure and function, GPCRs, ion channels, nuclear receptors, and transporters are all proteins that transmit a signal to effect a physiological result. Receptors of all classes usually exist as multiple subtypes, and the physiological result of signal transmission may differ with the subtype of receptor as well as its tissue distribution. Drugs acting at receptors can either increase or decrease the level of signal transmission. Many older drugs act at receptors, but their mechanism of action was discovered well after they were in clinical use. Natural ligands for receptors, as well as natural products, have been a valuable source of starting points for drug discovery projects that target receptors. As such, medicinal chemistry optimization approaches often rely on knowledge of the structure of the natural ligand or on pharmacophore modeling based on known ligands. Recent advances in various technologies have allowed high-throughput screening to be a more feasible approach in solving structures of receptors, and this has started to inform new approaches in structure-based drug design. However, the understanding of how to design agonists and antagonists of these highly complex systems *a priori* is still a very immature field. Most recently, drug discovery strategies target specific receptors or receptor subtypes, with the goal of generating efficacious molecules that do not exhibit many of the side effects that plague older compounds.

CASE STUDY 1 DISCOVERY OF RAMELTEON

Insomnia, characterized by an inability to go to sleep or to stay asleep, is estimated to affect about one-third of all Americans. Standard treatment includes benzodiazepines, zolpidem (Ambien), and eszopiclone (Lunesta), all of which act at the $GABA_A$ receptor. However, their long-term use is discouraged due to abuse potential. Sleep patterns are affected by melatonin, a hormone that has a role in regulating circadian rhythms. The hormone itself has a short half-life, and there is little data on its long-term safety, although it is sold over the counter and used as a dietary supplement to promote sleep. It binds to the melatonin receptor (MT), a GPCR that exists as three subtypes. Both the MT_1 and MT_2 subtypes are involved in regulating circadian rhythm patterns. The MT_1 subtype is expressed throughout the brain, while MT_2 is found primarily in the retina. A third subtype, MT_3, has been identified as a form of the enzyme quinone reductase, but its function is unknown. Osamu Uchikawa and others at Takeda Pharmaceuticals began a program to find melatonin receptor agonists that would have selectivity for the MT_1 receptor, based on the assumption that its localization in the brain made it the more important target for regulating sleep patterns. Their initial study, based on a CoMFA analysis of 133 compounds that bound to the melatonin receptor, indicated that the pyrrole of melatonin (**Figure 6.43**) was not necessary for binding but the methoxy and amide groups (shown in red) were needed. To this end, a series of bicyclic compounds were synthesized, with variations in ring size (n), levels and positions of saturation/unsaturation, aryl substitution (OR^1, R^2,), cycloalkyl (R^3) substituent, and amide substituent (R^4).

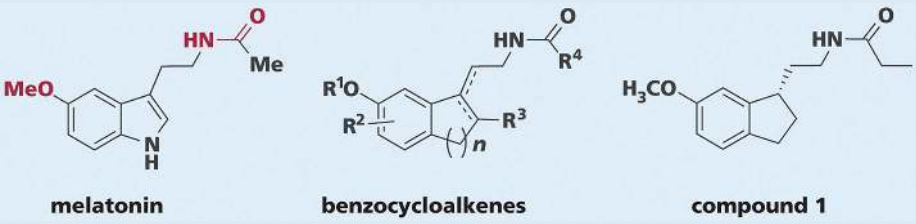


Figure 6.43 Structures of early analogs leading to the development of ramelteon. The natural ligand melatonin, along with 132 other compounds, were used in a CoMFA analysis. The analysis indicated that the methoxy and amide groups of melatonin were required (in red) but the pyrrole was not. This led to the synthesis of a series of bicyclic derivatives. The optimal structure in this series was compound **1** based on potency and selectivity.

The five-membered ring analogs ($n = 1$) were twice as active as those containing a six-membered ring ($n = 2$), and *S*-enantiomers of the saturated compounds exhibited the highest selectivity for MT_1 over MT_3. The optimal structure in this series was found to be **1**, with the *S*-configuration. Compounds were assayed for inhibition of binding of 2-[^{125}I]iodomelatonin to both MT_1 and MT_3 receptors, and **1** showed excellent selectivity (for MT_1, $K_i = 0.041$ nM; for MT_3, $K_i = 3{,}570$ nM).

Based on a homology model of the MT_1 receptor derived from bovine rhodopsin, the oxygen of the methoxy group is proposed to bind to a histidine residue (His195) via an H-bond, the amide NH binds to a serine OH side chain (Ser144), and the amide carbonyl participates in an H-bond with the OH of another serine (Ser110), as seen in **Figure 6.44**.

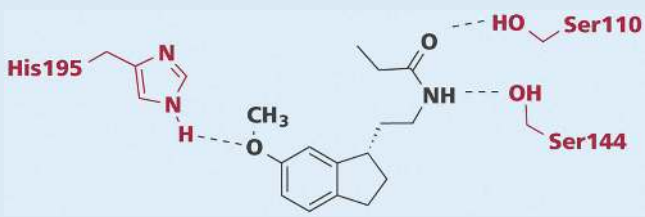


Figure 6.44 Proposed binding interactions between compound 1 and MT_1 receptor. Based on a homology model of the MT_1 receptor, the methoxy oxygen is proposed to hydrogen-bond to His195, with other H-bonds from the amide NH to Ser144 and the amide carbonyl to Ser110.

To optimize the orientation of the methoxy oxygen lone pair for hydrogen bonding, the next iteration of optimization added a ring constraint, generating tricyclic analogs with either a linear or angular fused system. The angular fused system was more potent than the linear fused system and had similar or slightly reduced potency and selectivity

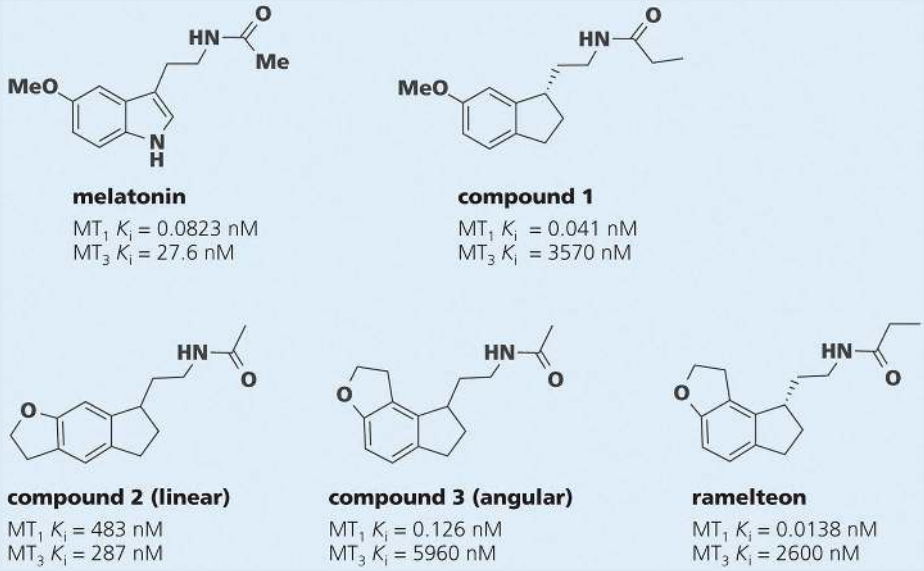


Figure 6.45 Structure and potencies of melatonin and synthetic analogs. Compound 1 is more potent and selective for MT_1 vs. MT_3 as compared to melatonin. A series of tricyclic analogs incorporated ring constraint to optimize the orientation of the ether oxygen. The linear tricyclic analog **2** is less potent and less selective than the angular tricyclic analog **3**. Ramelteon, containing the propanamide with the *S*-configuration, was approved in 2005 for insomnia.

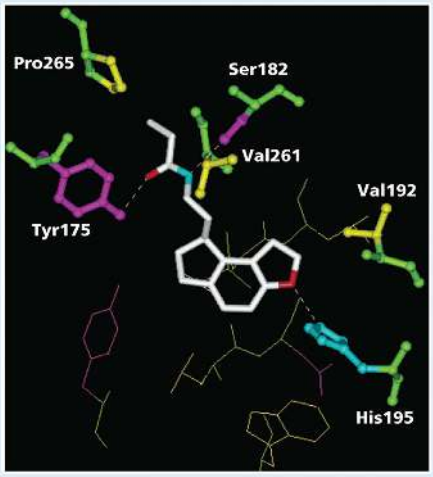


Figure 6.46 Model of ramelteon in binding site of MT_1 receptor. Using a homology model, key interactions include the furan oxygen to His195, and H-bonds of the amide NH and carbonyl to Ser185 and Tyr175, respectively.

Source: From Uchikawa, O.; Fukatsu, K.; Tokunoh, R.; et al. (2002). "Synthesis of a novel series of tricyclic indan derivatives as melatonin receptor agonists." *J Med Chem* 45:4,222–4,239. With permission from the American Chemical Society.

compared to **1**. For example, the linear compound **2** had about 4,000-fold less affinity for MT_1 than the angular compound **3** and had no selectivity vs. MT_3 (Figure 6.45).

Increasing the size of the tetrahydrofuran ring to a six-membered ring or substituting a nitrogen into that ring did not further improve activity. Overall, the optimized candidate based on potency and selectivity was ramelteon, which contained the *S*-configuration and the propenamide side chain. Docking ramelteon into an MT_1 homology model shows binding interactions between the furan oxygen and His195 and hydrogen bonding of the amide NH and carbonyl to serine and tyrosine residues (Figure 6.46).

In addition to M_1 vs. MT_3 selectivity, ramelteon showed no affinity for $GABA_A$ receptors or other neurotransmitter receptors. *In vivo* testing in cats showed sleep-promoting activity, with an increase in slow-wave sleep and rapid eye movement (REM) sleep. Ramelteon entered clinical trials in both younger and older adults with chronic insomnia. Both studies showed a decrease in the amount of time to enter sleep as compared to placebo. No evidence for abuse, withdrawal, or residual effects was found, and ramelteon was approved by the FDA in 2005.

CASE STUDY 2 DISCOVERY OF ROSIGLITAZONE

In the early 1990s, the most widely used drugs for treating non-insulin-dependent diabetes mellitus (NIDDM) were the sulfonylureas, which act by increasing insulin secretion. Undesirable side effects include hypoglycemia and weight gain. At Glaxo, the glitazone compound ciglitazone was found by screening compounds for their effects in rodent models of insulin resistance. Interestingly, earlier, in 1982, a different pharmaceutical company, Takeda, had reported the same compound (named ADD-3878) as having anti-diabetic activity in animal models (Figure 6.47). The Takeda scientists

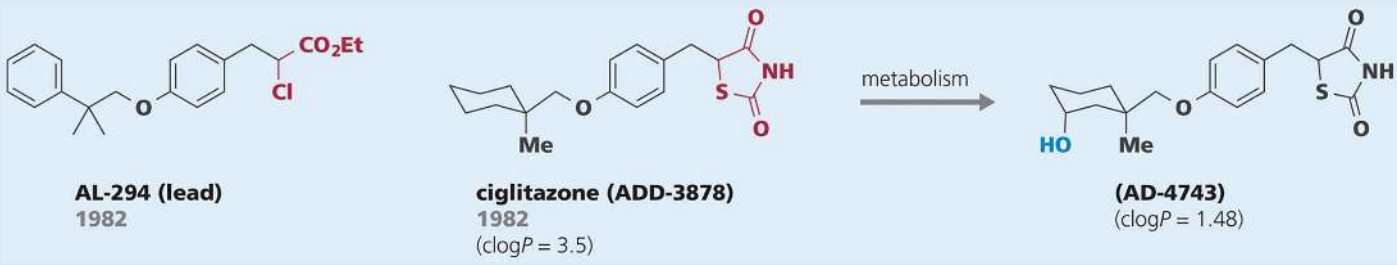
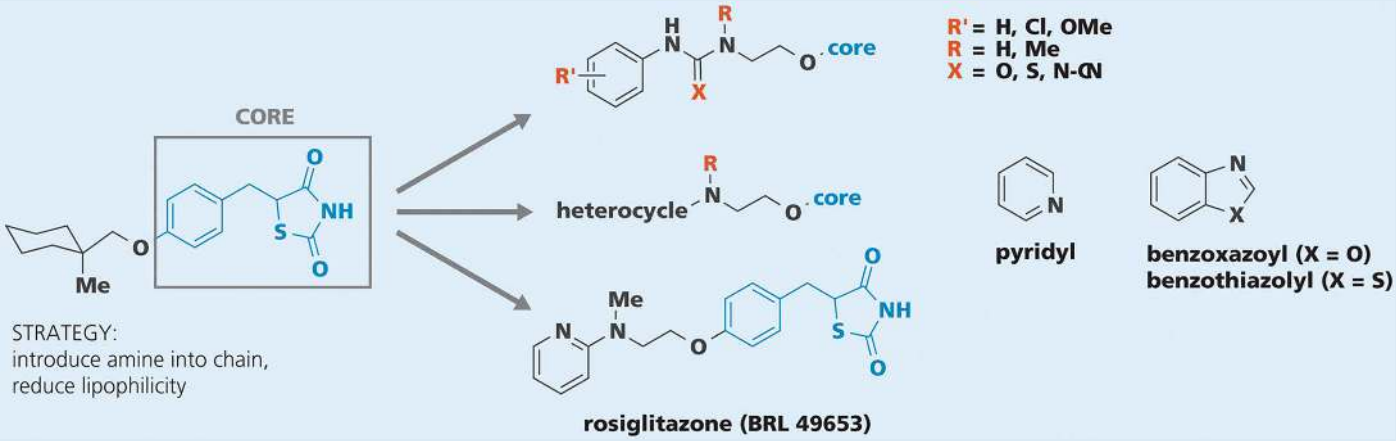


Figure 6.47 Early leads in the development of rosiglitazone. An initial lead, compound AL-294, was modified at Takeda by the substitution of a thiazolidinedione (red) for the chloroester to give ciglitazone (ADD-3878) the same compound identified as a lead in a screening program at Glaxo. The less lipophilic metabolite AD-4743 was more potent in mice.



had synthesized ADD-3878 based on their lead compound, AL-294, and reported that the substitution of the thiazolidinedione for the α-chloro ester improved hypoglycemic activity and decreased toxicity. The overall mechanism in insulin-resistant animal models was an increase in sensitivity to endogenous insulin rather than an increase in secretion. They also reported that a hydroxylated metabolite, AD-4743, was even more potent in the *in vivo* model.

Scientists at Glaxo pursued the observation that the less lipophilic (clogP = 1.48 vs. 3.5) AD-4743 was a more potent anti-hyperglycemic agent in genetically obese or glucose-intolerant mice than ciglitazone. They prepared a library of analogs with various amines/amine derivatives as replacements for the cyclohexyl methyl group (Figure 6.48), including ureas, thioureas and cyanoguanidines, and amino heterocycles. These replacements lowered lipophilicity; several aryl urea and thiourea analogs were as effective as ciglitazone, and heterocyclic analogs were up to 300 times more active than the parent compound AD-4743.

One of the most potent compounds, BRL 49653 (rosiglitazone), had a minimum effective dose of 3 μmol/kg of diet. Although the clog*P* of rosiglitazone is 1.53, it was found that other analogs with similar lipophilicity were 100 times less active, indicating that lipophilicity is not the sole determining factor for activity. Screening of several of the more potent analogs for toxicity was based on their ability to lower blood hemoglobin levels; rosiglitazone had the best profile, with no effect on hemoglobin at doses 100-fold greater than the anti-hyperglycemic dose level.

Although screening and drug development were based on animal models of anti-hyperglycemia, rosiglitazone was later found to act as a selective PPARγ agonist. This receptor is primarily expressed in adipose tissue, and activation leads to adipocyte differentiation and improved insulin signaling in mature adipocytes.

Rosiglitazone is one of several thiazolidinediones that received FDA approval to treat type II diabetes, with other examples including pioglitazone and troglitazone (Figure 6.49). Since their introduction, all have had some issues with side effects: troglitazone was withdrawn due to liver toxicity, pioglitazone conferred an increased risk for bladder cancer, and rosiglitazone increased the risk of heart attack.

Figure 6.48 Amine analogs of the ciglitazone lead. A library of amine analogs of ciglitazone was prepared in order to reduce lipophilicity by replacing the cyclohexyl while retaining the thiazolidinedione core (blue). Substitutions included phenylureas, thioureas, and cyanoguanidines (top) and heterocyclic amines (middle). The heterocycles were the most potent, and the pyridyl compound rosiglitazone was approved to treat diabetes.

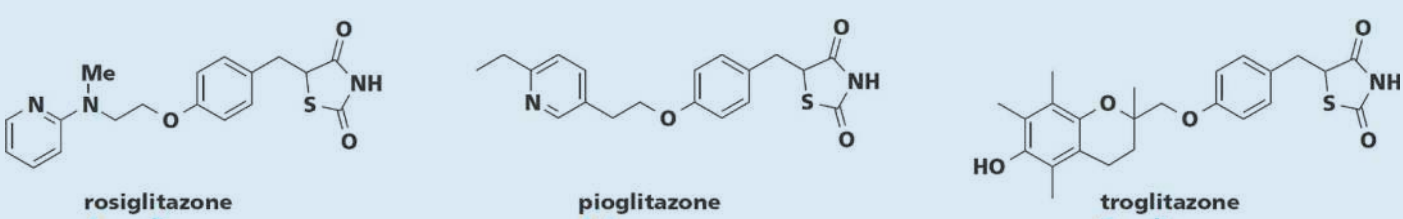


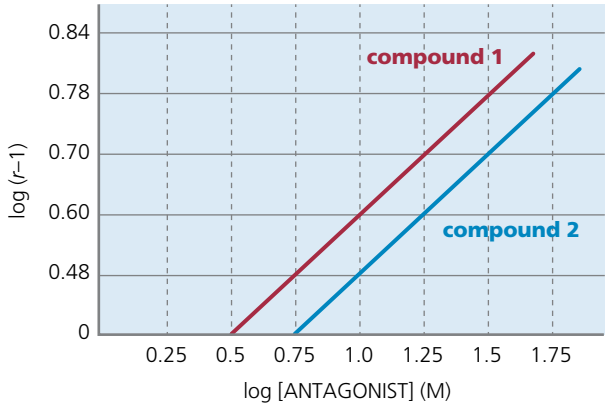
Figure 6.49 Structures of anti-diabetic thiazolidinediones. Rosiglitazone, pioglitazone, and troglitazone all received FDA approval to treat diabetes, although the latter two have been withdrawn due to toxicity issues.

6.7 REVIEW QUESTIONS

1. Summarize the basic differences between muscarinic and nicotinic acetylcholine receptors.
2. What structural feature makes salbutamol a longer-acting drug than isoprenaline?
3. What are the three major families of ligand-gated ion channels?
4. Name four members of the class 1 nuclear receptor superfamily.
5. Name three members of the VGIC family and a type of drug acting at each one.
6. What type of receptor is characterized in part by a zinc finger region that binds to DNA?
7. Name four major families (based on protein structure) of drug receptors, and give a specific example of each one.
8. Explain the difference between an agonist and an antagonist.
9. LGICs, VGICs, and neurotransmitter transporters all allow ions to pass through. Explain how they differ.
10. What hit identification strategies are typically used to start drug discovery programs with GPCR targets, ion channel targets, nuclear receptor targets, and transporter targets? For each specific target class, explain why that strategy is used more than others and why some strategies are not often used.

6.8 APPLICATION QUESTIONS

1. A Schild plot is shown for two antagonists, **1** and **2**. Which compound is more active? What is the pA_2 of each compound? Explain what *r* is.



2. Research the drug lorcaserin. What is its structure? What GPCR does it affect? What is its selectivity profile against related GPCRs? Is it an agonist or an antagonist? Compare its structure to that of the natural ligand of its GPCR target. Draw a pharmacophore that is consistent with the two structures.
3. In Case Study 1, a figure shows ramelteon docked into the putative active site of the MT_1 receptor. Using this figure and the structure of ramelteon, design three analogs that, when tested, can test the proposed binding hypothesis.
4. Using the specific examples of drugs and compounds described in this chapter, find at least one example of (a) a privileged structure, (b) an isostere, (c) a library synthesis, and (d) the use of conformational constraint.
5. Research the drug semaglutide. What is its structure, and how does it work? Research current efforts to identify small molecules that work through the same mechanism.

6. **Figure 6.39** shows the structures of four well-known selective serotonin reuptake inhibitors. Of the four structures, choose two of them and draw pharmacophore models based on their structures. Explain your rationale for the model and which features of the two chosen compounds fit this model. For the two compounds you did not choose, do they fit your pharmacophore model? Explain why or why not.
7. Natural products have been an abundant source of modulators of receptors. Research the structure of two different natural products (not described in this chapter) that exert their effects on receptors. What is their structure, what class of receptor do they target, and what is their physiological effect?
8. Design three new analogs of the PPAR agonist gemfibrozil in **Figure 6.37** by incorporating different acid isosteres.
9. Research the Nobel Prize in Chemistry in 2003. Who received the award, and for what work?
10. Do the structures of estradiol and raloxifene (shown in **Figure 6.33**) adhere to Lipinski's rules? Use an online program or ChemDraw to find the log$P$.

FURTHER READING

General

Alexander SPH, Kelly E, Marrion N et al. (2015) The concise guide to pharmacology 2015/16: overview. *Brit J Pharmacol* 172:5729–5743.

Hopkins AL & Groom CR (2002) The druggable genome. *Nat Rev Drug Discov* 1:727–730.

Kenakin T (2004) Principles: receptor theory in pharmacology. *Trends Pharmacol Sci* 25:186–192.

Leff P (1995) The two-state model of receptor activation. *Trends Pharmacol Sci* 16:89–97.

Maehle A-H (2004) Receptive substances: John Newport Langley (1852–1925) and his path to a receptor theory of drug action. *Med Hist* 48:53–174.

Neubig RR, Spedding M, Kenakin T & Christopoulos A (2003) International union of pharmacology committee on receptor nomenclature and drug classification. XXXVII: update on terms and symbols in quantitative pharmacology. *Pharmacol Rev* 55:597–606.

Overington JP, Al-Lazikani B & Hopkins AL (2006) How many drug targets are there? *Nat Rev Drug Discov* 5:993–996.

Santos R, Ursu O, Gaulton A et al. (2016) A comprehensive map of molecular drug targets. *Nat Rev Drug Discov* 16:19–34.

G Protein–Coupled Receptors

Bjarnadottir TK, Gloriam DE, Hellstrand SH et al. (2006) Comprehensive repertoire and phylogenetic analysis of the G protein-coupled receptors in human and mouse. *Genomics* 88:263–273.

Congreve M, Langmead CJ, Mason JS & Marshall FH (2011) Progress in structure based drug design for G protein-coupled receptors. *J Med Chem* 54:4283–4311.

Costanzi S (2008) On the applicability of GPCR homology models to computer-aided drug discovery: a comparison between *in silico* and crystal structures of the β_2-adrenergic receptor. *J Med Chem* 51:2907–2914.

del Alamo D, Sala D, Mchaourab HS & Meiler J (2022) Sampling alternative conformational states of transporters and receptors with AlphaFold2. *eLife* 11:e75751.

Kolb P, Kenakin T, Alexander SPH et al. (2022) Community guidelines for GPCR ligand bias: IUPHAR review 32. *Br J Pharmacol* 179:3651–3674.

Smit MJ, Leurs R, Alewijnse AE et al. (1996) Inverse agonism of histamine H2 antagonists accounts for upregulation of spontaneously active histamine H2 receptors. *Proc Natl Acad Sci USA* 93:6802–6807.

Smith JS, Lefkowitz RJ & Rajagopal S (2018) Biased signaling: from simple switches to allosteric microprocessors. *Nat Rev Drug Discov* 17:243–260.

Ion Channels

Anger T, Madge DJ, Mulla M & Riddall D (2001) Medicinal chemistry of neuronal voltage-gated sodium channel blockers. *J Med Chem* 44:115–137.

Ashcroft F, Benos D, Bezanilla F et al. (2004) Twenty questions. *Nat Rev Drug Discov* 3:239–278.

Bagal SK, Brown AD, Cox P et al. (2013) Ion channels as therapeutic targets: a drug discovery perspective. *J Med Chem* 56:593–624.

Pallotta B (1991) Single ion channel's view of classical receptor theory. *FASEB J* 5:2035–2043.

Unwin N (2005) Refined structure of the nicotinic acetylcholine receptor at 4 Å resolution. *J Mol Biol* 346:967–989.

Nuclear Receptors

Brzozowski AM, Pike ACW, Dauter Z et al. (1997) Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature* 389:753–758.

Glass CK & Ogawa S (2006) Combinatorial roles of nuclear receptors in inflammation and immunity. *Nat Rev Immunol* 6:44–55.

Kojetin DJ & Burris TP (2013) Small molecule modulation of nuclear receptor conformational dynamics: implications for function and drug discovery. *Mol Pharmacol* 83:1–8.

Transporters

Boudker O & Verdon G (2010) Structural perspectives on secondary active transporters. *Trends Pharmacol Sci* 31:418–426.

Focke PJ, Wang X & Larsson HP (2013) Neurotransmitter transporters: structure meets function. *Structure* 21:694–705.

Levine M, Cuendet MA, Khelashvili G & Weinstein H (2016) Allosteric mechanisms of molecular machines at the membrane: transport by sodium-coupled symporters. *Chem Rev* 116:6552–6587.

Ramelteon

Buysse D, Bate G & Kirkpatrick P (2005) Ramelteon. *Nat Rev Drug Discov* 4:881–882.

Fukatsu K, Uchikawa O, Kawada M et al. (2002) Synthesis of a novel series of benzocycloalkene derivatives as melatonin receptor agonists. *J Med Chem* 45:4212–4221.

Uchikawa O, Fukatsu K, Tokunoh R et al. (2002) Synthesis of a novel series of tricyclic indan derivatives as melatonin receptor agonists. *J Med Chem* 45:4222–4239.

Rosiglitazone

Cantello BCC, Cawthorne MA, Cottam GP et al. (1994) [[ω-(Heterocyclylamino)-alkoxy]benzyl]-2,4-thiazolidinediones as potent antihyperglycemic agents. *J Med Chem* 37:3977–3985.

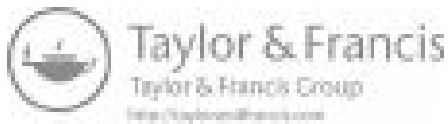
Edvardsson U, Bergstrom M, Alexandersson M et al. (1999) Rosiglitazone (BRL49653), a PPARγ-selective agonist, causes peroxisome proliferator-like liver effects in obese mice. *J Lipid Res* 40:1177–1184.

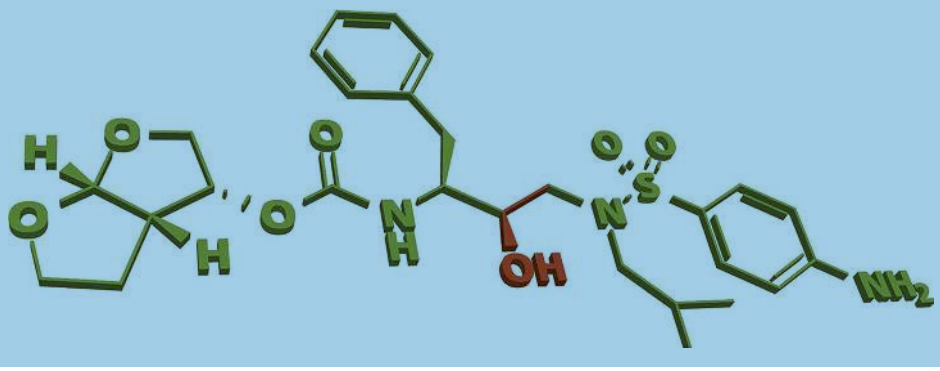
Sohda T, Mizuno K, Imamiya E et al. (1982) Studies on antidiabetic agents. II: synthesis of 5–[4-(1-methylcyclohexylmethoxy)-benzyl] thiazolidine-2,4-dione (ADD-3878) and its derivatives. *Chem Pharm Bull* 30: 3580–3600.

Willson TM, Brown PJ, Sternbach DD & Henke BR (2000) The PPARs: from orphan receptor to drug discovery. *J Med Chem* 43:527–550.

Websites

GPCR action: (a) Signal transduction pathways (2011) *Bozeman Science*. http://www.youtube.com/watch?v=qOVkedxDqQo; (b) G protein signaling (2009) *Garland Science*. http://www.youtube.com/watch?v=V_0EcUr_txk&feature=related; (c) GPCR Claymation (2011) http://www.youtube.com/watch?v=pRBg8Q2pR1s


Taylor & Francis
Taylor & Francis Group



Enzymes as Drug Targets

7

LEARNING OBJECTIVES

- Learn general mechanisms of several classes of enzymes, including proteases, kinases, polymerases, phosphodiesterases, and oxidoreductases.
- Understand the use of enzyme kinetics to characterize modes of enzyme inhibition.
- Recognize when to apply specific medicinal chemistry strategies according to the class of enzyme being targeted.
- Learn examples of drugs that target various enzyme classes.

7.1 INTRODUCTION

Enzymes are a broad class of proteins that catalyze chemical reactions involving the conversion of an enzyme substrate to a product. They are ubiquitous and are involved in numerous physiological processes, such as digestion, blood coagulation, energy production, neurotransmission, and DNA replication. Several of their specific functions include the following:

- Synthesis of amino acids, nucleic acids, fatty acids, coenzymes, vitamins, and proteins
- Degradation of proteins, lipids, carbohydrates, nucleic acids, and small-molecule xenobiotics
- Replication of DNA and RNA
- Activation or inactivation of proteins
- Processing of proteins and nucleic acids by adding or removing specific functional groups, such as carbohydrates, methyl groups, acetate groups, or prenyl groups

Enzymatic reactions range from very simple reactions, such as the hydrolysis of an ester substrate to generate an acid and alcohol catalyzed by an esterase, to complex multicomponent reactions, like those involved in cellular respiration. An example of a multicomponent enzyme is pyruvate dehydrogenase, a complex that consists of multiple copies of three different enzymes along with associated coenzymes and converts pyruvate to acetyl-CoA, which is then used in the citric acid cycle for cellular respiration. Some enzymes are highly specific, acting only on certain molecular substrates, while others are much less selective and, for example, will catalyze reactions at a certain position in the substrate, regardless of the specific structure. Enzymes that exhibit high specificity rely on recognition of the substrate and can often discriminate between very similar substrates on the basis of structural differences quite remote

DOI: 10.1201/ 9781003465607-9

from the active site. Localization of the enzyme in certain organs and tissues, or surrounding conditions, such as pH, can also impact selectivity in a broader sense.

All chemical catalysis is centered on reducing the activation energy of a reaction in order for it to proceed faster. With enzymatic catalysis, a reaction catalyzed by an enzyme (E) involves starting material(s), called substrate (S), and generates an enzyme–substrate complex (E-S). The reaction proceeds through a transition state and eventually forms an enzyme–product complex (E-P), finally releasing the product (P) and the regenerated enzyme (E) (Equation 7.1). The location in the enzyme where this process occurs is called the **active site**.

$$E + S \underset{k_d}{\overset{k_a}{\rightleftharpoons}} E \cdot S \overset{k_{cat}}{\rightleftharpoons} E \cdot P \rightleftharpoons E + P \qquad (7.1)$$

Scientists can describe an enzyme's activity mathematically, using equations that describe kinetics. In Equation 7.1, k_d represents the dissociation constant, which is a measure of the affinity between the enzyme and the substrate. The term k_{cat} is the rate constant for the catalytic step, measured in units of seconds^{-1}. This number reflects how many molecules of substrate one molecule of enzyme can turn over in one second when the enzyme is fully saturated with substrate. Other values used in describing the activity of an enzyme are V_{max}, the maximum rate of the reaction, and K_m (Michaelis–Menten constant), which reflects not only the affinity between the enzyme and substrate but also the rate of the reaction as well.

The role of the enzyme in most cases is to stabilize the transition state, effectively lowering the energy barrier for the reaction to occur (Figure 7.1). In some cases, enzymes increase the strain energy of the substrate, with the same overall result of reducing the activation energy.

There are several common features of enzyme-catalyzed reactions that are relevant to understanding how drugs work when they act as enzyme inhibitors. Enzymes enhance the rate of chemical reactions via **proximity and orientation effects**. Just bringing the reactants together (proximity) and in optimal alignment for reaction (orientation) in the active site provides a considerable rate enhancement compared to the uncatalyzed or nonenzymatic reaction. **Stabilization of the transition state** is another fundamental means by which enzymes accelerate reaction rates.

Each enzyme catalyzes distinct chemical reactions, but many individual enzymes and general classes of enzymes (several described in later sections) rely on some common mechanisms:

- Covalent bonds can be formed between a nucleophilic side chain of an amino acid, such as serine, threonine, tyrosine, cysteine, glutamic acid, aspartic acid, arginine, histidine, or lysine, and an electrophilic site in the substrate. Certain proteases utilize this **covalent catalysis** mechanism.
- Enzymes that utilize **metal ion catalysis** require a metal ion as a co-factor (commonly Zn^{2+}, Mg^{2+}, Fe^{2+}, Fe^{3+}, Cu^{2+}, Co^{2+}, Or Mg^{2+}) for reaction acceleration. The metal ion can play a variety of roles, including binding to substrate and appropriately orienting it for reaction, acting as a reversible oxidant or reductant, and stabilizing electrostatic charges by shielding negative charges on either substrates or developing transition states.
- **Acid–base catalysis** is a common mechanism used by enzymes to lower the activation barriers of certain reactions. Proton transfer and relay between amino acid side chains in the active site, substrates, and products is an efficient mechanism. Enzymes that rely on acid–base catalysis will be sensitive to pH.

Examples of some of these features and mechanisms are illustrated in the hydrolysis of the neurotransmitter acetylcholine, catalyzed by the enzyme acetylcholinesterase (Figure 7.2). The substrate acetylcholine is localized in a deep gorge within the active site of the enzyme, and the carbonyl of the acetate group is attacked by the enzyme's Ser203. This step is initiated by acid–base catalysis, where Glu334 deprotonates the nearby His447, which then removes a proton from Ser203 to allow the serine oxygen to form a covalent bond to the carbonyl in the substrate. The anionic transition state intermediate is stabilized by interaction with backbone amide N-H groups

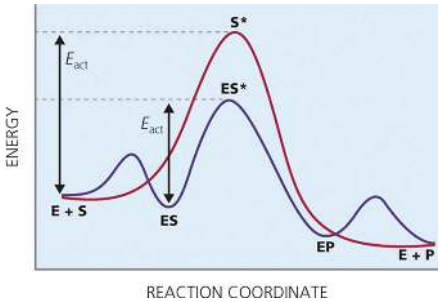


Figure 7.1 General energy diagram of an enzyme-catalyzed reaction vs. an uncatalyzed reaction. In an uncatalyzed reaction (red line), substrate (*S*) is converted to product (*P*) via transition state *S**, requiring a specific energy of activation (E_{act}). In an enzyme-catalyzed reaction (purple line), enzyme (*E*) forms an E–S complex and the transition state (*E–S**) has a lower E_{act} than the uncatalyzed reaction. The E–S complex proceeds to an E–P complex, which is converted to E + P.

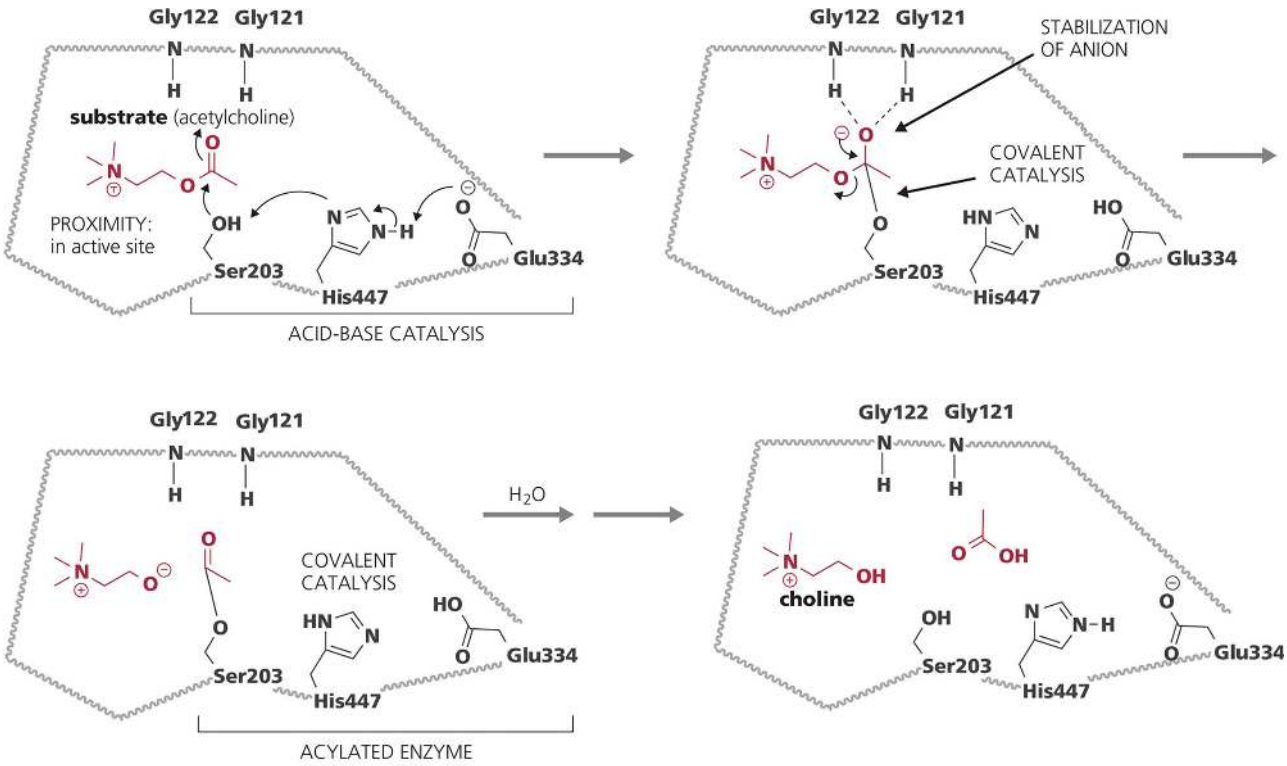


Figure 7.2 Mechanism for hydrolysis of acetylcholine catalyzed by acetylcholinesterase. The substrate (acetylcholine, shown in red) is localized in the active site; acid–base catalysis by Glu334 and His447 activates Ser203, which then forms a covalent bond to the carbonyl of the substrate. The resulting transition state anion is stabilized in the oxyanion hole by backbone NH groups of Gly121 and Gly122. Completion of the reaction occurs with the release of choline, as the oxyanion, and the acetylated enzyme. Further acid–base catalyzed hydrolysis (not shown) generates the protonated choline and activates a water molecule to hydrolyze the acylated enzyme and regenerate it.

of glycine residues (121 and 122) in what is termed the oxyanion hole. The collapse of the tetrahedral intermediate releases the choline product as the oxyanion and acetylated Ser203. Further acid–base catalysis is involved in donating a proton to the choline product, as well as activating a water molecule in a hydrolysis reaction of the acetylated enzyme, which then releases acetate and regenerates the enzyme.

Models to describe the interaction of an enzyme with a substrate have been developed, and since a number of drugs mimic the enzyme substrate, these models have implications for drug design and medicinal chemistry. One of the earliest, the **lock and key model** (Figure 7.3a), was proposed by Emil Fischer in 1894 to explain the high specificity exhibited by enzymes. This model suggests that the enzyme and its substrate are of rigid, precise complementary shapes. Only a substrate (key) of a precise shape and size will fit into the active site (lock) to allow the enzymatic reaction to occur. This model does not account for the dynamic nature of proteins or the concept that enzymes stabilize transition states. In the late 1950s, Daniel Koshland proposed the **induced-fit model** (Figure 7.3b), which addresses some of the shortcomings of the lock and key model. In this model, the enzyme and substrate are both flexible, and as the substrate starts to bind and make specific interactions with the enzyme, the active site moves to accommodate it until the catalytic groups are in the appropriate orientation for reaction, much like a hand fitting into a glove. X-ray crystallography

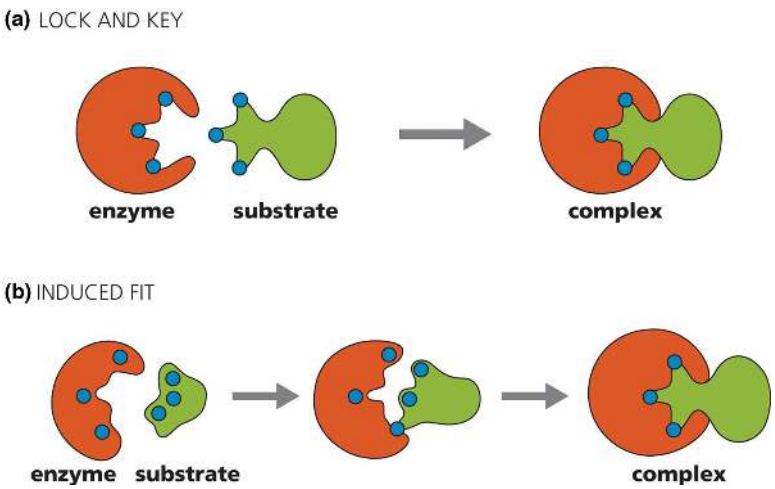


Figure 7.3 Models of enzyme–substrate binding. (a) In the lock and key model, both the substrate and the enzyme are static, and binding occurs without a change in either's conformation. (b) In the induced-fit model, both enzyme and substrate change conformation as binding occurs until a tight fit is achieved. Blue dots represent points of binding interactions.

often shows considerable conformation changes of a protein structure after substrate binding, thereby providing strong support for this model.

Unlike receptors, many enzymes can be isolated and purified and can perform their function as readily in a test tube as they do in the body. Ribonuclease was the first functional enzyme synthesized in a laboratory. Two different groups completed this feat: Bruce Merrifield and Bernd Gutte from Rockefeller University used solid-phase synthesis methods, and a separate research group from Merck Sharpe and Dohme, led by Ralph Hirschmann and Robert Denkewalter, used solution methods. This scientific breakthrough was reported on the front page of *The New York Times* on January 17, 1969. The ability to prepare and purify large quantities of enzymes allows them to succumb to structure elucidation via spectroscopic methods, such as NMR or X-ray crystallography. These types of studies have provided significant advances in understanding how specific enzymes function. In addition, access to large quantities of proteins enables a number of drug discovery and medicinal chemistry strategies, such as high-throughput screening and structure-based drug design.

7.2 EFFECT OF SMALL-MOLECULE MODULATORS ON ENZYMES

While the previous section described the function and effects of enzymes on their natural substrates, the focus in this section is on how small-molecule drugs change the course of the natural, enzyme-catalyzed conversion of substrate to product. As with receptor modulators, enzyme modulators may resemble the natural substrate or may have a significantly different structure. Regardless of the structure, however, most drugs that act on enzymes work through inhibition of the enzyme and are termed **inhibitors**. Drugs that activate an enzyme (**activators**) are far less common than inhibitors.

7.2.1 Modulators Are Classified by How They Interact with an Enzyme

Small-molecule enzyme modulators are classified by two primary characteristics: the location of their binding to the enzyme and the extent of their binding. The extent of binding is related to the reversibility of binding, and enzyme inhibitors (I) can be classified as either reversible or irreversible. **Reversible inhibitors** form a complex (often noncovalent, but not always) with the enzyme (E–I), and both the enzyme and the inhibitor can be recovered intact (**Figure 7.4**). Reversible inhibitors may also be **slow, tight-binding inhibitors** which have such a slow off rate that they can be considered irreversible. The inhibitor constant (or inhibition constant) (K_i) is a measure of the affinity of the inhibitor for the enzyme. As described in Chapter 3, K_i can be directly related to the Gibbs free energy of binding. Another value frequently used to quantify enzyme inhibition, IC_{50}, is the concentration of inhibitor at which 50% of the enzyme activity is inhibited. This value will vary according to the conditions of the assay, such as substrate concentration. In contrast, the K_i is constant, regardless of conditions.

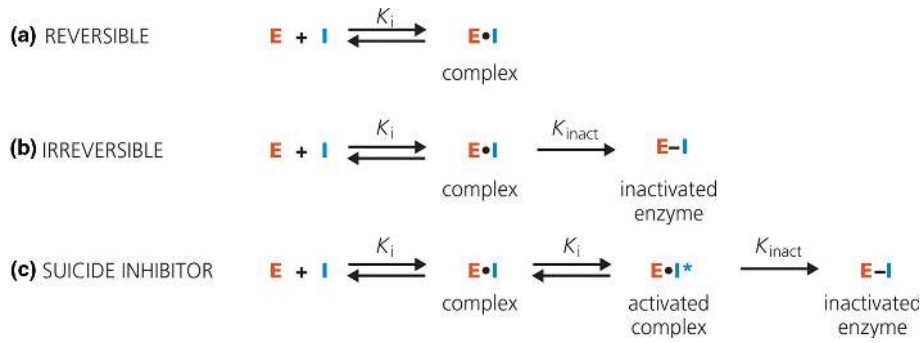


Figure 7.4 Three types of interaction of an enzyme with an inhibitor. (a) Reversible interaction of an enzyme (*E*) with an inhibitor (*I*) gives an enzyme–inhibitor complex E–I in equilibrium with E and I. (b) Irreversible inhibition leads to the complex E–I and then the formation of a permanent covalent bond to generate covalently modified enzyme E–I. (c) A suicide inhibitor leads to the complex E–I and then the formation of an activated complex E–I*, which proceeds to inactivated enzyme E–I.

Classical irreversible inhibitors bind through covalent interactions and do not dissociate from the enzyme. In these cases, K_{inact} is an indication of the rate of formation of the inactivated enzyme. A specific type of irreversible inhibitor is a **suicide inhibitor** or **mechanism-based inactivator**. This is usually a modified version of the substrate, and the initial E–I complex is acted upon by the enzyme to generate an activated complex E–I*, which then covalently modifies the enzyme, resulting in its irreversible inhibition.

The location of binding is another way in which enzyme modulators are characterized (Figure 7.5). **Active site inhibitors** bind at the reaction site and prevent substrate from binding (see Figure 7.5b). **Allosteric modulators** are small molecules that bind to a site remote from the active site and affect enzymatic activity through either inhibition or activation. Allosteric modulators can function through multiple mechanisms; Figure 7.5c illustrates one way in which binding at a remote site can inhibit enzymatic activity.

In addition to location and extent of binding, enzyme inhibitors are also classified with respect to how they influence binding of the natural substrate(s), as shown in Figure 7.6. **Competitive inhibitors** are active-site inhibitors; they bind to the enzyme at the same site as the substrate and compete with it for binding (see Figure 7.4a, 7.5b, and 7.6a). When the enzyme has multiple substrates, the substrate with which the inhibitor is competitive is noted. In these cases, increasing the amount of substrate will decrease the degree of inhibition by the inhibitor.

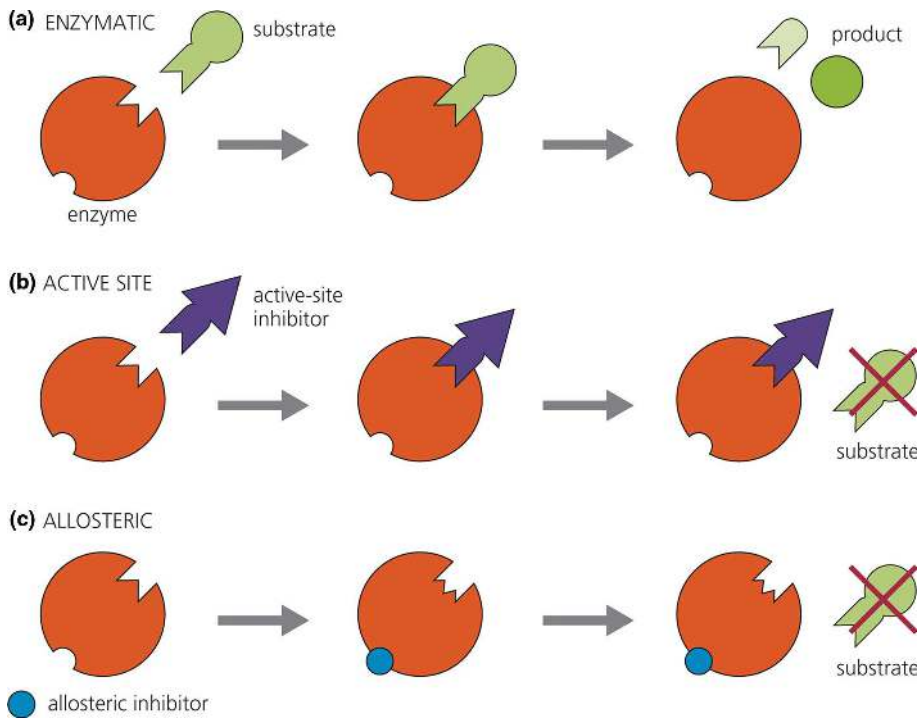


Figure 7.5 Modes of inhibition based on the location of binding. (a) With no inhibition, enzyme and substrate bind, and product is formed. (b) An active site inhibitor binds at the same site as the substrate so that substrate cannot bind and no product is formed. (c) One type of allosteric inhibitor binds at a site remote from the active site and changes the conformation of the active site such that substrate can no longer bind efficiently; no product is formed.

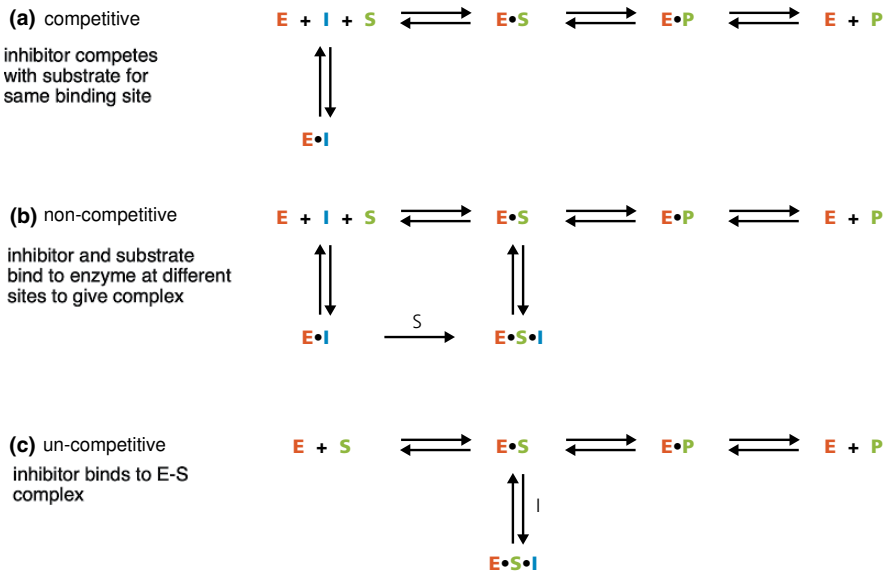


Figure 7.6 Effect of types of inhibition on the reaction of an enzyme with a substrate in the presence of an inhibitor. (a) Competitive inhibitors compete directly with substrate to give an E–I complex. (b) Noncompetitive inhibitors can bind to both the E and the E–S complex, producing an E–I and E–S–I complex, thereby reducing the amount of E–P complex, and eventually product formed. (c) An uncompetitive inhibitor binds to the E–S complex, giving an E–S–I complex, reducing the amount of E–P complex.

Noncompetitive inhibitors are usually allosteric inhibitors but can also bind near the active site without competing with the substrate for binding. They bind to the enzyme and the E–S complex with equal affinity, resulting in a situation where the activity of the enzyme is reduced but substrate binding is not affected (see **Figure 7.6b**).

Uncompetitive inhibitors are allosteric inhibitors that bind to the E–S complex and, on the basis of Le Chatelier's principle, drive the equilibrium away from unbound substrate and enzyme and toward the E–S complex, where binding of the inhibitor to the E–S complex effectively limits conversion to product (see **Figure 7.6c**). Another category of enzyme inhibitors is that of **mixed inhibitors**, where the inhibitor may bind with unequal affinity to both free enzyme (competitively) and E–S complex (uncompetitively).

7.2.2 Effects on Enzyme Kinetics Are Used to Characterize the Mechanism of Small-Molecule Modulators

The different types of inhibition described in the previous section are elucidated by studying the effect of inhibitors on enzyme reaction kinetics. The activity of many enzymes can be described by the Michaelis–Menten equation, which relates initial reaction velocity (v) to the concentration of substrate [S]. In this equation, the Michaelis–Menten constant K_m is substrate concentration [S] at the point of one-half of V_{max} (the maximum rate of the reaction) **(Figure 7.7)**. Velocity can be determined by measuring either the decrease in substrate concentration over time or the increase in product concentration over time. The Michaelis–Menten constant may be derived using a **saturation curve** plotting velocity vs. substrate concentration for a specific enzymatic reaction, as shown in **Figure 7.7b**, where concentration of substrate is increased to the point where increasing concentration no longer results in change in velocity (that is, saturation). A double-reciprocal (Lineweaver–Burk) plot of 1/[S] vs. $1/v$ gives a linear plot that can also be used to obtain the same information, based on the reciprocal of the Michaelis–Menten equation. In this case, the y-intercept is equal to $1/V_{max}$, the x-intercept equals $-1/K_m$, and the slope of the line is K_m/V_{max} (**Figure 7.7c**).

When the Michaelis–Menten equation is used to distinguish between competitive and noncompetitive inhibitors, the control curve or plot when no inhibitor is present is compared to the curve or plot in the presence of a fixed amount of inhibitor. If the inhibitor is competitive, then K_m will increase, since more substrate is required to reach a given velocity of 1/2 V_{max}, but V_{max} will not change, since at a high enough [S], the substrate will displace all the inhibitor and the reaction will proceed as it does without inhibitor present. When graphed, the x-intercept ($-1/K_m$) will change, as will

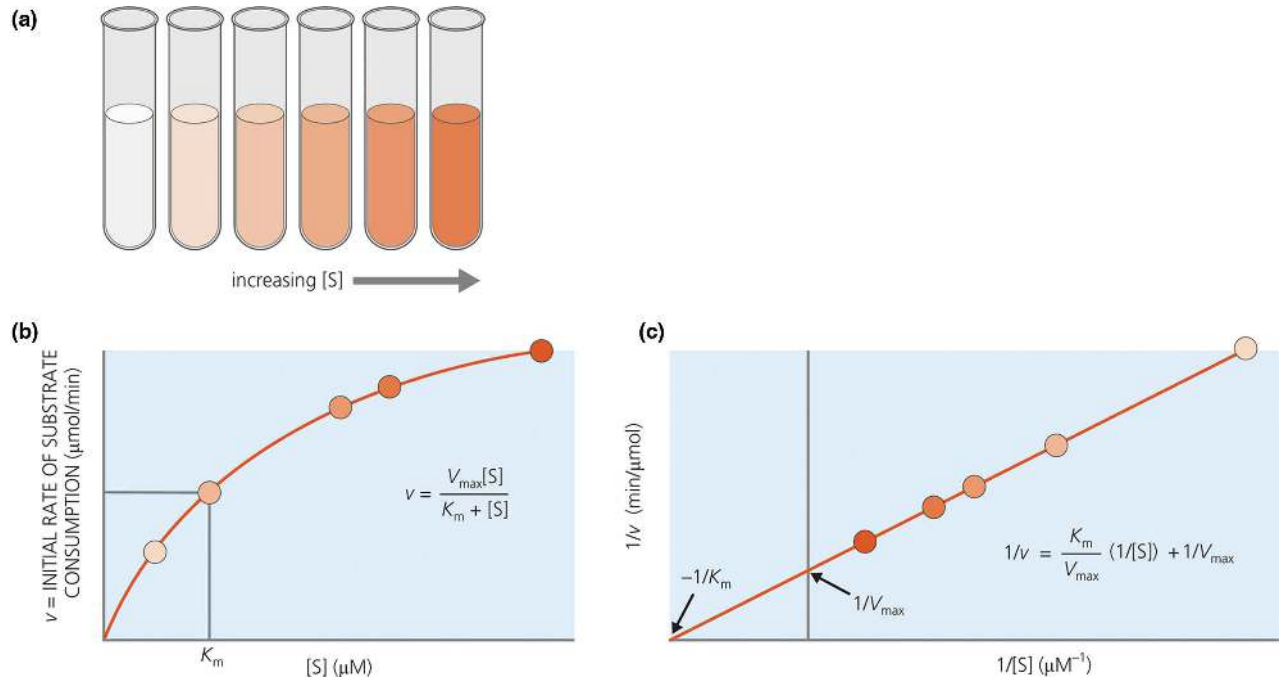


Figure 7.7 Equations and graphs used to characterize enzyme kinetics. (a) Representation of increasing substrate concentration [*S*]. (b) Michaelis–Menten equation and representation of a saturation curve for an enzymatic reaction, plotting substrate concentration [*S*] vs. velocity (*v*). The Michaelis constant K_m is the substrate concentration at half of the maximal reaction velocity. (c) Reciprocal Michaelis–Menten equation and linear (Lineweaver–Burk) plot for an enzymatic reaction. In this plot, the slope = K_m/V_{max}, the *x*-intercept = $-1/K_m$, and the *y*-intercept = $1/V_{max}$.

Source: From Alberts, B.; Bray, D.; Hopkin, K.; et al. (2013). *Essential Cell Biology*, 4th ed. Garland Science.

the slope (K_m/V_{max}) of the line compared to the graph of substrate alone. In contrast, if the inhibitor is noncompetitive, $-1/K_m$ will remain the same but V_{max} can never be reached, so in a Lineweaver–Burk plot, the slope of the line will change due to the change in V_{max}, with the reciprocal $1/V_{max}$ being a larger number with inhibitor present, but with the *x*-intercept ($-1/K_m$) remaining the same. If the inhibitor is uncompetitive, K_m and V_{max} will both decrease, changing both the *x*- and *y*-intercepts, but the slope of the line will remain the same. These examples are illustrated in **Figure 7.8.**

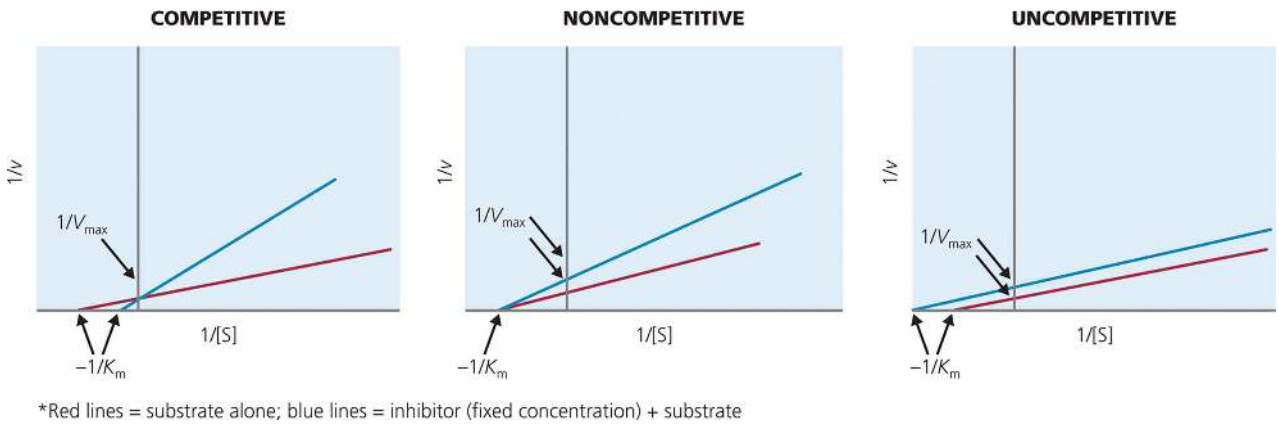


Figure 7.8 Lineweaver–Burk plots describing competitive, noncompetitive, and uncompetitive inhibition. Red lines show increasing concentration of substrate alone; blue lines show fixed concentration of inhibitor + increasing concentration of substrate. For a competitive inhibitor, adding increasing amounts of substrate allows V_{max} to be reached, so it remains the same, but K_m will increase since more substrate is required to reach 1/2 V_{max}. In the presence of a noncompetitive inhibitor, V_{max} cannot be reached regardless of increasing amounts of substrate, so K_m remains the same but the slope and V_{max} change. With an uncompetitive inhibitor, K_m and V_{max} both change, and the slope remains the same.

Molecules that activate an enzyme (activators) could do so by increasing the affinity of the enzyme for the substrate (K_d) or affecting the maximum rate of the enzymatic reaction (V_{max}). A more detailed discussion of enzyme biochemistry is beyond the scope of this chapter; however, characterization of enzyme inhibitors is an essential component of medicinal chemistry and drug discovery programs.

In the context of drug discovery, human diseases can be caused or exacerbated by enzymes that are overactive, underactive, or misactive. The goal, then, is to find a drug that will target only the particular human enzyme that contributes to or causes the disease targeted. Given the similarity in function and structure of various enzyme classes, this is a challenging prospect. In addition to targeting human enzymes, infectious agents, such as bacteria and viruses, may rely on their unique enzymes in order to replicate. Therefore, bacterial and viral enzymes are frequently targeted as a treatment for these infections. In diseases that result from overactive enzymes, as well as with bacterial or viral enzymes, drug discovery efforts focus on inhibitors. Underactive enzymes require activators, and depending on the specific circumstances, diseases caused by misactive enzymes can be treated by either inhibitors or activators.

7.2.3 Enzymes Are Frequent Targets of Drug Action

As described previously, because of their roles in numerous physiological and pathophysiological processes, enzymes are an important class of drug targets and are considered readily druggable (particularly inhibitors) for the following reasons:

- Enzymes are often isolable, well-characterized, and available in significant quantity; therefore, the development, and the implementation, of assays for hit identification and optimization is usually straightforward.
- The characterization of inhibitors and activators can rely on established biochemistry and enzyme kinetics principles.
- Enzyme substrate mimics and transition state analogs, as described in Chapter 4, are often excellent starting points for hit and lead generation.
- Enzymes and enzyme–inhibitor complexes are amenable to structure elucidation via methods such as X-ray crystallography, cryo-EM, and NMR, which affords opportunities for virtual screening, fragment-based screening, and structure-based drug design approaches.

However, there are also challenges in targeting enzymes that are distinct from other target classes, which include the following:

- Specificity for one enzyme over many related enzymes may be difficult to achieve.
- Certain enzyme substrates occur in very high concentrations in the body, thereby requiring extremely potent modulators to compete with the natural substrate.
- Physical properties of enzyme inhibitors are often at odds with druglike properties, such as permeability and metabolic stability.

Some common classes of enzymes that are frequent medicinal chemistry targets, as well as specific examples of drugs that affect a specific enzyme in that class, are listed in **Table 7.1**.

In the following sections, some specific enzyme classes are described along with representative examples of enzyme inhibitor drugs. To date, the only approved drug (in China) that works by enzyme activation is Dorzagliatin, although numerous probe molecules exist, and multiple glucokinase activators have advanced to clinical trials as potential anti-diabetic drugs, although most were terminated after Phase II trials.

Table 7.1 Common Enzyme Drug Targets, the Reaction Catalyzed, Substrates, and Select Drug Examples

			Select Drug Examples		
Class of Enzyme	**Reaction Catalyzed**	**Substrate**	**Drug**	**Mechanism**	**Indication**
Kinase	Phosphorylation	Protein, lipid, or small molecule	Imatinib	Bcr-Abl kinase inhibitor	Cancer
			Tofacitinib	JAK inhibitor	Autoimmune diseases
Glucokinase	Phosphorylation	Glucose	Dorzagliatin	Glucokinase activator	Diabetes
Protease	Hydrolysis of peptide bond	Peptide or protein	Captopril	ACE inhibitor	Cardiovascular disease
			Indinavir	HIV protease inhibitor	HIV infection
			Penicillin	Penicillin-binding protein inhibitor	Bacterial infection
Polymerase or Reverse Transcriptase	Formation of phosphodiester bond	DNA or RNA	AZT	HIV reverse transcriptase inhibitor	HIV infection
			Sofosbuvir	HCV RNA polymerase inhibitor	Hepatitis C infection
Esterase and Phosphodiesterase	Hydrolysis of carboxylic acid and phosphate esters	Small-molecule second messengers, lipids, proteins, nucleic acids	Orlistat	Lipase inhibitor	Obesity
			Donepezil	Acetylcholinesterase inhibitor	Alzheimer's disease
			Sildenafil	Phosphodiesterase inhibitor	Erectile dysfunction
Oxidoreductase	Oxidation and reduction	Small-molecule building blocks for endogenous molecules, proteins, lipids	Atorvastatin	HMG-CoA reductase inhibitor	Cardiovascular risk, cholesterol lowering
			Methotrexate	Dihydrofolate reductase inhibitor	Cancer; autoimmune diseases
Topoisomerase	Cleavage and formation of nucleotide bond	DNA and RNA	Ciprofloxacin	DNA gyrase inhibitors	Bacterial infection
			Camptothecins, etoposide	Topoisomerase II inhibitors	Cancer

7.3 KINASES AND KINASE INHIBITOR DRUGS

Kinases are enzymes that catalyze phosphorylation of a substrate. The number of distinct human kinases is estimated to be >500, and these play essential roles in numerous physiological events. Many are receptor kinases; that is, they span the plasma membrane and contain an extracellular binding site and an intracellular catalytic site. There is also a smaller family of non-receptor intracellular kinases. Many receptor kinases are an integral part of a signaling cascade and regulate cell growth, proliferation, and differentiation, making them frequent targets for cancer drugs. In addition to cancer, other disorders that are marked by uncontrolled cell proliferation, such as inflammatory diseases and autoimmune disorders, may also be treated by kinase inhibitors. Figure 7.9 illustrates a receptor tyrosine kinase. Typically, receptor kinases are composed of an extracellular ligand–binding domain, a membrane-spanning region, and an intracellular domain containing the phosphorylating enzymatic activity. Upon binding of a ligand, two receptors dimerize, resulting in the activation by autophosphorylation (tyrosine in figure, but can also be serine or threonine) using ATP as the phosphate source. The active, phosphorylated form then phosphorylates substrates, activating signaling pathways and leading to a downstream effect, such as cell proliferation.

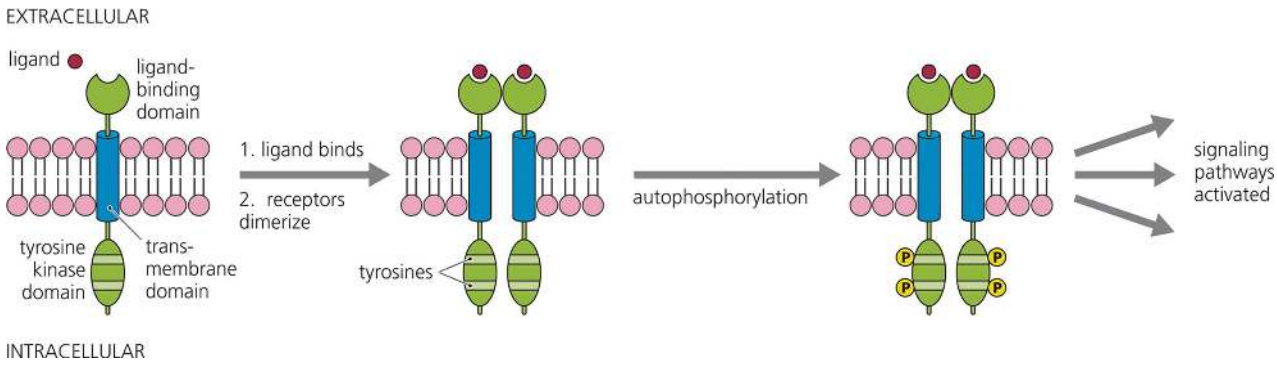


Figure 7.9 Structure and function of a receptor tyrosine kinase. A receptor tyrosine kinase has an extracellular ligand–binding domain, a transmembrane region, and an intracellular tyrosine kinase domain. Upon ligand binding, two receptors dimerize, resulting in activation by self-phosphorylation. The active phosphorylated form then acts as a kinase, phosphorylating substrates and resulting in downstream effects by activating signaling pathways.

The kinase substrate may be a protein or a bioactive small molecule, such as a nucleic acid, a carbohydrate, or a lipid, that contains a free hydroxyl group. When the substrate is a protein, phosphorylation usually occurs on a serine, threonine, or tyrosine hydroxyl group; these kinases are termed serine/threonine kinases or tyrosine kinases, respectively. The reaction involves the transfer of a phosphate group from a nucleoside triphosphate, usually adenosine triphosphate (ATP), but sometimes guanosine triphosphate (GTP), to a substrate containing a hydroxyl group. The products of the reaction are the phosphorylated substrate and the nucleoside diphosphate, ADP (or GDP), as shown in Figure 7.10.

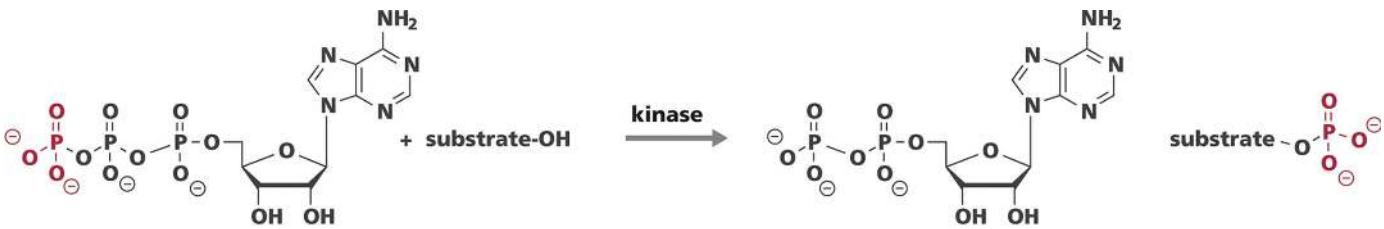


Figure 7.10 Phosphorylation of the hydroxyl of a substrate catalyzed by a kinase. A substrate (either a small molecule containing a hydroxyl group or a protein containing a serine, threonine, or tyrosine) is phosphorylated. In this example, ATP is a second substrate and donates the phosphate group (shown in red), producing ADP and phosphorylated substrate.

Kinase-mediated phosphorylation proceeds through a complex mechanism. A condensed version is illustrated in Figure 7.11, using an ATP-dependent kinase as an example. It involves complexation of the triphosphate of ATP by two magnesium ions and also to amino acids in the protein substrate. This coordination (metal ion catalysis) appropriately orients and activates the terminal phosphate toward nucleophilic attack. The substrate hydroxyl group is deprotonated by an amino acid (*Asp* in figure) of the kinase via acid–base catalysis. Phosphate transfer from ATP to the substrate generates a pentavalent transition state, which leads to the newly formed ADP molecule and phosphorylated substrate. When the substrate is a protein or peptide, kinases often recognize up to four amino acid residues to either side of the phosphorylation site—meaning that phosphorylation will not occur unless this specific consensus sequence is present. However, while the kinase may accept a smaller peptide substrate containing the consensus sequence, the entire protein substrate contributes to binding and recognition, and enzymatic reactions are most efficient on the natural substrate.

The structure of the active site of kinases is highly conserved, albeit very flexible, and is located in the kinase domain (also called the catalytic domain) of the much larger protein. A model of a typical kinase bound to ADP, based on an X-ray structure

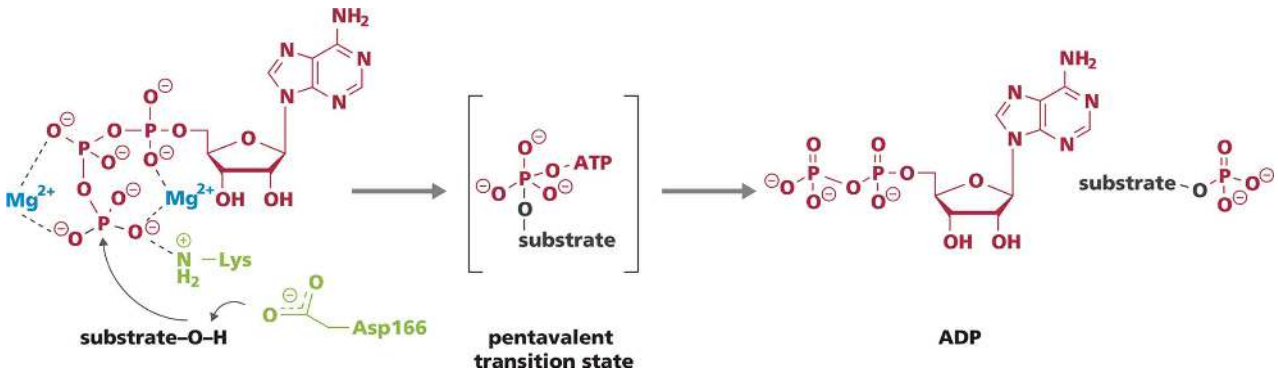


Figure 7.11 Condensed reaction mechanism of ATP-dependent kinase phosphorylation of a substrate. Metal and amino acid coordination to the phosphate groups in the active site and acid–base catalysis enable a reaction between the substrate and ATP to generate a pentavalent transition state. The completion of the reaction generates phosphorylated substrate and ADP. *Note:* red = ATP, ADP, transferred phosphate; blue = Mg^{2+}; green = protein kinase residues; black = substrate.

of aurora kinase, is shown in **Figure 7.12**. The ATP binding site is buried deep inside the protein, in a cleft between two lobes of the protein. One lobe consists of the N-terminal region, composed of β-sheets, a glycine-rich region, and the αC-helix, and the other is the α-helix-rich C-terminal region. The area between the two lobes is termed the **hinge region**, and the adenine base of ATP makes a series of H-bonds to this hinge region of the kinase, while the furanose and triphosphate of ATP sit in a hydrophilic pocket that extends toward the substrate binding site. A **gatekeeper residue** (not shown) sits close to the hinge region/ATP binding site. An activation loop (A-loop), beginning with the conserved amino acid triad DFG (Asp-Phe-Gly), sits near the substrate binding site and acts as a door for the substrate, with the aspartate residue coordinating with magnesium ions that stabilize the triphosphate during catalysis. This loop is highly flexible, with multiple conformations of it observed during structural

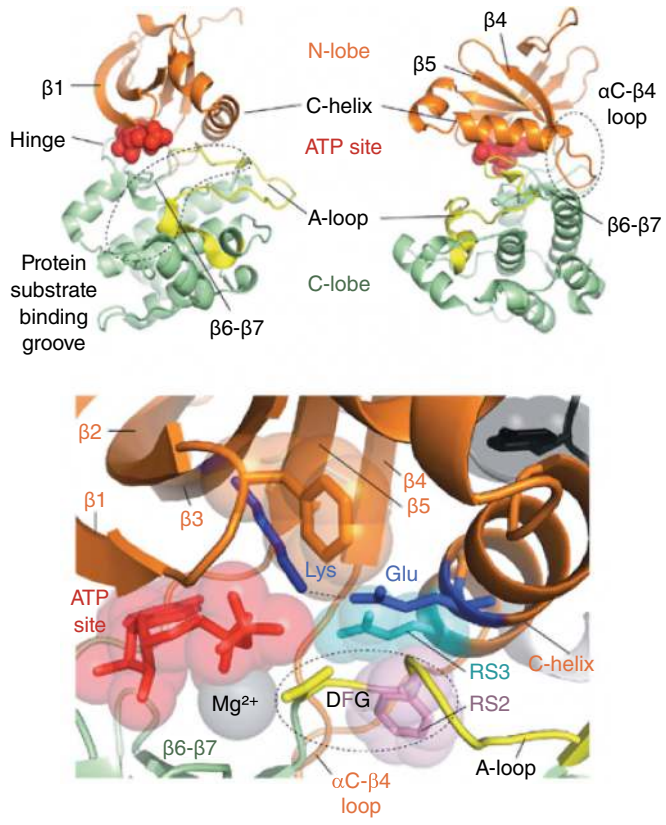


Figure 7.12 Model of a kinase domain based on the structure of aurora kinase, and some of its key features. *Top*, two views of the kinase. The N-terminal lobe and C-terminal lobes are separated by a hinge region, where the ATP binds and catalysis occurs. The regulatory A-loop and C-helix are also labeled. *Bottom*, close-up view of kinase active site with ADP (red) modeled into ATP site, bound to Mg^{2+} (gray). Highlighted is the A-loop (yellow), which contains the RS2 residue Phe (pink) and the RS3 residue, Glu (blue), from the C-helix.

Source: From Arter, C.; Trask, L.; Ward, S.; Yeoh, S.; and Bayliss, R. (2022). *J. Biol. Chem.* 298(8), 102,247.

analyses. Residues in this loop are often phosphorylated as part of a regulation mechanism, and this modification can substantially increase enzymatic efficiency. The R spine (regulatory spine) consists of four hydrophobic amino acids, labeled RS1–RS4, that are in different locations of the kinase domain but are aligned and stacked to create a "spine" that spans the catalytic cleft; this alignment is essential for catalytic activity and is characteristic of an active conformation of the kinase. Two residues of particular interest for drug design are RS2, which is within the A-loop and the central residue of the conserved DFG motif, and RS3, from the C-helix. These residues are displaced when the enzyme is in an inactive conformation and the "spine" is disrupted, leaving RS2 and RS3 sites unoccupied.

In addition to phosphorylation, kinases are regulated by other mechanisms. In many cases, an effector protein is required to bind, often at a site remote from the active site, to the kinase to render it capable of phosphorylation. Mechanisms that control kinase activity include other modifications of the protein, such as acylation or isoprenylation and localization. These regulatory elements can also be targeted by small molecules to modify the activity of the kinase.

Recent advances in solving high-resolution structures of kinases bound to inhibitors and substrates, as well as the continued interest in kinase inhibitors as drug targets and their therapeutic success, have led to an explosion in the medicinal chemistry of small-molecule inhibitors. Traditional high-throughput screening can be used for hit identification. However, focused screening approaches, which test a small library of hundreds to thousands of molecules that are either designed or known to target the kinase ATP binding domain, are more common. The similarity across kinase binding domains virtually assures that hits will be identified; however, optimizing selectivity for the desired kinase is often a significant effort. Due to the availability of X-ray structures and/or homology models of kinases, fragment-based screening and virtual screening strategies are often also successful. The application of transition state mimetic strategies for kinase inhibition has not been particularly useful due to the fact that mimicking a highly charged pentavalent phosphate transition state (see **Figure 7.11**) is not readily achieved with standard organic small molecules.

7.3.1 Kinase Inhibitors Work through Multiple Binding Modes and Are Effective Anti-Cancer Drugs and Anti-Inflammatory Agents

Since the approval of the first kinase inhibitor, imatinib, in 2001, there are now >80 approved drugs that act on over 20 different kinases. Of these drugs, the majority are inhibitors of receptor tyrosine kinases and bind at or near the active site. Allosteric modulators and activators are less common, as are inhibitors of kinases that act on small-molecule substrates. However, significant advances in the field have led to the approval of the first small-molecule allosteric modulators, and some activators have advanced to clinical trials. Inhibitors of kinases that phosphorylate lipids (such as sphingosine) or nucleosides have also been reported. While most kinase inhibitors are used to treat cancers that either overexpress certain growth factor receptor kinases or have mutant forms of those kinases, currently about 30% are used to treat other conditions, such as rheumatoid arthritis and psoriasis. Those drugs primarily inhibit Janus kinases (JAKs), which are cytoplasmic non-receptor kinases.

With the availability of high-resolution co-structures of small molecule kinase inhibitors bound to their target, there has been a greater appreciation that multiple binding modes can lead to potent inhibition. Furthermore, different conformations of the kinase itself are prone to different modes of inhibitor binding. In an effort to characterize these, kinase inhibitors were classified according to "types" (e.g., types I–VI), and as additional structural information has become available, subclassifications within each type have been proposed. It is likely that even further refinement and/or expansions of these taxonomies will be necessary. The following descriptions are based on those proposed by Roskowski (in 2016, 2021, 2024) and by Bayliss (2022). Of note, the same kinase inhibitor may bind via different modes to different kinases.

The most common class of drugs targeting kinases are competitive inhibitors that make interactions with the hinge region within the ATP site but are differentiated by whether they bind to an active or inactive form of the enzyme and also by which specific binding sites they occupy. The kinase conformation with **DFG-in**, the αC helix

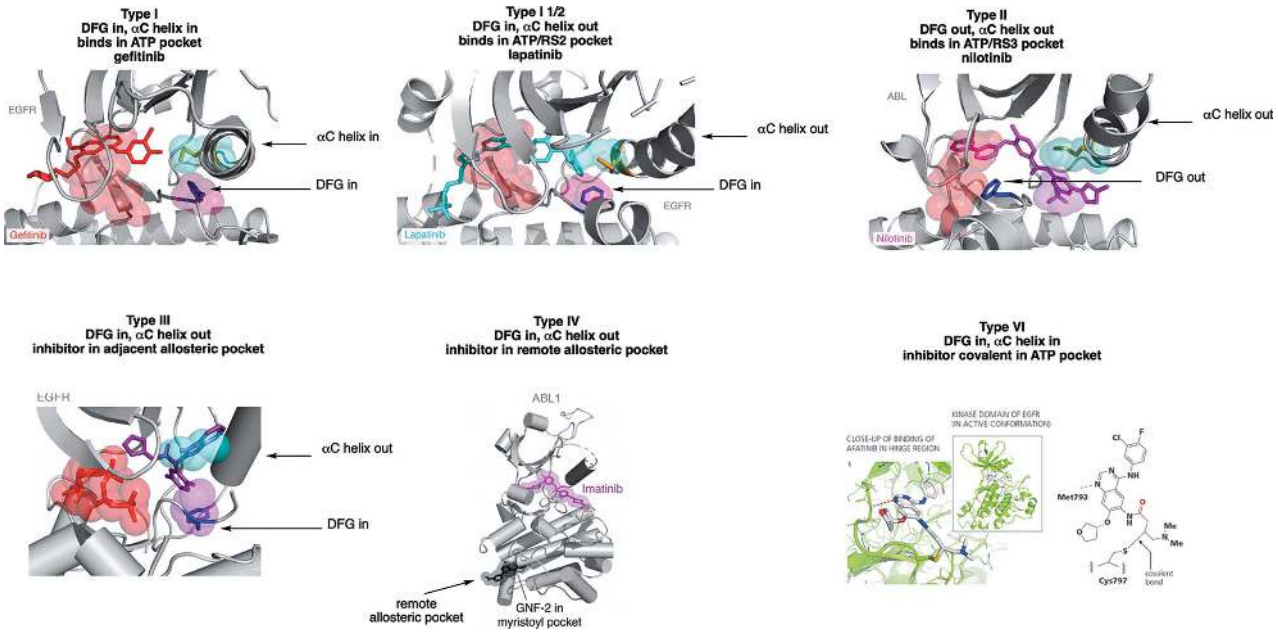


Figure 7.13 Detailed structures of binding modes of different types of kinase inhibitors (ATP site in red, RS2 site in purple, and RS3 site in blue). The type I inhibitor gefitinib (red) binds in the ATP pocket; type I 1/2 inhibitor lapatinib (teal) occupies the ATP pocket, extending into the RS3 pocket; type II inhibitor nilotinib (purple) sits in the ATP pocket, extending into the RS2 pocket. A type III allosteric inhibitor (purple) binds only in the RS3 pocket, and a type IV allosteric inhibitor (GNF-2) binds in the remote myristoyl pocket of Abl1 kinase with the ATP site occupied (in this case by imatinib). Finally, a type VI covalent inhibitor, afatinib, is shown, with a bond to Cys797 in the ATP pocket. The chemical structures of gefitinib (type I) and lapatinib (type I 1/2) and nilotinib (type II) are shown in **Figure 7.14**; afatinib in **Figure 7.16**.

Source: Types I, I 1/2, II, III, and IV from Arter, C.; Trask, L.; Ward, S.; Yeoh, S.; and Bayliss, R. (2022). *J. Biol. Chem.*, 298, 102,247. Type VI from Solca, F.; Dahl, G.; Zoephel, A.; et al. (2012). *J Pharmacol Exp Ther* 343:342–350. With permission from the American Society for Pharmacology and Experimental Therapeutics.

"in," and therefore the R-spine stacked, is considered an active conformation. In this conformation, only the ATP site is accessible to an inhibitor, since the Phe of DFG occupies its RS2 pocket and a hydrophobic side chain of the αC helix occupies the RS3 pocket. A number of clinically used kinase inhibitors, such as gefitinib, bind at the hinge region in the active kinase conformation and make similar H-bonding contacts as the adenine moiety of ATP. These are classified as type I inhibitors (**Figure 7.13**).

Type I 1/2 inhibitors also bind in the ATP site, but when the kinase is in a DFG-in, αC helix "out," inactive conformation, and they typically also occupy the now-empty RS3 pocket. Laptanib is an example of a type I 1/2 inhibitor (**Figure 7.13**). This class has been further classified according to whether or not they bind to additional clefts near the ATP site: type I 1/2A inhibitors also extend into a site past the gatekeeper residue, while Type I 1/2B inhibitors do not.

Type II inhibitors are also active site–competitive inhibitors but bind to the inactive **DFG-out** conformation and occupy the RS2 pocket that is vacated by the Phe residue when the A-loop moves to an inactive conformation. In **Figure 7.13**, nilotinib is an example of a type II inhibitor. As with the subdivisions within the type I 1/2 inhibitors, type IIA inhibitors extend into the back cleft, while the B subtype does not.

Allosteric inhibitors include type III, which bind adjacent to the ATP site, and type IV, which bind remote from the ATP binding pocket. Type V inhibitors are bivalent, although no approved drugs bind in this fashion. Type VI are covalent inhibitors, which form a bond with a cysteine residue in the vicinity of the ATP site. **Figure 7.13** shows afatinib's covalent binding to EGFR kinase domain. **Table 7.2** shows representative examples of kinase inhibitors of various types.

Structures of representative competitive inhibitors, which will be discussed further in Chapter 10, are shown in **Figure 7.14**. Type I inhibitors erlotinib and gefitinib inhibit epidermal growth factor receptor (EGFR) kinase activity and therefore prevent receptor signaling in cells and are approved to treat certain cancers that overexpress EGFR. Activities against other kinases have also been reported.

Table 7.2 Summary of Kinase Inhibitor Types, Based on Kinase Conformation and Binding Location

Type	Kinase Conformation Features	Pocket Occupied	Examples (Kinase Target)
Type I: competitive	Active; DFG_{in}/**α**C $helix_{in}$	ATP	Gefitinib (EGFR)
Type I 1/2: competitive	Inactive; DFG_{in}/**α**C $helix_{out}$	Type A: ATP and RS3	A: Lapatinib (EGFR); dasatanib (BCR-ABL)
		Type B: ATP	B: Erlotinib (EGFR); sunitinib (CDK2)
Type II: competitive	Inactive; DFG_{out}/**α**C $helix_{in\ or\ out}$	Type A: ATP and RS2	A: imatinib (ABL, Kit), nilotinib
		Type B: ATP	B: Sunitinib (VEGFR, Kit)
Type III: allosteric	Inactive; DFG_{in}/**α**C $helix_{out}$	Adjacent to ATP and RS3	Trametinib (MEK), cobimetinib (MEK)
Type IV: allosteric		Remote from ATP	Asciminib (ABL1)
Type V: bivalent		ATP and remote pocket	None approved
Type VI: covalent, competitive		ATP; covalent bond with Cys	Afatinib (EGFR), ibrutinib

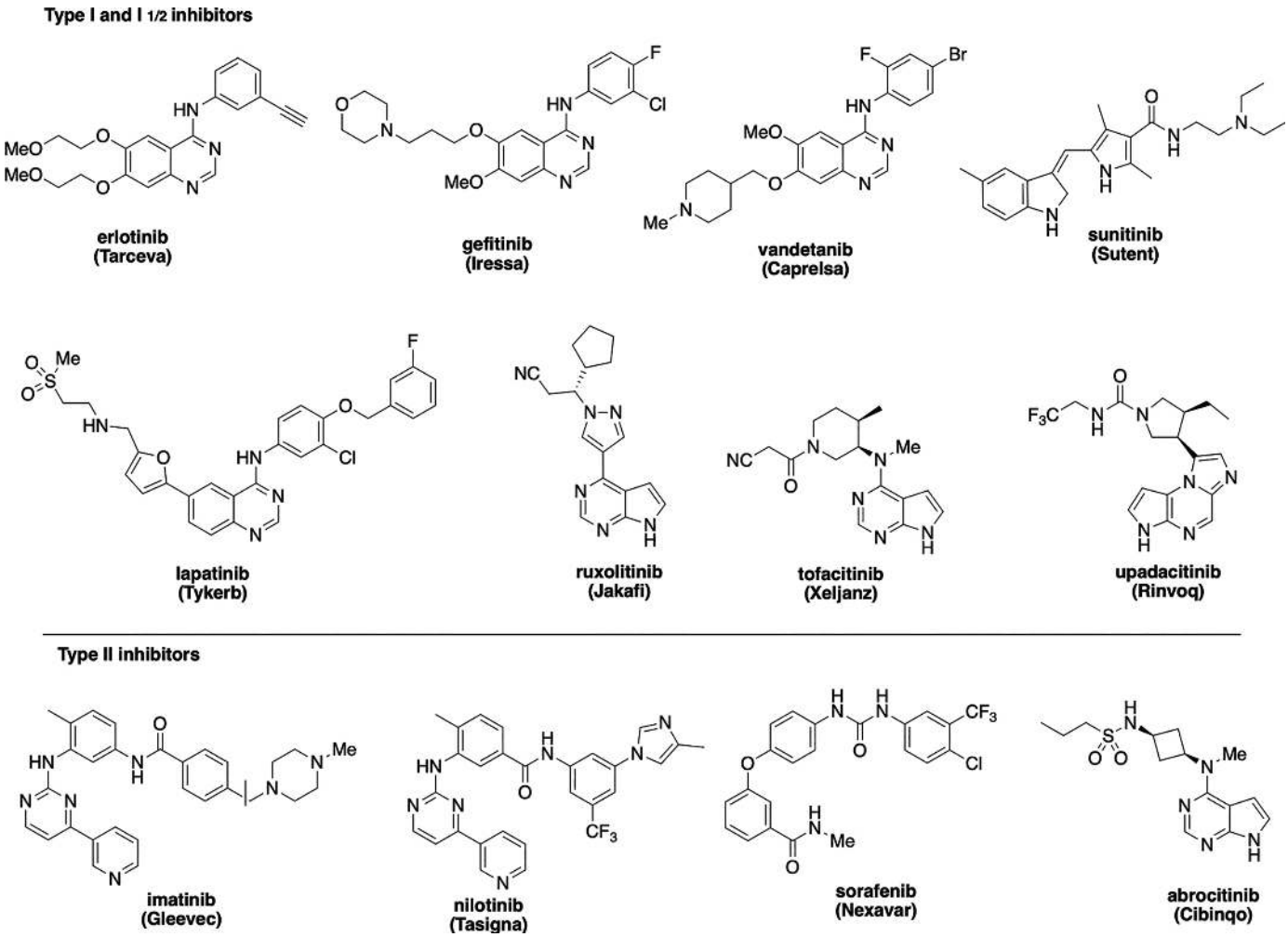


Figure 7.14 Structures of types I, I 1/2, and II kinase inhibitors. Type I inhibitors erlotinib and gefitinib inhibit epidermal growth factor receptor (EGFR) kinase activity and are approved to treat certain cancers that overexpress EGFR. Vandetanib and sunitinib inhibit a number of receptor tyrosine kinases, including platelet-derived growth factor (PDGF) and vascular endothelial growth factor (VEGF). Lapatinib is a type I 1/2 inhibitor used to treat solid tumors that target EGFRs. Ruxolitinib, tofacitinib, and upadacitinib are all type I JAK inhibitors and are used to treat inflammatory diseases. Type II inhibitors imatinib and sorafenib target multiple kinases, including VEGFR, PDGFR, and RAF, and are approved to treat certain cancers. Abrocitinib is a JAK inhibitor approved to treat atopic dermatitis.

Vandetanib and sunitinib inhibit a number of receptor tyrosine kinases, including platelet-derived growth factor (PDGF) and vascular endothelial growth factor (VEGF). Lapatinib is a type I 1/2 inhibitor that targets several EGFR's and is used to treat solid tumors. The Janus kinases (JAKs) are tyrosine kinases and are part of the JAK-STAT signaling pathway, which is involved in the regulation of the immune system, with some mutant forms involved in certain types of cancer. Type 1 JAK inhibitors include ruxolitinib, tofacitinib, and upadacitinib, which are approved to treat inflammatory diseases, such as rheumatoid arthritis, psoriatic arthritis, and atopic dermatitis (eczema), among other conditions.

Inhibitors that bind to the ATP site have the potential to be non- or only partially selective since they bind to a site that is highly conserved across all kinases. It is generally believed that allosteric modulators have the potential to be more selective than those acting at or near the active site, since these sites may not be as widely conserved; however, they may also be prone to rapidly generating mutations and resistance.

Trametinib, a type III inhibitor, targets mitogen-activated protein kinase kinase (MEK) and is approved for use in certain cancers (**Figure 7.15a**). It was optimized from a hit identified in a high-throughput screen. A similar type III inhibitor is PD318088 (also a MEK inhibitor). Common structural features of some allosteric MEK inhibitors that bind to the ATP-bound form of the enzyme include a diaryl amine (see, for example, deucravacitinib and PD318088), often with a para-halogen substituent on

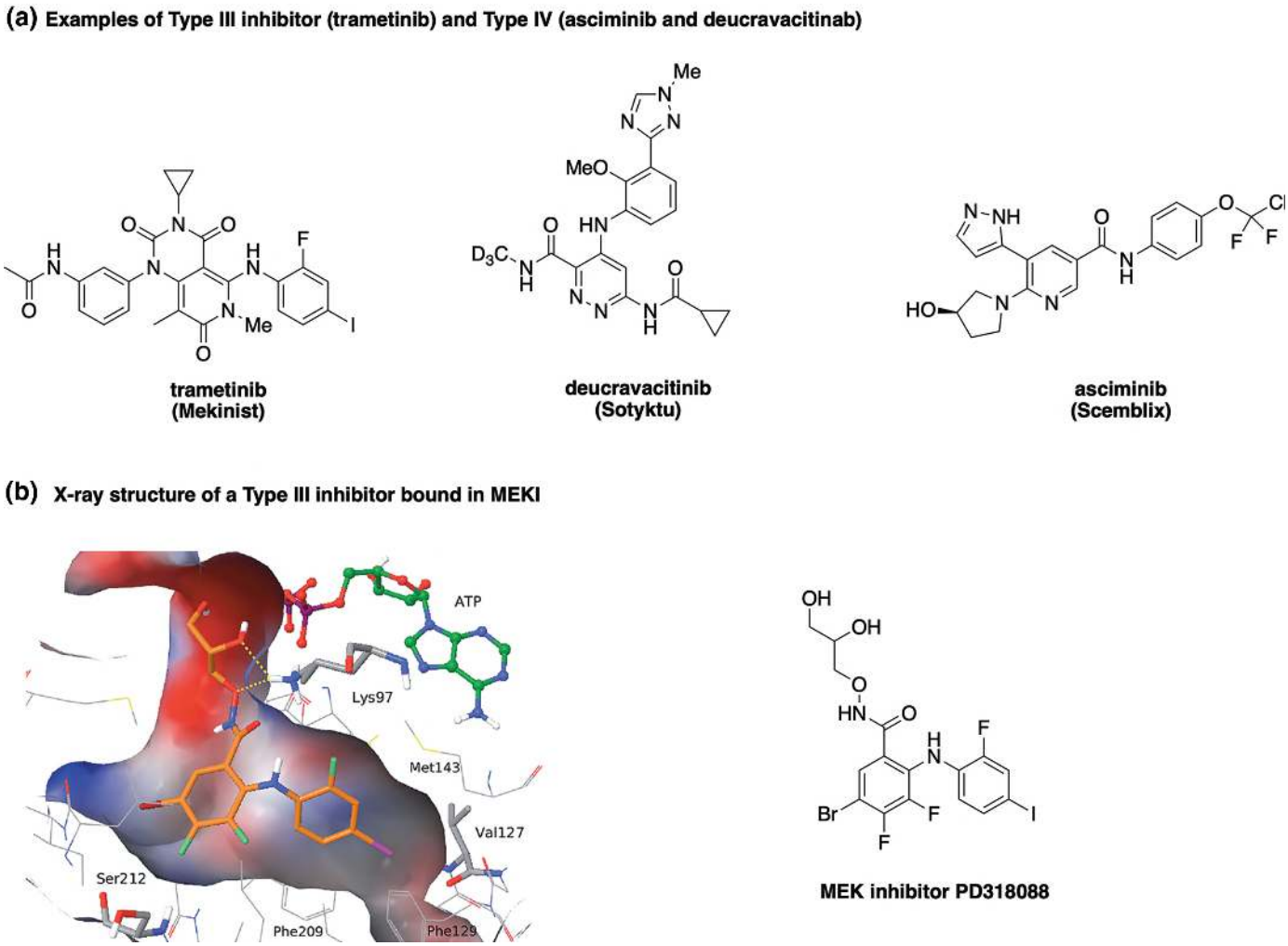


Figure 7.15 Structures of type III and type IV allosteric inhibitors, and X-ray structure of a type III inhibitor bound near the active site of MEK. (a) Structures of Types III and IV allosteric kinase inhibitors. Trametinib is a type III inhibitor and binds adjacent to the ATP pocket, while asciminib, like deucravacitinib, is a type IV inhibitor specific to the myristoyl pocket of ABL kinase, and deucravacitinib is selective for TYK2 kinase. (b) Prototypical type III MEK inhibitor PD318088 (orange) structure and X-ray bound at allosteric site adjacent to ATP binding site.

Source: Adapted from Heald, R. A.; Jackson, P.; Savy, P.; et al. (2012). *J Med Chem* 55:4,594–4,604. With permission from the American Chemical Society.

one of the aryl groups. The X-ray co-structure of MEK bound to PD318088 shows binding to the kinase when ATP is also bound, at a site adjacent to the ATP site (**Figure 7.15b**). The trisubstituted aryl ring makes a number of van der Waals interactions in a hydrophobic pocket. Other important interactions include one between the para-iodine of the trisubstituted aryl ring and a backbone carbonyl of Val127 in the hinge region, and a hydrogen bond between one of the fluorine substituents on the penta-substituted aryl ring and Ser212. There are also interactions between the polar hydroxyls and the ATP terminal phosphate and Lys97. It is suggested that binding of these types of molecules stabilizes an inactive conformation of the kinase. As predicted on the basis of their binding to an allosteric site, these compounds are highly selective for MEK1 and MEK2, and they do not inhibit other kinases or members of the same kinase family. This selectivity is likely because the inhibitors bind to a site that has little sequence homology to other kinases.

Type IV inhibitors, such as deucravacitinib and asciminib, bind remotely from the ATP binding site. Deucravacitinib is a selective inhibitor of TYK2 (a Janus kinase) and is used to treat psoriasis. It binds about 46 Å away from the ATP pocket. In the case of Abl kinase, a myristoyl group is involved in autoregulation of its enzymatic activity, and asciminib binds in that myristoyl pocket. This drug is selective for Abl kinase and is used to treat chronic myelogenous leukemia that is resistant to earlier inhibitors due to mutations in the ATP binding site. Using asciminib in combination with ATP-competitive inhibitors, inhibition against resistant mutations is restored. Asciminib was discovered by scientists at Novartis using a fragment-based screening approach specifically designed to identify inhibitors binding in the myristoyl pocket (GNF-2 was an early analog in this program).

While there are no type V bivalent inhibitors yet approved, there are now a number of type VI covalent inhibitors. The therapeutic success of targeting kinases mutated in cancers has been tempered by the development of resistance to the drugs. This is due to secondary mutations that form after months of treatment. In an effort to combat these secondary mutations, one area of interest has been in developing covalent inhibitors. These type VI inhibitors were designed to bind in the ATP site, like many of the types I and II inhibitors, but contain an electrophile, usually an acrylamide, that reacts with a noncatalytic cysteine via a Michael reaction. Several examples of covalent kinase inhibitors are shown in **Figure 7.16**, along with a common mechanism of covalent bond formation.

This type of inhibition results in a longer residence time of the inhibitor, which usually results in improved efficacy and can also limit the emergence of certain resistant mutations. **Figure 7.13** shows an X-ray structure of afatinib, the first covalent kinase inhibitor approved by the FDA, bound to the ATP site of EGFR kinase. Afatinib is a close structural analog of gefitinib, and its quinazoline core makes a hydrogen bond to the hinge region of the protein. The protein's nearby cysteine reacts with the crotonamide to make a covalent bond. Ibrutinib is approved for B-cell malignancies and targets Bruton's tyrosine kinase (BTK), which is essential for B-cell development. Dacomitinib and neratinib are examples of two others approved to treat certain cancers, and many other compounds that act by way of covalent modification of the ATP site have been developed or are in active development.

This strategy to inhibit kinases has become more prevalent and relies on the identification of potent noncovalent binders and then, based on structural information, incorporates reactive functionality at an appropriate site to engage a nearby nucleophile (usually sulfur). However, some liabilities of compounds such as these, including inhibition of wild-type as well as oncogenic mutated proteins, have been reported, and further study of optimal reactivity toward covalent bond formation is an active area of research.

Drugs that work by activating a kinase are rare, although several have reached clinical trials, and one has been approved for use in China. One target is glucokinase, a hexokinase that phosphorylates glucose to generate glucose-6-phosphate, induces glucose metabolism, and reduces glucose levels. Drugs that activate glucokinase can be used to treat diabetes. Dorzagliatin (**Figure 7.17**) was developed by Hua Medicine and approved in China in 2022 for the treatment of type 2 diabetes. Several other glucokinase activators, such as cadesigliatin, are in clinical trials.

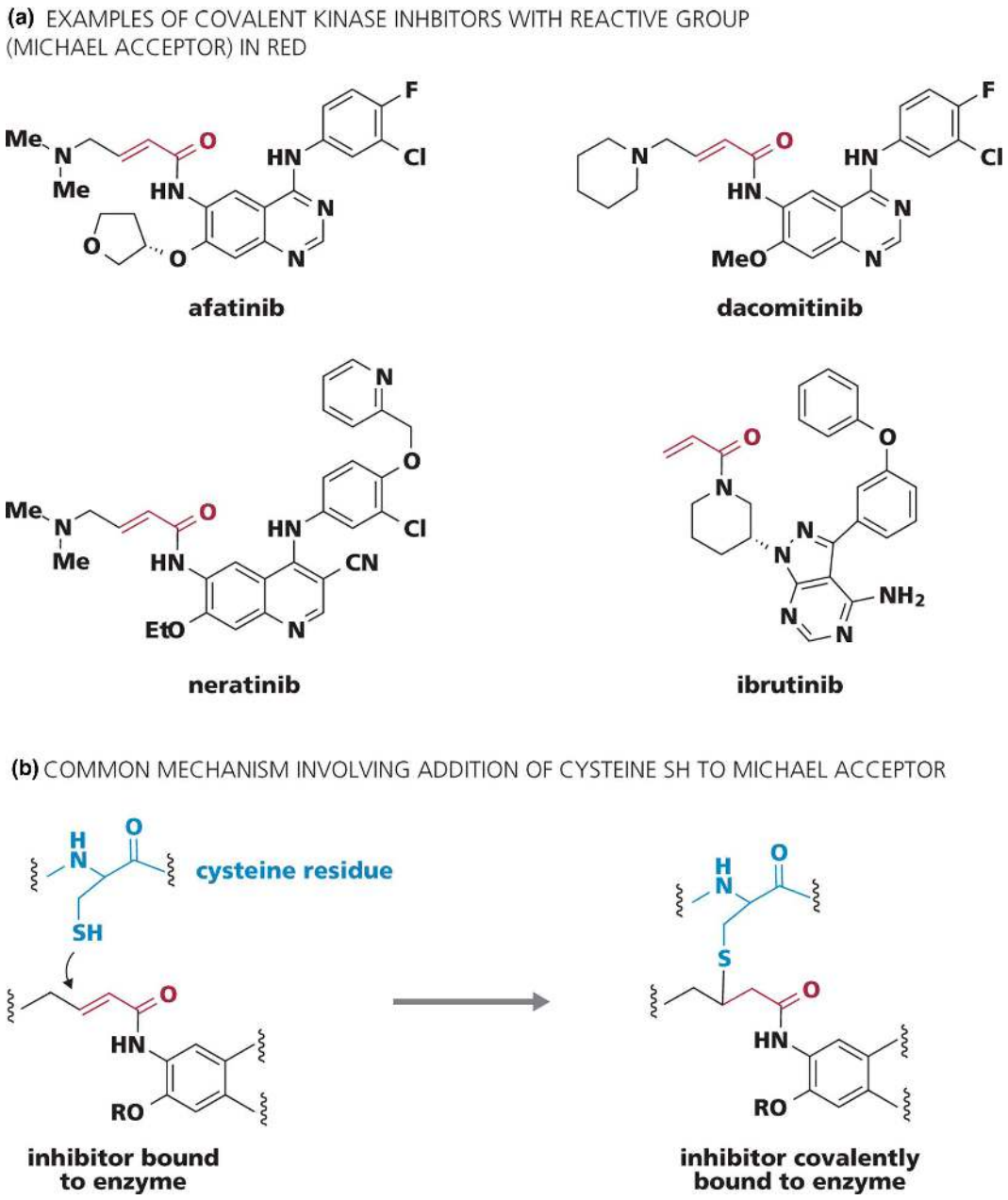


Figure 7.16 **Covalent kinase inhibitors.** (a) Structures of afatinib, dacomitinib, neratinib, and ibrutinib; all are approved to treat certain cancers. (b) Common mechanism of inhibition involving the addition of a cysteine thiol (blue) near the inhibitor binding site to a Michael acceptor (red) to give a covalent adduct.

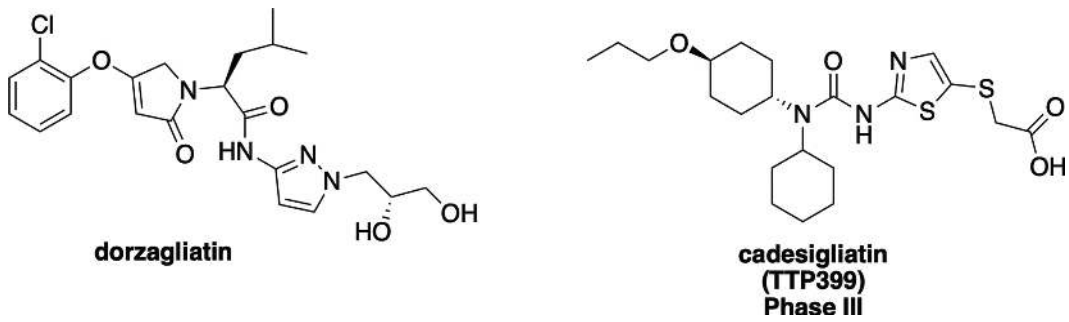


Figure 7.17 **Structures of glucokinase activators dorzagliatin and cadesigliatin.**

The kinase inhibitors are excellent examples of combining screening, particularly focused screening, with structure-based design in order to optimize drug candidates. The availability of numerous X-ray structures of various kinases bound to small molecules, as well as the ability to characterize their modes of action, allows rapid progress in improving their biological, pharmaceutical, and physical properties. Small molecules that modulate the activity of kinases (inhibitors, active site as well as allosteric inhibitors, as well as activators) will continue to be an active area of research.

7.4 PROTEASES AND PROTEASE INHIBITORS

Proteases catalyze the hydrolysis of an amide bond in a highly controlled and regulated manner. Like kinases, they are ubiquitous in the body and are involved in numerous physiological processes, such as digestion, blood coagulation, catabolism, apoptosis, and regulation of signaling. Their overactivity or misactivity is associated with disease states, making them frequent drug discovery targets. In addition, infectious agents, such as bacteria, viruses, and parasites, also rely on proteases for their replication. Drugs that target these proteases, such as antibacterial β-lactam compounds and antiviral HIV protease inhibitors, have found widespread use.

At neutral pH, the amide bond is highly stable, with a half-life in water estimated at up to 1,000 years. This stability ensures the fidelity of peptides and proteins essential for normal physiological processes. However, when essential for a particular physiological process, proteases catalyze the cleavage of specific amide bonds in a peptide or protein. By convention, the bond that a protease cleaves is termed the **scissile bond**. Amino acid residues at the N-terminal side are numbered P1, P2, P3, etc., moving away from the site of reaction, and amino acid residues at the C-terminal side are numbered P1′, P2′, P3′, etc. (Figure 7.18). These correspond to binding sites S1–S3,

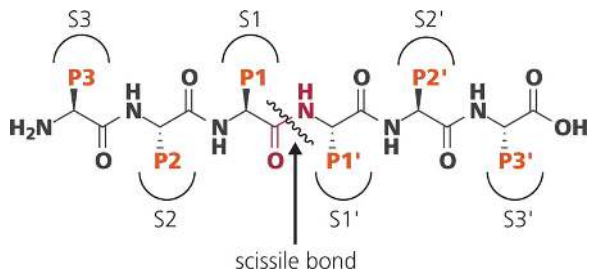


Figure 7.18 Protease substrate nomenclature and conventions. The amide bond that is cleaved by a protease is termed the *scissile bond* (red). For a six–amino acid peptide cleaved by a protease, the three amino acids at the N-terminal are named P1–P3 (red), and P1′–P3′ at the C-terminal (orange). These correspond to S1–S3 and S1′–S3′ sites on the enzyme, respectively.

etc., and S1′–S3′, etc., on the enzyme. Proteases can be highly specific or very general. For example, exopeptidases cleave a terminal amino acid, regardless of its identity; the serine protease chymotrypsin is selective for cleavage of any peptide bond that contains an aromatic amino acid (phenylalanine, tyrosine, or tryptophan) or large hydrophobic amino acid (such as methionine) at the P1 site. In contrast, the highly specific HIV protease recognizes eight specific amino acids in its substrate.

A very general mechanism for the protease reaction is shown in Figure 7.19. A nucleophile (Nu^-) attacks the carbonyl group of the amide and generates a tetrahedral intermediate. Collapse of the intermediate generates two peptide fragments, one containing the newly formed acyl-nucleophile.

Proteases are classified according to their reaction mechanism, and their active sites are highly conserved. Serine and cysteine proteases (Figure 7.20) are similar in that the nucleophile is a serine or cysteine residue found at the protease's active site. In these cases, as seen in the mechanism shown in Figure 7.19, the first products are the newly formed amine and an acyl-enzyme product, which is rapidly hydrolyzed in a second step to generate the carboxylic acid fragment and regenerate the enzyme. Key residues at the active site (Asp and His in the case of serine proteases, His in the case of cysteine proteases) are involved in acid–base catalysis to activate the nucleophile (serine and cysteine residue, respectively) by deprotonation. Aspartyl proteases and metalloproteases use water as the nucleophile (Nu = OH in Figure 7.19) and generate

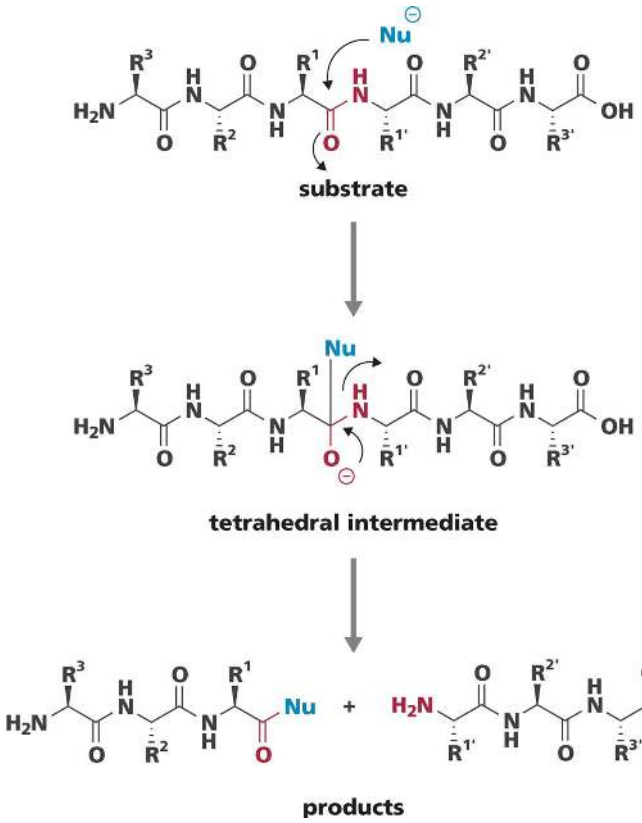


Figure 7.19 Abbreviated mechanism for hydrolysis of an amide bond in a peptide by a protease. The addition of the nucleophile (Nu^-; blue) to a specific amide carbonyl leads to a tetrahedral intermediate, which collapses to give two smaller peptide fragments, the nucleophile being incorporated into the N-terminal fragment. The scissile bond is shown in red.

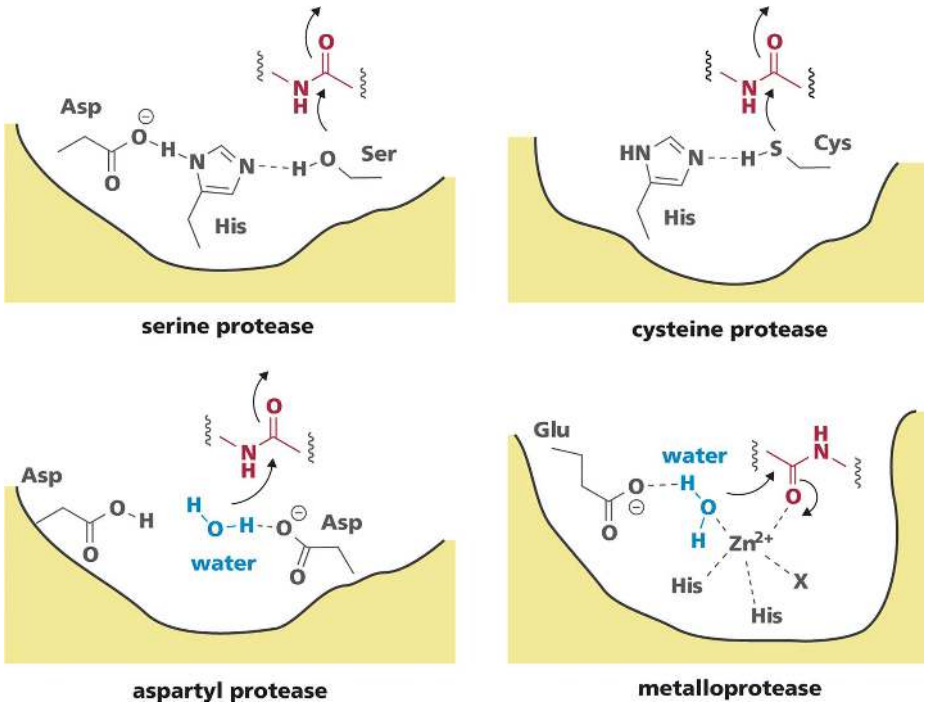


Figure 7.20 Active sites of serine, cysteine, aspartyl, and zinc metalloproteases (substrate in red). In a serine protease active site, acid–base catalysis, mediated by Asp and His residues, activates a serine side chain, which acts as a nucleophile and reacts with the peptide carbonyl (shown in red). Cysteine proteases act by a related mechanism and rely on a histidine residue to activate a cysteine residue to accelerate the reaction with the substrate carbonyl. Aspartyl proteases use two aspartates to deprotonate a water molecule in the active site, which then adds to the substrate. Metalloproteases use a metal to coordinate to the carbonyl of the substrate as well as to an activated water molecule in the active site; this nucleophile then reacts with the peptide carbonyl.

the amino and carboxyl fragments directly. Aspartyl proteases are characterized by the presence of two aspartic acid residues at the active site that act as general acid–base catalysts to deprotonate a water molecule. Metalloproteases contain a conserved metal cation (usually Zn^{2+}) that can serve as a Lewis acid to activate the water nucleophile and/or amide carbonyl as well as stabilize various intermediates via metal ion catalysis.

7.4.1 Protease Inhibitors Are Designed on the Basis of the Enzyme's Mechanism of Action and Are Useful for Treating HIV and Bacterial Infections, as Well as Cardiovascular Disease

Drug design targeting proteases relies on understanding the mechanism of action of the specific protease of interest. For example, drugs that inhibit **aspartyl proteases** often mimic the tetrahedral intermediate of the enzyme substrate, as described in Chapter 4. These include a series of HIV protease inhibitors, such as darunavir, lopinavir, saquinavir, and indinavir (Figure 7.21). The HIV protease is an essential enzyme that is required for the maturation of the human immunodeficiency virus and its replication. It cleaves at very specific sites that are unique to HIV, including one that contains Phe-Pro at the scissile site. In these examples, a substrate-like peptidic molecule was used as a starting point, and the scissile bond was replaced with various transition state mimetics, such as hydroxyethylene. Further optimization, remote from the transition state mimetic, led to the final drugs. Some of the structures in Figure 7.21, particularly saquinavir and indinavir, still bear some resemblance to the original Phe-Pro peptide starting point. Chapter 11, on antiviral agents, will describe these compounds in more detail.

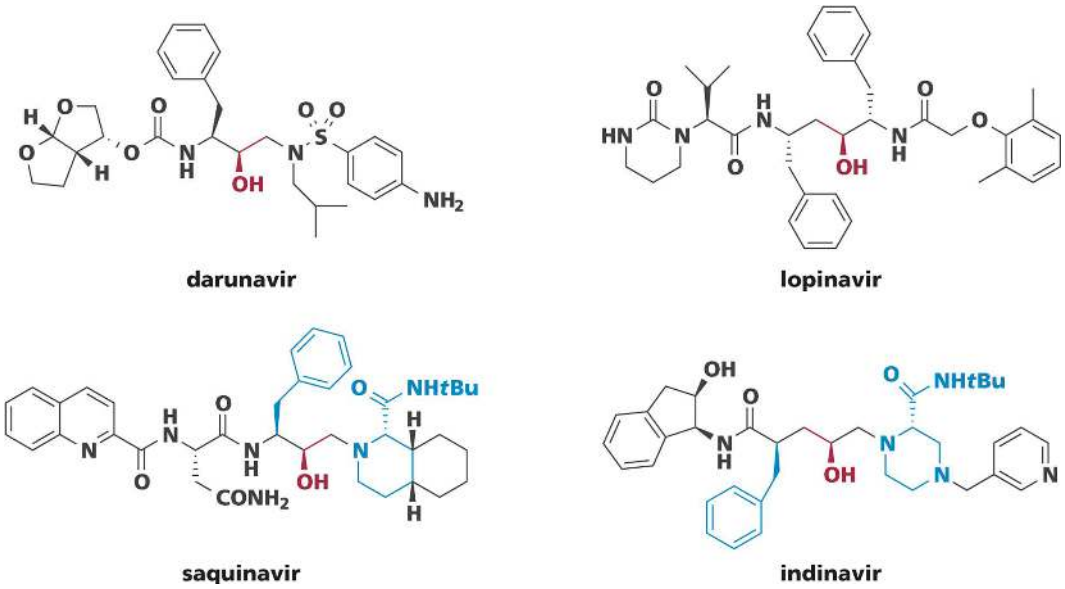


Figure 7.21 Structures of HIV protease inhibitors. Transition state mimetics found in the HIV drugs darunavir, lopinavir, saquinavir, and indinavir are shown in red; vestiges of the original substrate containing Phe-Pro at the scissile bond evident in saquinavir and indinavir are shown in blue.

Drugs targeting **serine/threonine proteases** include the penicillins and other β-lactam antibacterial compounds, which will be discussed further in Chapter 12, as well as the proteasome inhibitors described in Chapter 10. Like other protease inhibitors, the molecules mimic the natural substrate; however, they contain a **warhead**, an electrophilic moiety that reacts with the active site nucleophile (serine/threonine hydroxyl) and prevents the hydrolysis step in the enzymatic reaction. In contrast to aspartyl protease inhibitors, most serine/threonine protease inhibitors that act at the active site make a covalent bond to the enzyme. Some of these covalent bonds are reversible (slow, tight-binding inhibitors), but most are irreversible. While few approved drugs target **cysteine proteases**, strategies to inhibit them are similar to those used to design drugs to inhibit serine/threonine proteases and include the introduction of a warhead into a substrate-like molecule. Examples of warheads known or reported to generate effective inhibitors include β-lactams, boronic acids, epoxides, aldehydes, nitriles, α-heteroatom carbonyls, and vinyl sulfones (Figure 7.22).

Figure 7.23a shows the reaction of penicillin-binding protein (PBP), a serine protease, with its natural substrate, the D-Ala-D-Ala terminus of a peptidoglycan strand in bacterial cell walls. The reaction results in a modified PBP that is essential

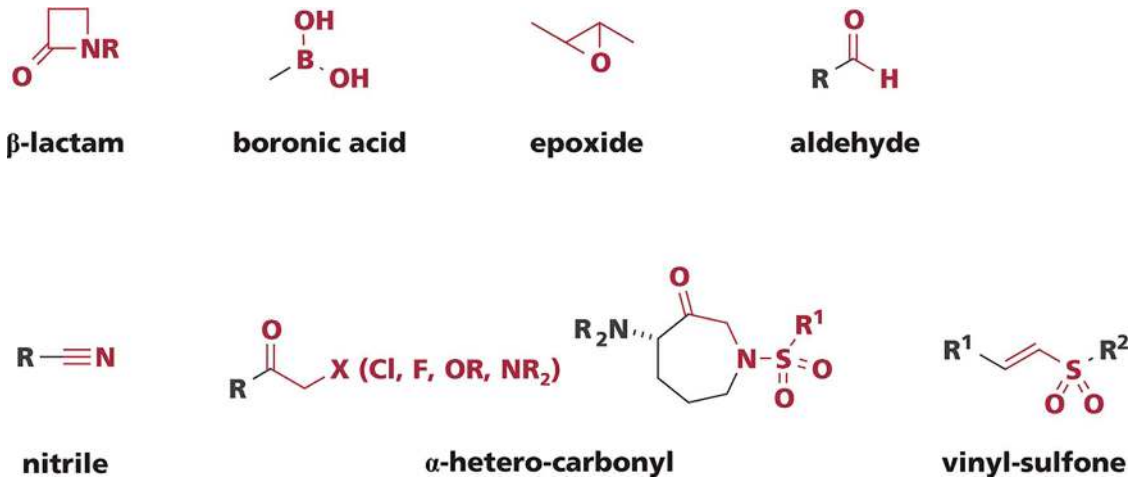


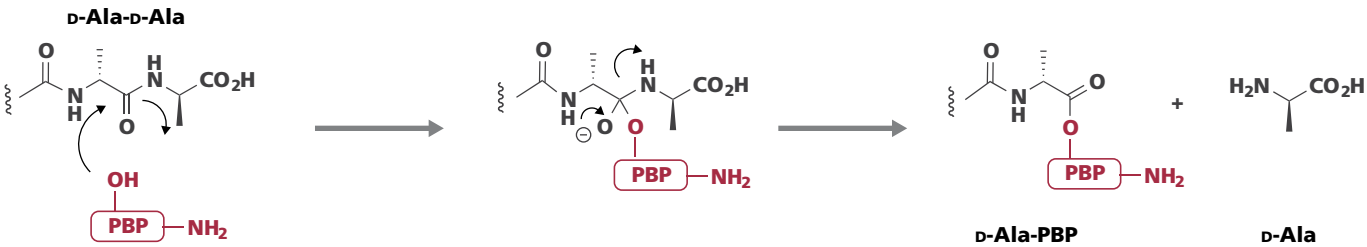
Figure 7.22 Examples of electrophilic warheads (shown in red) incorporated into serine/threonine and cysteine protease inhibitors. All the functional groups shown are electrophilic warheads that will react with a nucleophilic hydroxyl or thiol found in serine/threonine proteases and cysteine proteases, respectively.

Figure 7.23 Reactions of penicillin-binding protein. (a) Reaction of the serine hydroxyl of penicillin-binding protein (PBP, shown in red) with natural substrate peptidoglycan terminated in D-Ala-D-Ala. (b) Reaction of penicillin-binding protein with penicillin G at the β-lactam carbonyl generates an acylated, inactivated PBP.

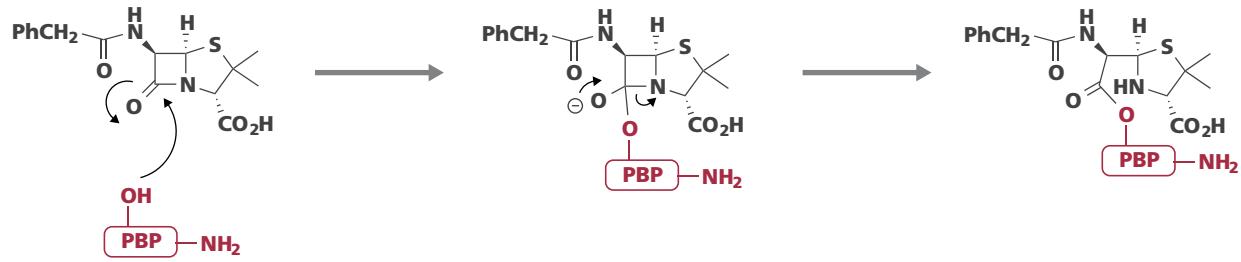
for cross-linking and synthesizing the bacterial cell wall. In the presence of penicillin, which contains a reactive β-lactam warhead (Figure 7.23b), the PBP's active site serine attacks the electrophilic carbonyl to generate a chemically modified, and therefore inhibited, protein, thus stopping cell wall biosynthesis. These are covalent irreversible inhibitors.

Clavulanic acid is another serine protease inhibitor that incorporates the same β-lactam warhead; however, once acted upon by its target, further reaction makes this compound an example of a suicide inhibitor. This natural product does not have antibacterial activity, but it reacts with the penicillin-destroying β-lactamase enzymes. These are serine proteases that evolved from the PBPs as a resistance mechanism toward penicillins, and their substrates are β-lactams. The initial reaction of clavulanic acid with β-lactamases involves the addition of a serine hydroxyl (Ser70) to the β-lactam carbonyl of clavulanic acid, followed by ring opening, as in the penicillin mechanism (Figure 7.24). The resultant acyl-enzyme intermediate is unstable and opens to form an iminium ion, a reactive intermediate. This iminium ion reacts with a second serine (Ser130) in β-lactamase to ultimately give a bridged, inactivated enzyme.

(a) REACTION OF PENICILLIN-BINDING PROTEIN (PBP) WITH D-Ala-D-Ala TERMINUS OF PEPTIDE SUBSTRATE



(b) REACTION OF PENICILLIN-BINDING PROTEIN (PBP) WITH PENICILLIN G



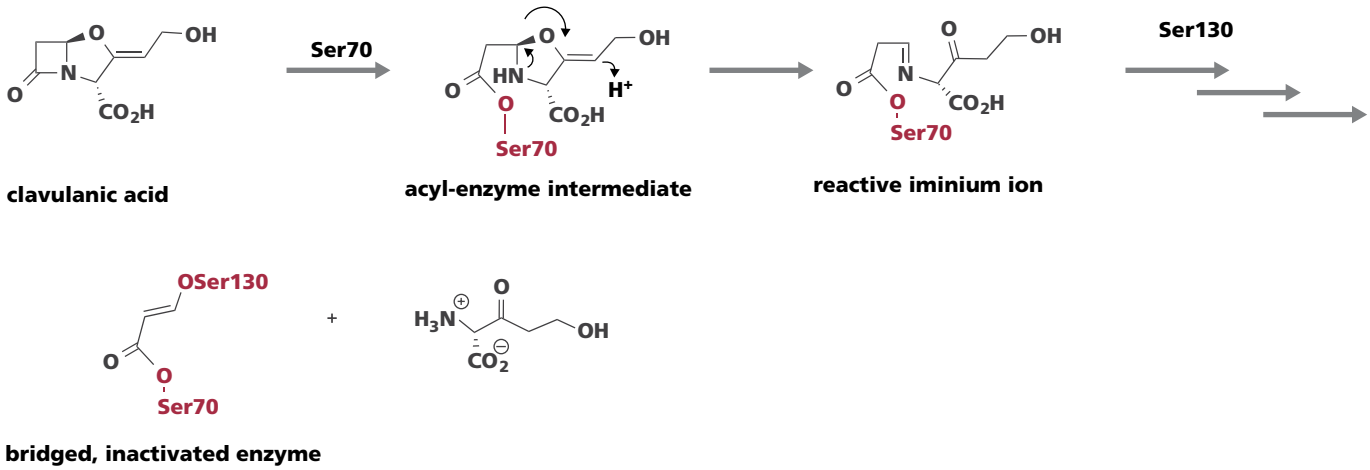


Figure 7.24 Mechanism of inhibition of β-lactamase by the suicide inhibitor clavulanic acid. Ser70 of β-lactamase adds to the carbonyl in the β-lactam, and the four-membered ring opens, generating an acyl enzyme intermediate. The unstable five-membered ring then opens to give a reactive iminium ion, which reacts further with the β-lactamase Ser130, forming a second covalent bond to the enzyme to give a bridged, inactivated enzyme.

Metalloprotease inhibitors take advantage of not only the structure of the substrate but also the presence of a metal cation at the active site to anchor binding. Inhibitors of angiotensin-converting enzyme (ACE), such as captopril and lisinopril, are examples of metalloprotease inhibitors (Figure 7.25). In these cases, the thiol of captopril and the central carboxylic acid of lisinopril bind to zinc in the active site. Other zinc-binding moieties frequently incorporated include hydroxamic acids, phosphonates, and hydantoins, among others.

Overall, the inhibition of proteases has been a very successful strategy in drug design, beginning with the introduction of β-lactam antibiotics. The examples included in this chapter cover a few classes of protease inhibitors, but there are many other drugs that work through this mechanism. In many of the most recent examples, structure-based design has been particularly successful. For example, the success of HIV protease inhibitors in treating AIDS led to the investigation of other proteases unique to viruses. This has recently come to fruition with the approval of HCV NS3/4A protease inhibitors to treat hepatitis C (see Case Study 2 and Chapter 11).

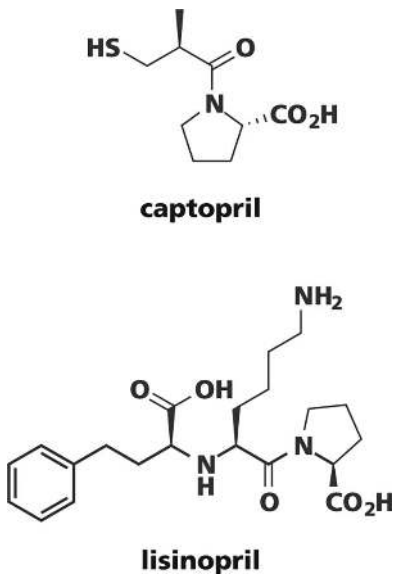


Figure 7.25 Structures of ACE inhibitors captopril and lisinopril. The thiol of captopril and the carboxylic acid of lisinopril bind to a zinc ion in the active site of the metalloprotease enzyme.

7.5 POLYMERASES AND POLYMERASE INHIBITORS

Polymerases are enzymes that catalyze the formation of a phosphodiester bond between two nucleotides, adding the 5′- end of one nucleotide to the 3′- end of a DNA (or RNA) strand, thus elongating the strand (Figure 7.26) and releasing diphosphate. Through a highly complex mechanism involving numerous other enzymes, such as topoisomerase, integrase, and helicase, polymerases are essential for nucleic acid replication. Humans use DNA polymerases to replicate and repair DNA, and these as well as other accompanying enzymes are often targets for anti-cancer therapies. Similarly, infectious organisms require polymerase-like enzymes for their replication, and the inhibition of these enzymes has been a highly successful therapeutic approach.

Polymerases work through a complex mechanism, abbreviated in Figure 7.26. A template strand of DNA (not shown) serves to direct which of the four nucleoside triphosphates is accepted through Watson–Crick base pairing. The two substrates in the reaction are a primer strand of DNA, which is eventually elongated, and the incoming nucleoside triphosphate (red). Divalent cations (usually magnesium) are also required at the active site; they coordinate with aspartates in the active site (blue) and bind to both substrates in order to activate them. The formation of the new phosphodiester bond is accompanied by the release of diphosphate.

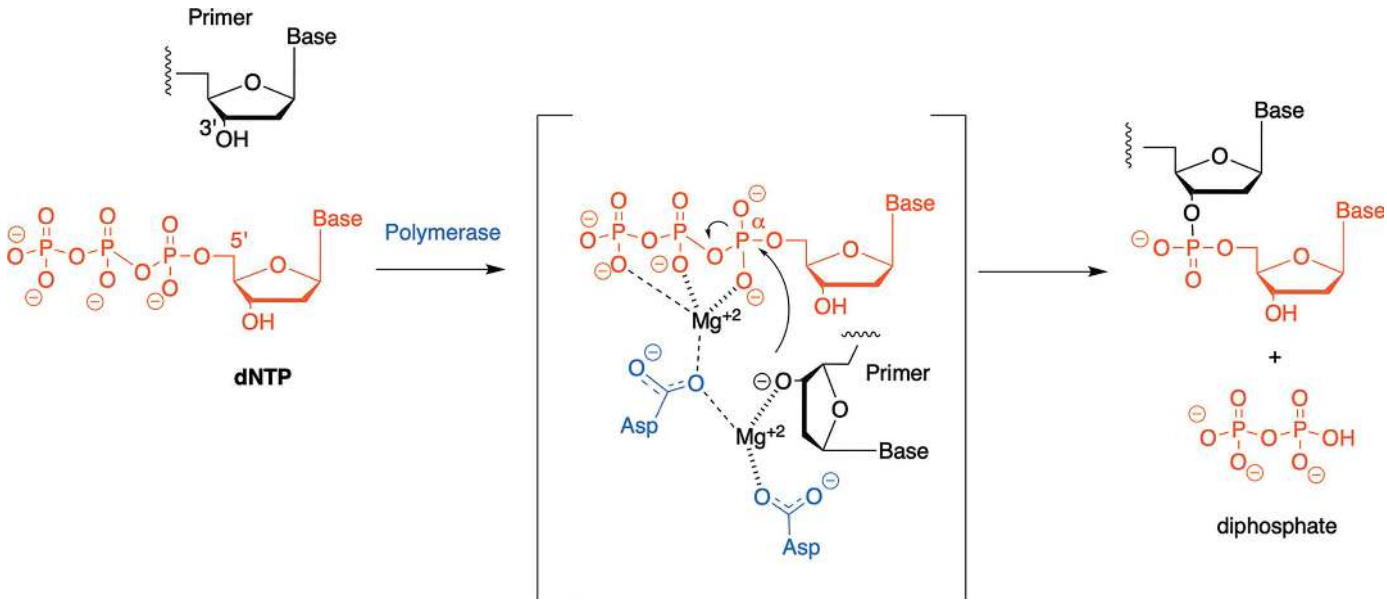


Figure 7.26 Polymerase-catalyzed reaction to form a phosphodiester bond. The phosphodiester bond is formed between a 3′-OH on the primer strand and the α-phosphate of an incoming deoxynucleoside triphosphate to extend the primer strand by one nucleotide. Abbreviated mechanism: A template strand (not shown) directs the identity of the base on the incoming deoxynucleoside triphosphate (dNTP; red) through Watson–Crick base pairing. Two magnesium cations (Mg^{2+}) at the active site coordinate to the incoming triphosphate, the reacting 3′-OH group on the primer terminal (black), and aspartates in the enzyme (blue). Nucleophilic attack of the 3′-O^- on the α-phosphate, followed by release of diphosphate, extends the primer chain by one nucleotide.

7.5.1 Inhibitors of Polymerases Can Mimic the Substrate or Bind at an Allosteric Site

Drug design for inhibition of RNA and DNA polymerases or their analog, reverse transcriptase, used by RNA viruses to make DNA from RNA, has taken a slightly different tack than relying on the characteristics of the active site and preventing substrate binding. In the most successful examples, inhibitors of the HIV reverse transcriptase are **chain terminators** (or **antimetabolites**). A nucleoside-like triphosphate molecule, competing with the natural substrate, is accepted as a substrate of the enzyme and incorporated into the growing nucleic acid chain. The inhibitor, however, does not contain the 3′-hydroxyl group that is required for further elongation, and therefore, the growing nucleic acid chain is terminated, effectively inhibiting further activity of the enzyme since no substrate is available (Figure 7.27).

Examples of nucleoside reverse transcriptase inhibitors (NRTIs) that inhibit HIV reverse transcriptase through a chain termination mechanism include abacavir, emtricitabine (2′-deoxy-5-fluoro-3′-thiacytidine, FTC), and azidothymidine (AZT), among others (Figure 7.28). All require initial reaction with kinases to form the corresponding triphosphate at the 5′-hydroxyl group, which is the actual enzyme substrate, such as AZT triphosphate. Similar classes of viral polymerase inhibitors, which are nucleoside analogs requiring prior activation to the triphosphate, have been developed for other viruses, such as hepatitis C. These will be described in further detail in Chapter 11 on antiviral agents.

Some polymerases, such as HIV reverse transcriptase, are also susceptible to allosteric inhibition. High-throughput screening identified small molecules that act at an allosteric site and inhibit HIV reverse transcriptase; these are known as nonnucleoside reverse transcriptase inhibitors (NNRTIs). Compounds such as nevirapine, efavirenz, and rilpivirine are allosteric inhibitors (Figure 7.29). They bind to a site about 10 Å away from the active site and force a conformational change of the enzyme, making it incompetent for enzymatic activity. Patients (and cells) treated solely with NNRTIs rapidly develop resistance to them, likely because of their allosteric mechanism of action.

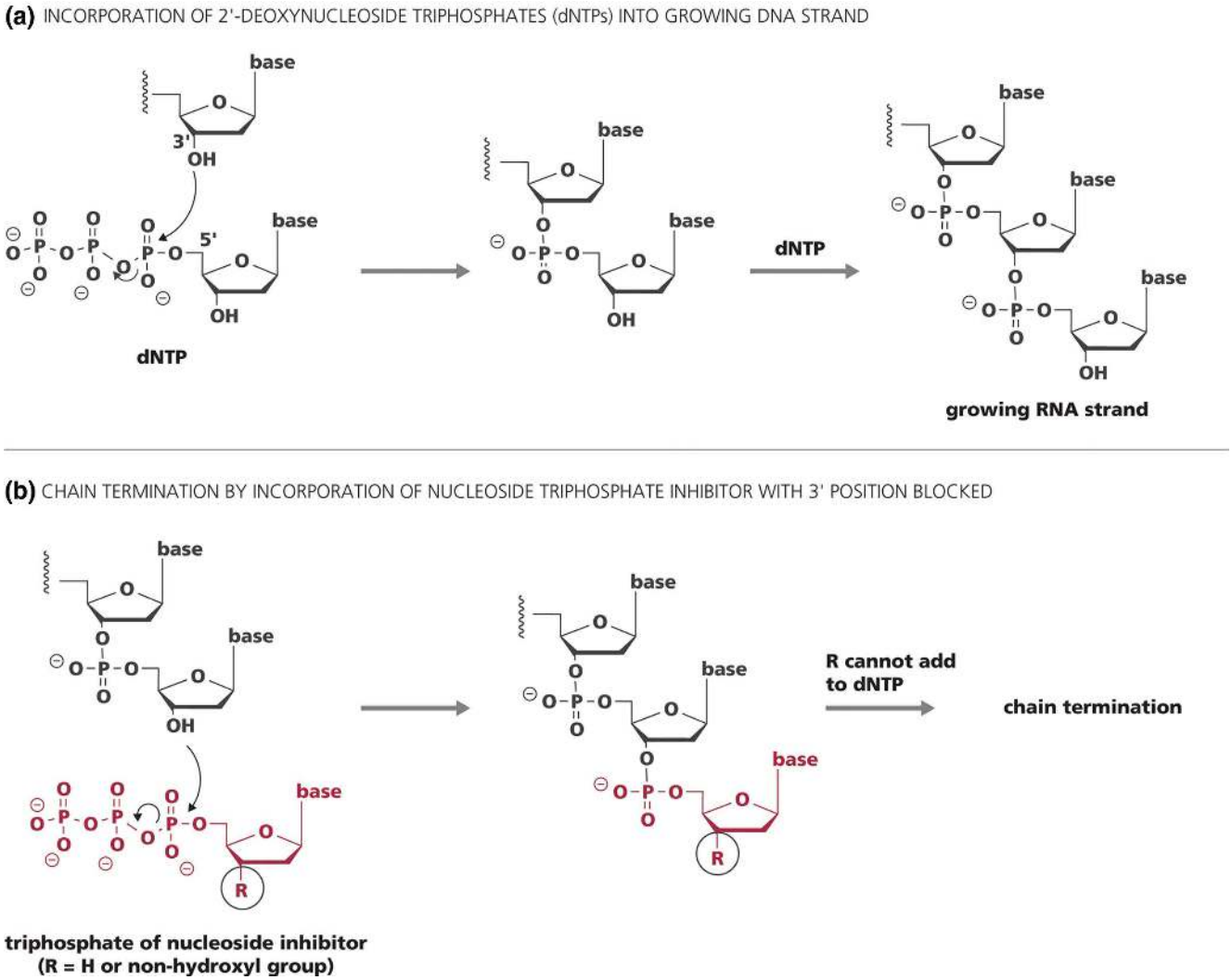


Figure 7.27 Chain termination mechanism showing the incorporation of a nucleoside reverse transcriptase inhibitor (NRTI) into a growing DNA strand. (a) The typical process of chain elongation occurs by the addition of a 3′-hydroxyl to a 5′-triphosphate. (b) Chain terminators work by the incorporation of a triphosphate of a nucleoside inhibitor (red) lacking a 3′-hydroxyl group (*R*); after its incorporation, chain elongation cannot proceed.

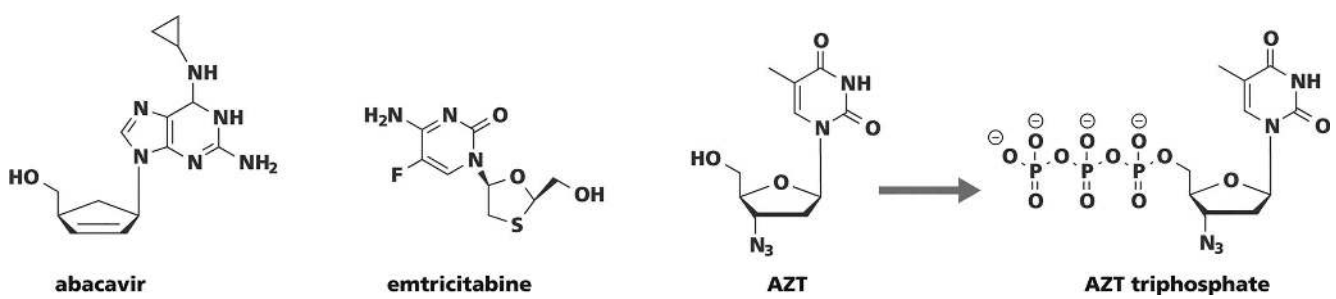


Figure 7.28 Structures of HIV drugs abacavir, emtricitabine, and AZT. These anti-HIV drugs are converted to the corresponding triphosphates, such as AZT triphosphate, and these triphosphates are substrates for HIV reverse transcriptase; they then act via a chain termination mechanism.

The inhibition of DNA and RNA polymerases was among the earliest strategies for treating viral infections with small-molecule drugs (Chapter 12) and continues to be an important approach, in some cases leading to cures of viral diseases. However, selectivity for viral polymerase enzymes over human enzymes continues to be challenging, as does finding an appropriate means to generate and deliver the active triphosphate.

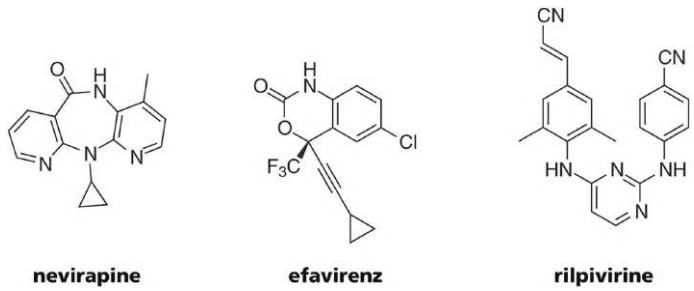


Figure 7.29 Structures of nevirapine, efavirenz, and rilpivirine. These anti-HIV drugs are nonnucleoside reverse transcriptase inhibitors (NNRTIs), acting as noncompetitive allosteric inhibitors.

7.6 ESTERASES, PHOSPHODIESTERASES, AND THEIR INHIBITORS

Esterases are a large family of enzymes that hydrolyze a carboxylic ester into an acid and an alcohol. These enzymes typically contain serine at their active site, and the mechanism is very similar to that of a serine protease. Acetylcholinesterase (see Figure 7.2) is an example of a therapeutically important esterase enzyme. Inhibitors of acetylcholinesterase are used to treat Alzheimer's disease and will be discussed in Chapter 13.

Phosphodiesterases (PDEs) catalyze cleavage of phosphodiester bonds and fall into several categories. Exonucleases and endonucleases cleave phosphodiester linkages on polynucleotide strands of DNA and RNA. However, for therapeutic strategies, the most important class is the cyclic nucleotide PDEs that cleave cAMP or cGMP (Figure 7.30).

The active site of these PDEs has both a nucleotide recognition pocket (structurally unique in each PDE subtype) and a hydrolysis site. There are designated regions in the active site; among them are the M site (metal ions), Q site (core pocket, which likely binds the cyclic nucleotide substrate), and the H site (hydrophobic pocket). The orientation of a specific glutamine (Q453 in Figure 7.31a), which forms hydrogen bonds to the purine, is involved in determining whether the PDE is selective for cGMP or cAMP. A mechanism for hydrolysis has been elucidated from X-ray crystal structures, as shown in Figure 7.31b for PDE9.

During hydrolysis, two metal ions (M1 and M2) coordinate to a water molecule, deprotonating it (W0), and they also form a complex with the phosphodiester while the activated nucleophilic hydroxide adds to the phosphorus atom in the phosphodiester group. Both Glu423 (E423) and His252 (H252) are involved in acid–base catalysis. A pentavalent enzyme–substrate complex forms, which proceeds to the tetravalent enzyme–product complex upon cleavage of the 3′-O–P bond. The addition of water hydrolyzes the complex to give GMP. During catalysis, a hydrophobic clamp in the H region, composed of Phe and Leu residues, interacts with the nucleobase of the substrate, and the amide of a conserved glutamine (Q453) forms hydrogen bonds to either adenine or guanine.

There are 11 subtypes of cyclic nucleotide PDEs, some of which cleave both cAMP and cGMP and some of which are selective. Phosphodiesterases 4, 7, and 8 are selective

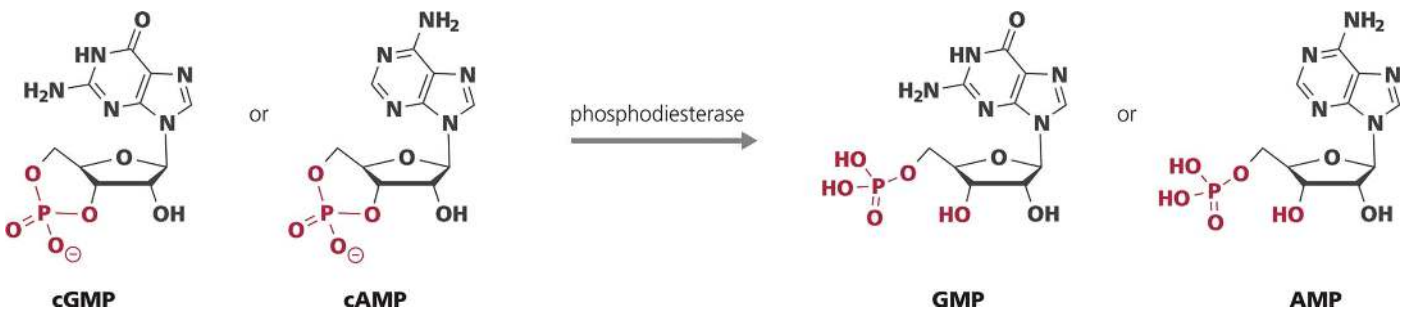


Figure 7.30 Hydrolysis of cGMP or cAMP catalyzed by phosphodiesterase. Phosphodiesterase enzymes generate monophosphates (GMP and AMP) from phosphodiesters cGMP and cAMP. The phosphodiester group hydrolyzed, and the resultant phosphate and hydroxyl groups are shown in red.

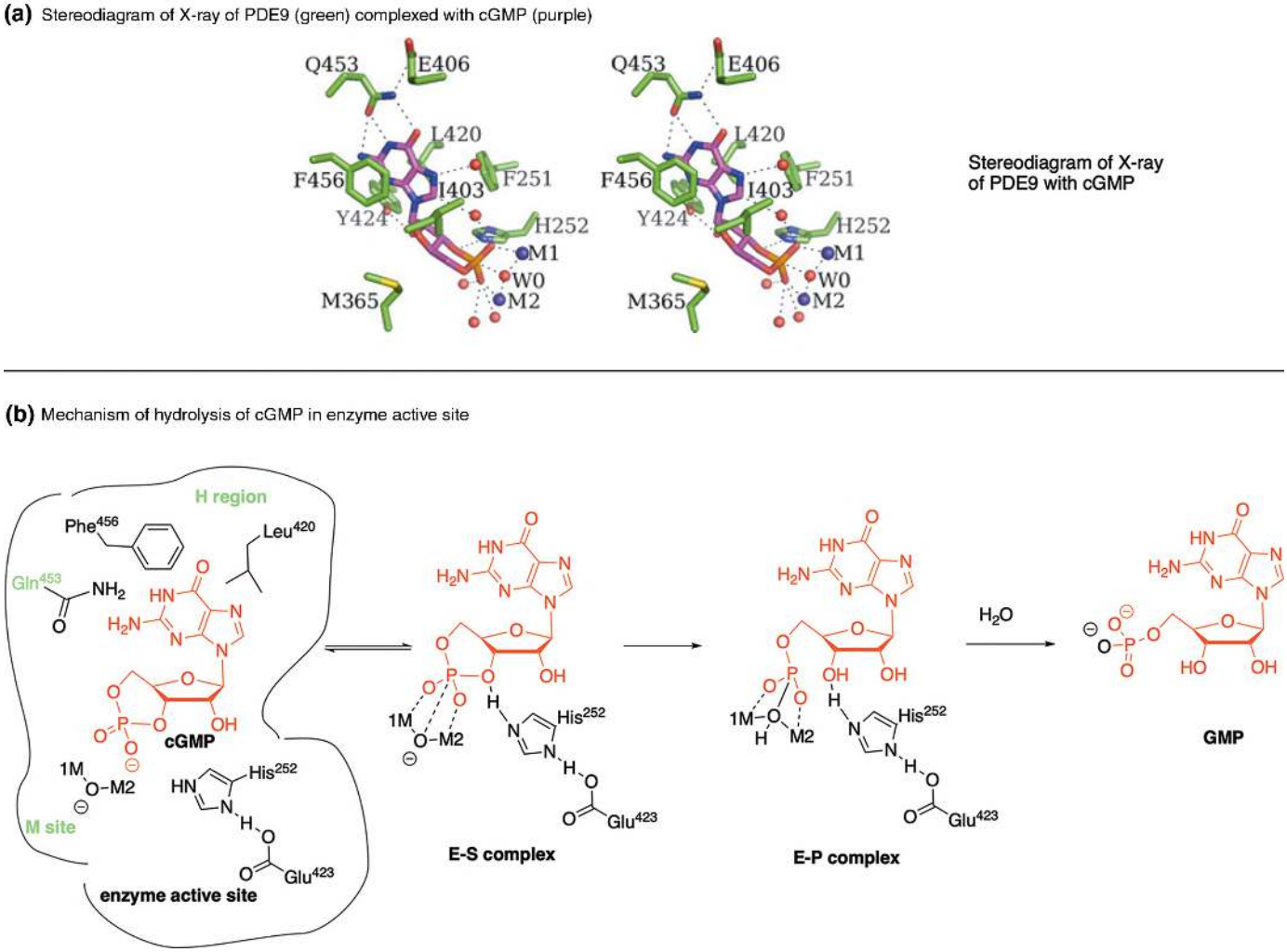


Figure 7.31 X-ray structure of cGMP in the active site of PDE9. (a) Stereodiagram of X-ray structure of PDE9 (green) complexed with cGMP (purple). M1 (Zn^{2+}) and M2 (Mg^{2+}) form a complex with a water molecule (W0) in the hydrolysis center; Q453 recognizes the guanine nucleobase. (b) Hydrolysis mechanism showing acid–base catalysis by Glu423, His252, and water (shown as OH^-) complexed to metal ions M1 and M2 in the M site. The metal ions form a complex with the cyclic phosphate and deliver the nucleophilic water to give the E–S complex; cleavage of the 3′-O–P bond gives the E–P complex; and the addition of water affords the hydrolyzed GMP. Gln453 forms H-bonds with the nucleobase, and Phe456 and Leu420 residues in the H-site, among others, engage in hydrophobic interactions with it.

Source: From Liu, S.; Mansour, M. N.; Dillman, K. S.; et al. (2008). *Proc Natl Acad Sci USA* 105:13,309–13,314. With permission from the National Academy of Sciences.

for cAMP, while PDEs 6 and 9 are selective for cGMP. Since cAMP and cGMP are important second messengers and have many roles in physiological processes, increasing their levels by inhibiting their degradation via PDE inhibition can have profound effects.

7.6.1 Inhibitors of Phosphodiesterases Treat Cardiovascular Disease, Respiratory Disease, and Erectile Dysfunction

Drug discovery targeting PDEs has focused on subtype selectivity, since global inhibition would generate numerous undesirable side effects. Selective PDE inhibitors are used to treat heart failure (PDE3), chronic obstructive pulmonary disease (COPD) (PDE4), and erectile dysfunction (PDE5). PDE inhibitors were found in the 1970s by phenotypic screening campaigns searching for compounds that increased heart muscle contraction, as an improvement over the cardiac glycosides digoxin and digoxigenin (see Chapter 1). At that time, diuretics and cardiac glycosides were the first-line therapy used to treat heart failure, in spite of their very narrow therapeutic index. Scientists at Sterling-Winthrop initiated a search for alternatives and phenotypic screening for positive inotropic activity (increasing heart muscle contraction)

and identified bipyridines as promising hits. Amrinone (now called inamrinone) was effective in increasing heart contraction and had no adverse effects in animal studies (Figure 7.32). At the time of its discovery (1978), amrinone's mechanism of action was unknown, but several years later, it was found to be a selective PDE3 inhibitor. Inhibition of PDE3, which is found in cardiac muscle, results in an increase of intracellular cAMP levels. This effect then results in increased intracellular calcium levels, which promotes contraction of cardiac muscle and vasodilation of vascular smooth muscle.

Later clinical trials of inamrinone revealed side effects such as thrombocytopenia, fever, and anemia after long-term use. Analogs such as milrinone have fewer side effects. A number of other analogs have been investigated and led to drugs such as enoximone, as well as the pyridazinones meribendan and pimobendan, which are all competitive reversible inhibitors. A pharmacophore model (Figure 7.32), proposed in the 1980s, has a dipolar region, corresponding to the carbonyl of active compounds, with an adjacent acidic proton. There is also a small lipophilic area; these are linked to a hydrogen bonding region through a mostly planar surface. All these features are found in the more recent PDE3 inhibitors.

Inhibitors of PDE4, primarily expressed in inflammatory and immune cells, have been developed to treat chronic obstructive pulmonary disease (COPD), a disease that affects over 200 million people. Asthma affects another 300 million, and together these diseases are an enormous burden on the health-care system. The natural product theophylline (found in cocoa beans) (Figure 7.33) has been used to treat asthma since the 1920s. It is a nonselective PDE inhibitor, with additional activity as an adenosine receptor antagonist. Its side effects include increased heart contractions and blood pressure, as well as nausea. Newer selective PDE4 inhibitors were based on the compound rolipram, originally investigated at Schering-Plough as an antidepressant on the basis of its downstream effect of reducing levels of cAMP in the brain. Rolipram was later found to be a selective PDE4 inhibitor, and analogs such as piclamilast and roflumilast (see Figure 7.33a) were developed to treat COPD.

The development of rolipram was discontinued during clinical trials due to adverse effects, such as nausea and vomiting. According to a docking model, rolipram binds in the active site of PDE4, as seen in Figure 7.33b. Multiple H-bonds in the binding model include ones between the pyrrolidinone carbonyl and the hydroxyl group of Thr407, the pyrrolidinone nitrogen and Asn395, the methoxy oxygen and Cys499, and the cyclopentyl ether oxygen and His234. The core of

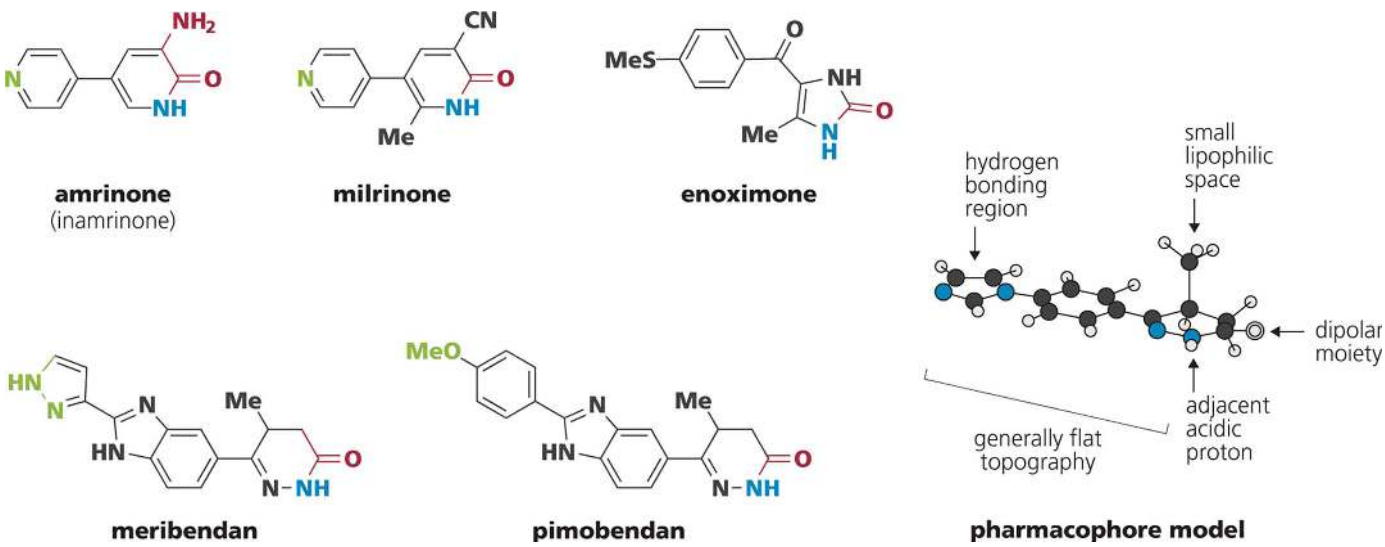


Figure 7.32 Structures of PDE3 inhibitors with pharmacophore model. Structures of amrinone, milrinone, enoximone, meribendan, and pimobendan are depicted. The pharmacophore model contains a dipolar region, corresponding to the carbonyl of active compounds, with an adjacent acidic proton, corresponding to the amide N-H. In the model, there is an area for a small lipophilic group (occupied by methyl groups in meribendan and pimobendan), and a planar-like linker joins these features to a hydrogen bonding region. Common dipolar regions of drugs are shown in red, acidic proton is in blue, and hydrogen bonding regions are shown in green.

Source: From Bristol, J. A.; Sircar, I.; Moos, W. H.; et al. (1984). *J Med Chem* 27:1,099–1,101. With permission from the American Chemical Society.

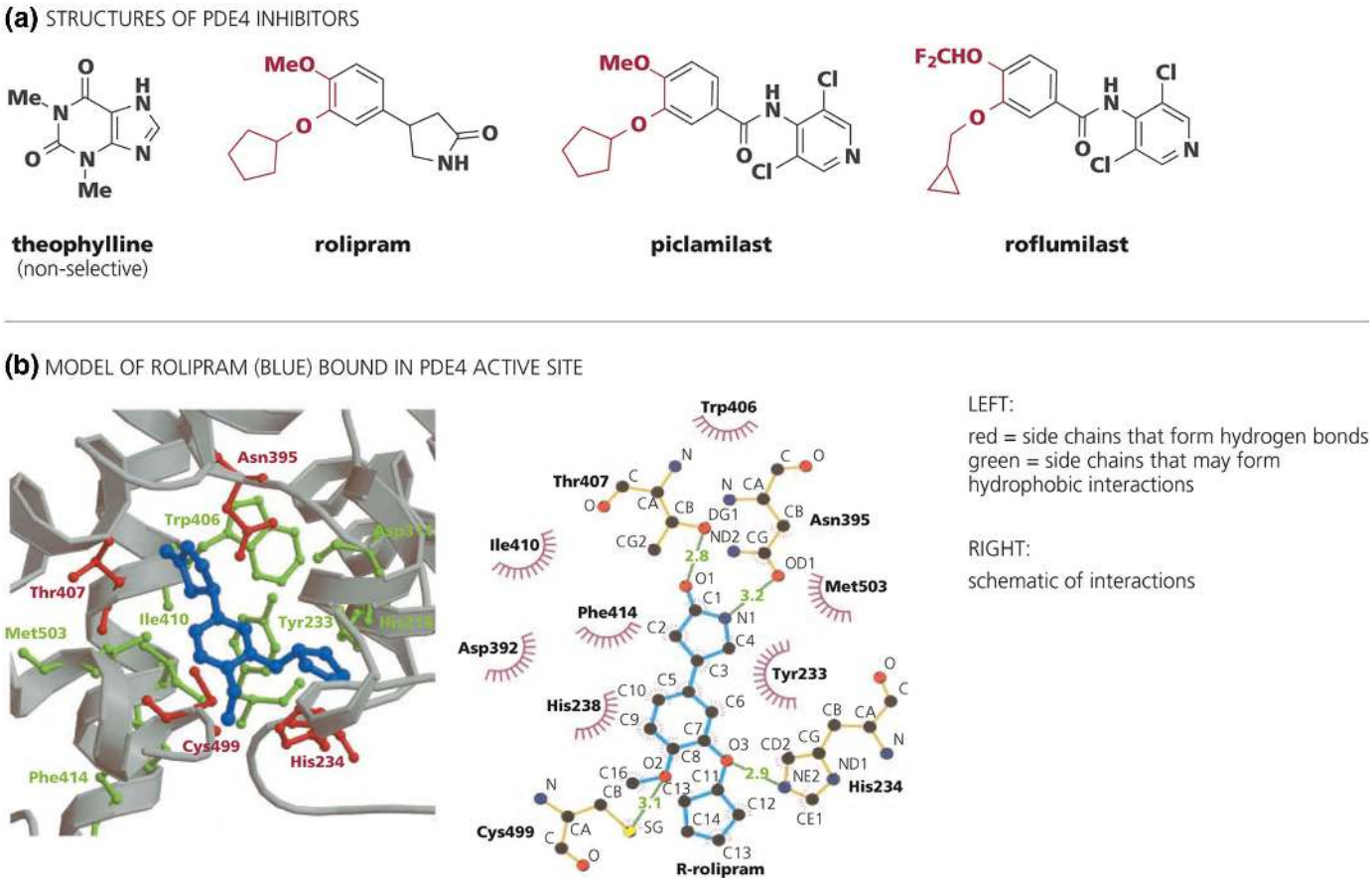


Figure 7.33 Structures of PDE4 inhibitors and model of binding. (a) Structures of theophylline (nonselective inhibitor) and the selective inhibitors rolipram, piclamilast, and roflumilast (dialkoxy catechol substituents in red). (b) *Left*, model of selective PDE4 inhibitor rolipram bound in PDE4 active site. Key interactions include H-bonds from pyrrolidinone carbonyl to Thr407 hydroxyl, pyrrolidinone nitrogen to Asn395, methoxy oxygen to Cys499, and cyclopentyl ether oxygen to His234 (residues that participate in H-bonds in red); hydrophobic interactions include Tyr233, Phe414, and His238 with the core of rolipram (residues participating in hydrophobic interactions in green. *Right*, schematic of specific interactions: rolipram in blue, H-bonding residues in yellow, and hydrophobic residues in red.

Source: From Dym, O.; Xenarios, I.; Ke, H.; and Colicelli, J. (2002). *Mol Pharmacol* 61:20–25. With permission from the American Society for Pharmacology and Experimental Therapeutics.

rolipram can make hydrophobic interactions with Tyr233, Phe414, and His238. Rolipram also interacts with a second binding site in the enzyme (not shown), and binding to that additional site is believed to be responsible for the side effects. Later analogs maintain the important 3,4-dialkoxycatechol of rolipram, which forms two key hydrogen bonds to PDE4, but lack the pyrrolidinone. The 3,5-dichloropyridines piclamilast and roflumilast were based on the bioisosteric replacement of the pyrrolidinone with a pyridine ring, and while piclamilast was discontinued, roflumilast is used clinically.

Inhibitors of PDE5 were originally developed at Pfizer to treat cardiovascular disorders, such as angina and hypertension. PDE5 is found in lungs, platelets, and smooth muscle, as well as in the corpus cavernosum (penile tissue). Drug discovery efforts at Pfizer were based on a report in 1975 on zaprinast, a weak and nonselective PDE inhibitor that had been prepared as an anti-allergy medicine. A number of 2-alkoxyphenyl heterocycles were investigated, with the pyrazolo[4,3-*d*]pyrimidin-7-one series showing good selectivity for PDE5. Improvement of the initial lead was based on the rationale that the heterocycle resembled the guanine base of cGMP, and therefore, adding a larger group at the 3-position would occupy the pocket in PDE filled by the ribose of cGMP (**Figure 7.34**). This approach led to the 3-propyl analog with improved PDE5 selectivity and high binding affinity. This

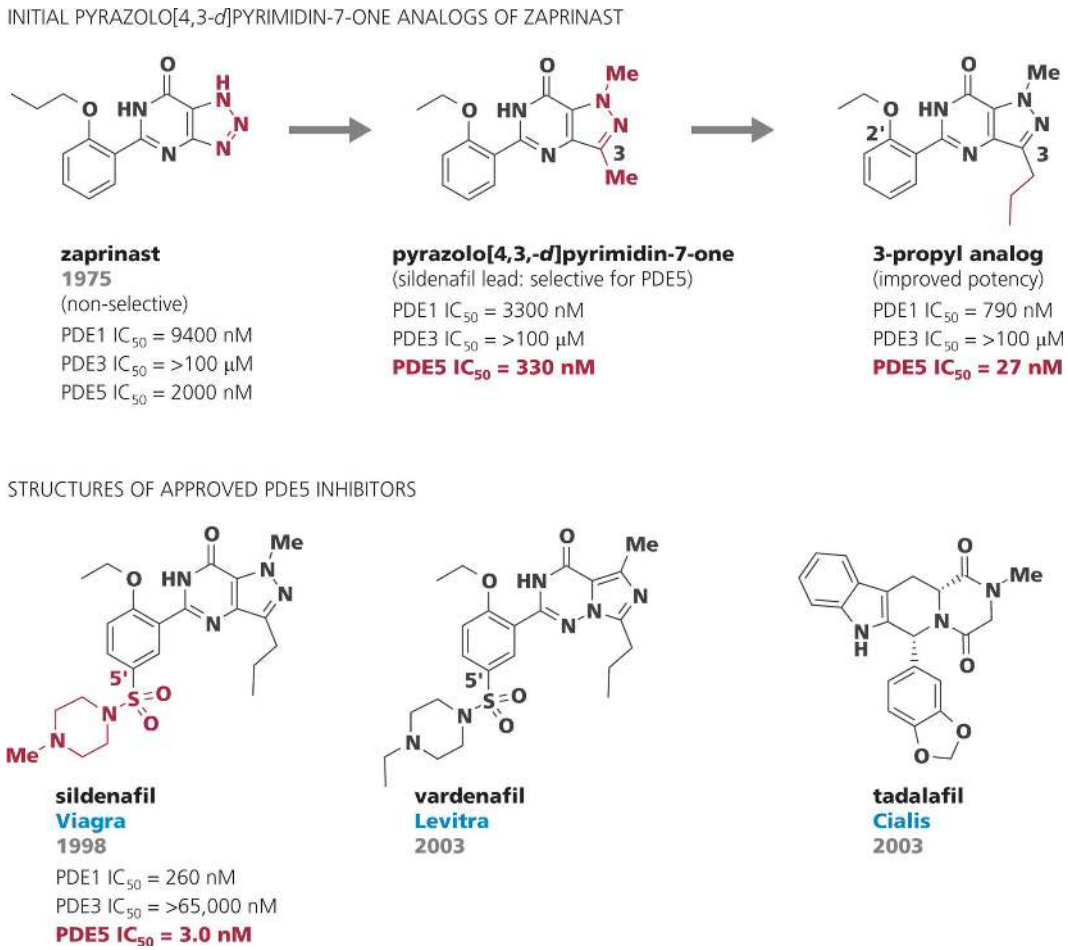


Figure 7.34 Lead development of sildenafil and other PDE5 inhibitors. Zaprinast, a nonselective hit, was modified, which led to the pyrazolo[4,3-*d*]pyrimidin-7-one as a more selective lead; progression to the 3-propyl analog was based on mimicking the ribose of cGMP (regions changed in each are colored red). Further modification by adding the piperazinyl sulfonamide led to sildenafil. Current PDE5 drugs include the structurally related vardenafil, as well as tadalafil.

analog was then further modified by the addition of the piperazinyl sulfonamide to afford sildenafil.

While the initial focus on this series was a search for compounds to treat angina and hypertension, their activity as a treatment for male erectile dysfunction, based on clinical observations, refocused the therapeutic use. Sildenafil received FDA approval in 1998, and later, PDE5 inhibitors vardenafil and tadalafil were approved in 2003. While vardenafil is similar in structure to sildenafil, tadalafil was based on a series of hydantoin PDE5 inhibitors discovered at GlaxoSmithKline.

An X-ray structure of sildenafil bound in PDE5 is shown in Figure 7.35. As proposed in the design, the 3-propyl group sits in the same binding site as the ribose of cGMP. Additional interactions include binding of the *N*-methylpiperazine to Tyr664, which contributes to PDE5 selectivity, and hydrophobic interactions in the H region between the ethoxyphenyl group and Phe786, and other hydrophobic side chains. The amide of the heterocycle hydrogen-bonds to Gln817. Similar binding interactions are seen in X-ray structures of tadalafil and vardenafil and are important to maintain in the design of these competitive reversible inhibitors.

In summary, selective inhibitors of PDE are used to treat a variety of conditions. Although early drugs such as amrinone and theophylline were discovered using phenotypic assays, more recent drugs have been discovered by a combination of screening for enzyme inhibition and structure-based drug design.

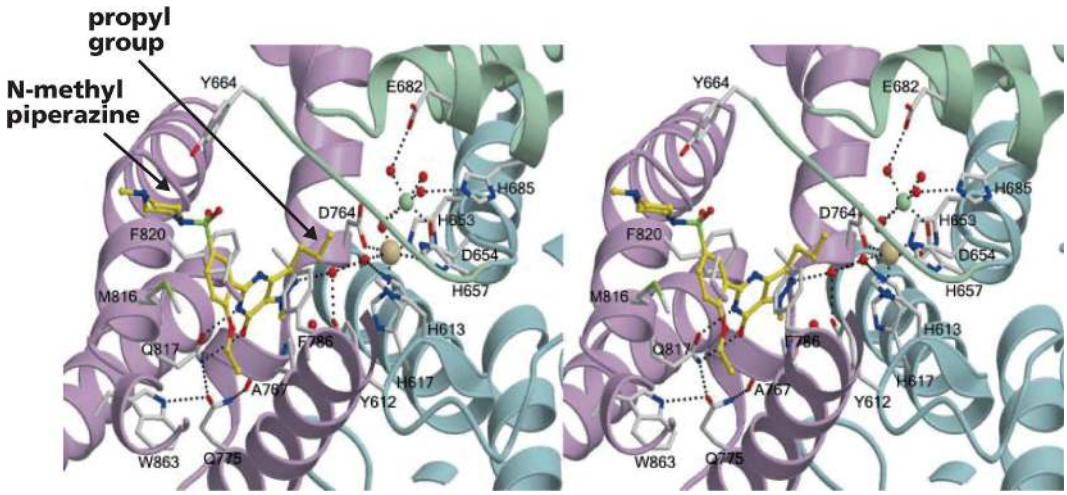


Figure 7.35 Stereodiagram of X-ray structure of sildenafil bound in PDE5. The amide within the heterocycle H-bonds to the glutamine Q817, and the propyl group sits in the same site as the ribose of cGMP. Hydrophobic interactions include Phe786 with the ethoxyphenyl and Phe820 (pi stacking) with the heterocyclic core. Sildenafil is shown in yellow; metal ions zinc (tan) and magnesium (green) are depicted as spheres.

Source: From Sung, B.-J.; Hwang, K. Y.; Jeon, Y. H.; et al. (2003). *Nature* 425:98–102. With permission from Macmillan Publishers Ltd.

7.7 OXIDOREDUCTASES, CYCLOOXYGENASES, AND INHIBITORS

Numerous enzymes catalyze oxidation and reduction reactions, and the oxidoreductase enzymes are one example. These enzymes act on diverse substrates, but the unifying feature of this family is transfer of electrons and/or hydrogens from one molecule to another. These enzymes rely on diverse mechanisms, for example, oxidases that use molecular oxygen as a hydrogen or electron acceptor or dehydrogenases that transfer hydrogen from a substrate to an enzyme system. Conversely, reductases may transfer hydrogens to a substrate. These enzymes may rely on coenzymes such as dihydronicotinamide adenine dinucleotide phosphate (NADPH) or flavin adenine dinucleotide (FAD) for hydrogen transfer. For instance, in the folic acid biosynthesis pathway, dihydrofolate reductase (DHFR) catalyzes the addition of two hydrogens to dihydrofolate, resulting in tetrahydrofolate, and uses NADPH as a co-factor and the source of one of the hydrogen atoms (**Figure 7.36a**). Inhibition of this pathway is important in both cancer and antibacterial chemotherapy and will be described in later chapters.

Two other examples of oxidoreductase-catalyzed reactions are the NADPH-dependent reduction of testosterone (**Figure 7.36b**) and the multi-step oxidative aromatization of testosterone to estrogen (**Figure 7.36c**). While the 5-alpha reductase enzyme proceeds through a mechanism similar to DHFR, the aromatase enzyme is a heme-dependent cytochrome P450 that uses molecular oxygen as an oxidant in a complex five-step mechanism.

While there are multiple therapeutic effects of inhibition of these enzymes, the diverse substrates, coenzymes, and mechanisms used by this class of enzymes make it difficult to generalize specific medicinal chemistry strategies that have been applied. Therefore, two specific examples are discussed in the following sections.

7.7.1 Inhibitors of HMG-CoA Reductase Are Used to Lower Cholesterol Levels

The early, rate-determining step in the biosynthesis of cholesterol is the conversion of 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) to mevalonate, which is in equilibrium with its cyclic form, mevalonolactone. The reduction is catalyzed by NADPH-dependent HMG-CoA reductase, which adds hydrogen atoms in two steps (**Figure 7.37**).

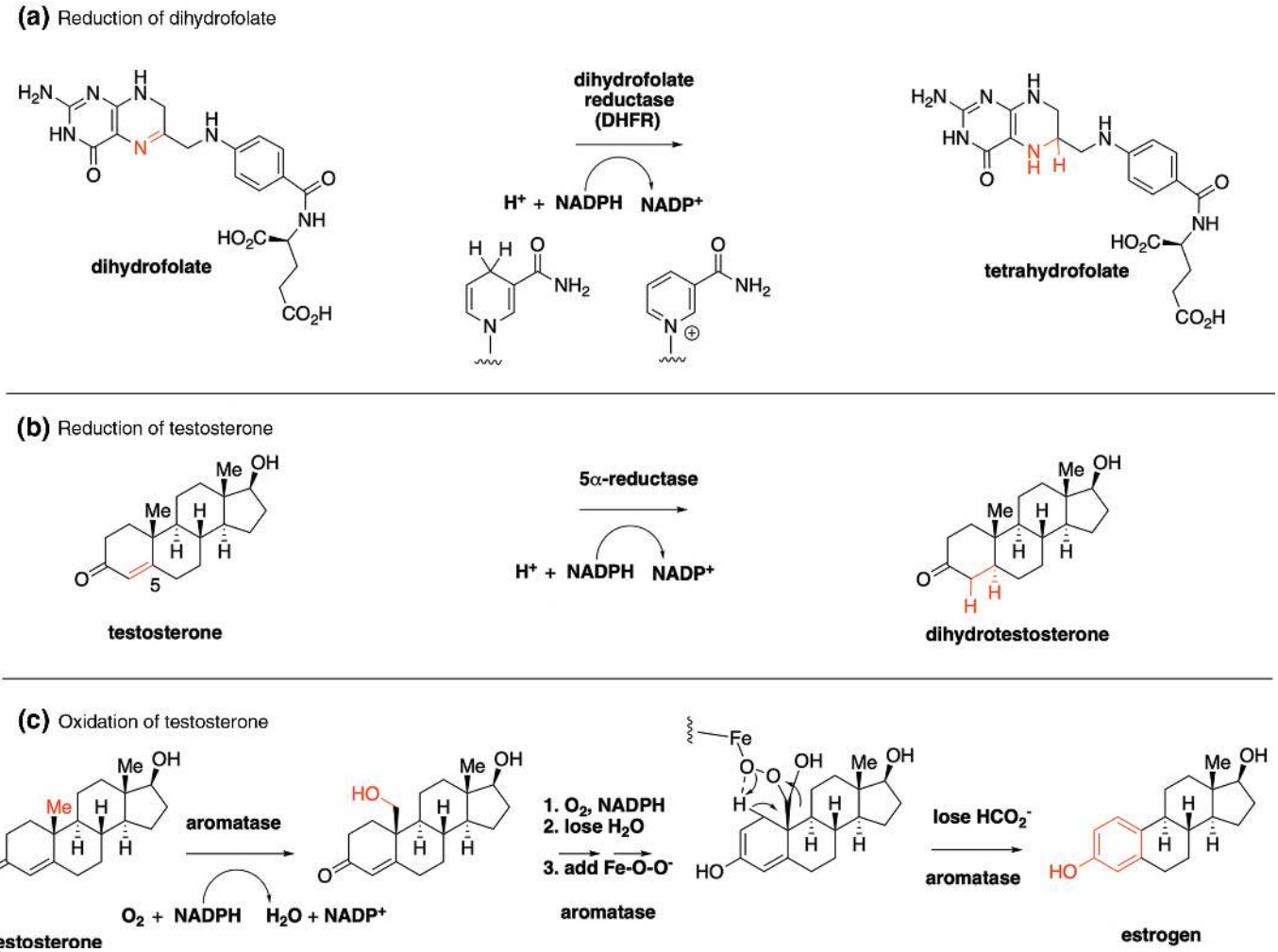


Figure 7.36 Three examples of reactions catalyzed by oxidoreductases: (a) Dihydrofolate is reduced to tetrahydrofolate by dihydrofolate reductase, via transfer of two hydrogens (red), one of which is donated by the conversion of NADPH to $NADP^+$. (b) Testosterone is reduced to dihydrotestosterone by 5α-reductase, by transfer of hydrogens to C4/C5. This reaction also requires NADPH to be converted to $NADP^+$ as a source of one of the H atoms. (c) Testosterone is oxidized by aromatase to estrogen in several steps, involving the oxidation of a methyl group to an alcohol, further oxidation to the acetal and then to the aldehyde; this is followed by the iron-catalyzed removal of hydrogen and the loss of formate. Changes highlighted in red.

On the basis of evidence that buildup of cholesterol is a risk factor for cardiovascular disease, a number of organizations developed inhibitors of the rate-determining step in cholesterol biosynthesis. While a series of ester analogs of HMG-CoA were found to be weak HMG-CoA reductase inhibitors, more success was found with screening microbial broths, and through this approach, compactin was reported by two separate groups (Beecham Pharmaceuticals and Sankyo Research Laboratories) in 1976 (Figure 7.38a). The ring-opened form of the lactone can be viewed as a transition state analog of the first reduction step.

Researchers at Merck discovered a close structural analog, mevinolin (lovastatin), in 1978 that contained a 3α-methyl group on the bicyclic core. Both mevinolin

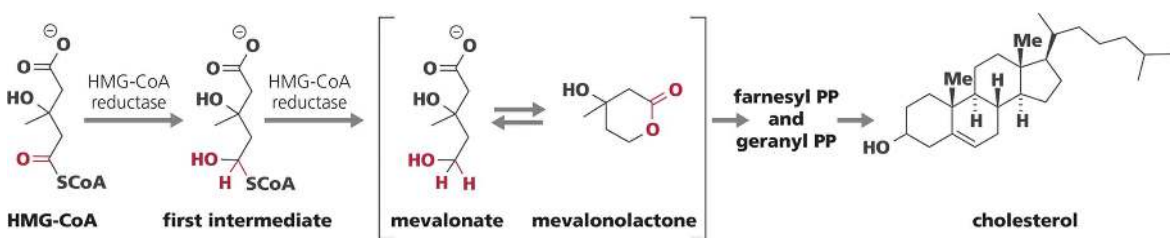


Figure 7.37 Formation of mevalonate, catalyzed by HMG-CoA reductase and conversion to cholesterol. This reaction is the rate-determining step in cholesterol biosynthesis. Overall, two hydrogens are added to give mevalonate, the precursor to farnesyl pyrophosphate (PP) and geranyl PP, which are precursors to cholesterol (reaction site in red).

and compactin showed excellent efficacy in lowering low-density lipoprotein (LDL) cholesterol with no apparent side effects (at low doses), and mevinolin/lovastatin received FDA approval in 1987. A semisynthetic analog, simvastatin, was approved soon after, as was pravastatin, the 3-hydroxy metabolite of compactin. While these compounds, termed statins, are natural or semisynthetic products, later analogs were based on an understanding of the structural elements required for activity, which include a dihydroxyheptanoate moiety linked to a lipophilic group (**Figure 7.38b**). This model was based on a number of synthetic analogs reported in the literature by groups at Merck and Sankyo. The indole analog fluvastatin, from Novartis, was approved in 1994, and at Pfizer and Shionogi, other heterocycles were chosen as synthetically accessible linkers, and a series of analogs were synthesized, leading to atorvastatin, also known as Lipitor (pyrrole linker) and rosuvastatin (pyrazine linker). These synthetic compounds all have in common a dihydroxyheptanoate mimic of mevalonate, a heterocyclic linker replacing the decalin, a small lipophilic group (isopropyl), and a 4-fluorophenyl group that is proposed to occupy the same space as the butyric ester of compactin. The statins, first introduced in 1987, are the leading treatment for the reduction of LDL cholesterol. For a time, Lipitor was the top-selling drug in the United States.

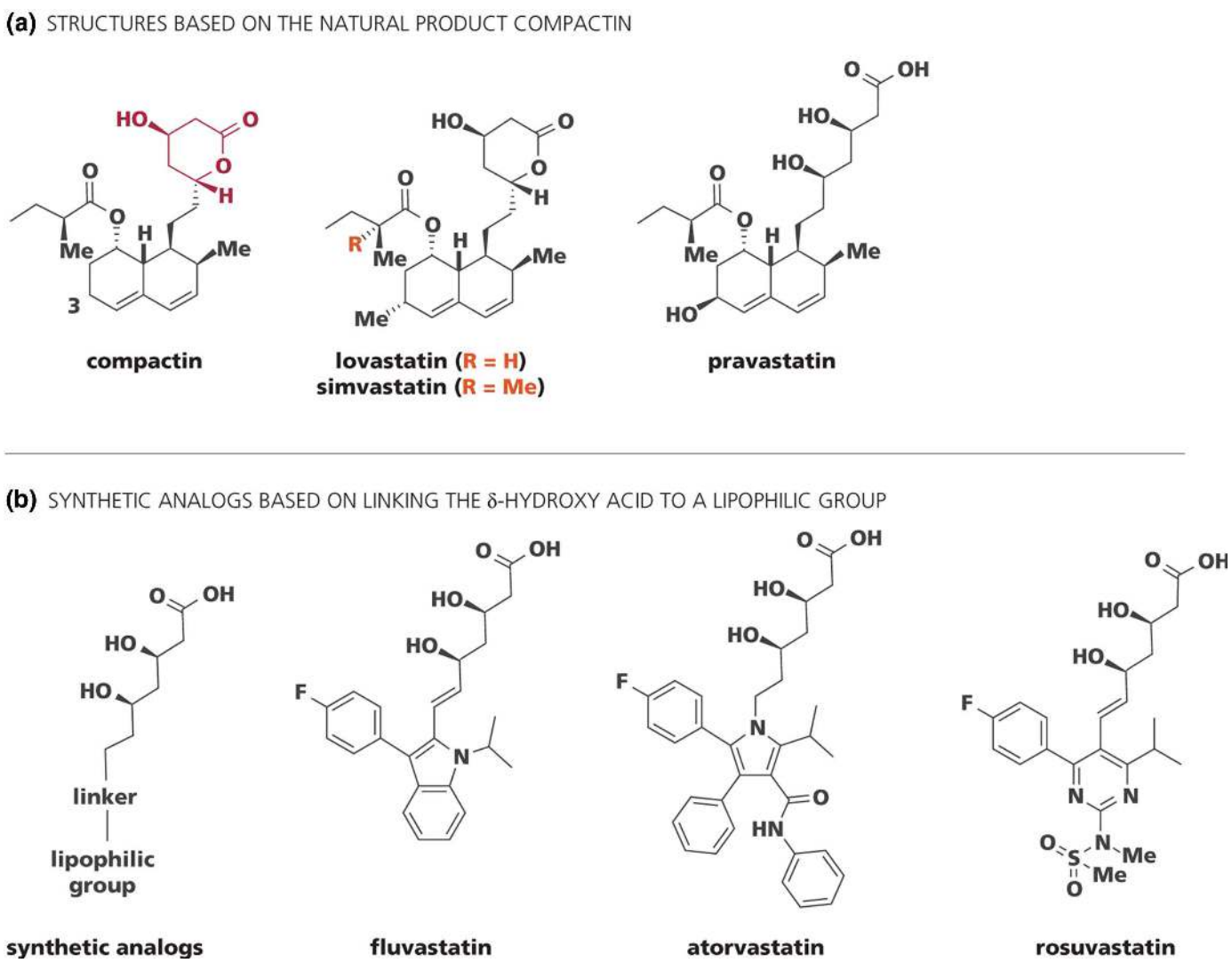


Figure 7.38 Structures of HMG-CoA reductase inhibitors. (a) Structures of compactin, lovastatin, simvastatin, and pravastatin, the structural features in compactin that mimic the transition state when in ring-open form, are shown in red. (b) Model of HMG-CoA reductase inhibitors and synthetic derivatives. Fluvastatin, atorvastatin, and rosuvastatin. All can be mapped to linking a mevalonate moiety (δ-hydroxy acid) to a lipophilic group. Other common structural features include a 4-fluorophenyl group that occupies the same site as the butyric acid in compactin, a heterocycle that mimics the decalin unit in compactin, and an isopropyl group.

7.7.2 Inhibitors of Cyclooxygenase Are Effective Anti-Inflammatory Drugs

Cyclooxygenase (COX), or prostaglandin-endoperoxide synthase (PTGS), is a heme-associated enzyme that catalyzes the conversion of the fatty acid arachidonic acid to prostaglandin H_2 (PGH_2), as shown in **Figure 7.39**. PGH_2 is the precursor to all other prostaglandins as well as thromboxane. The prostaglandins mediate multiple physiological effects, including inflammation, uterine contraction, and protection of the stomach lining. Thromboxane is derived from prostaglandins in the platelets and plays a role in platelet aggregation. There are two isoforms of the cyclooxygenase enzyme, COX-1 and COX–2. COX-1 is constitutively active in almost all tissues. The expression of COX-2 is more limited. It is produced at a consistent level in some tissues, such as brain and kidney, but in many tissues, it is up-regulated in response to inflammatory events, such as cytokine production. In both isoforms, there are two components to the catalytic site: The cyclooxygenase (COX) function converts the substrate to prostaglandin G_2 (PGG_2), and the heme-associated peroxidase (POX) completes the conversion from PGG_2 to PGH_2. The mechanism of the COX reaction is shown in **Figure 7.39**. Electron transfer from a tyrosine to the heme in COX initiates the formation of a tyrosyl radical (Tyr385), which abstracts the 13-pro-(*S*) hydrogen of the substrate. The addition of molecular oxygen to this species gives an 11-peroxy radical, which cyclizes and adds oxygen at the C-15 position to generate PGG_2. PGG_2 then diffuses to the POX site and is reduced to PGH_2 (not shown).

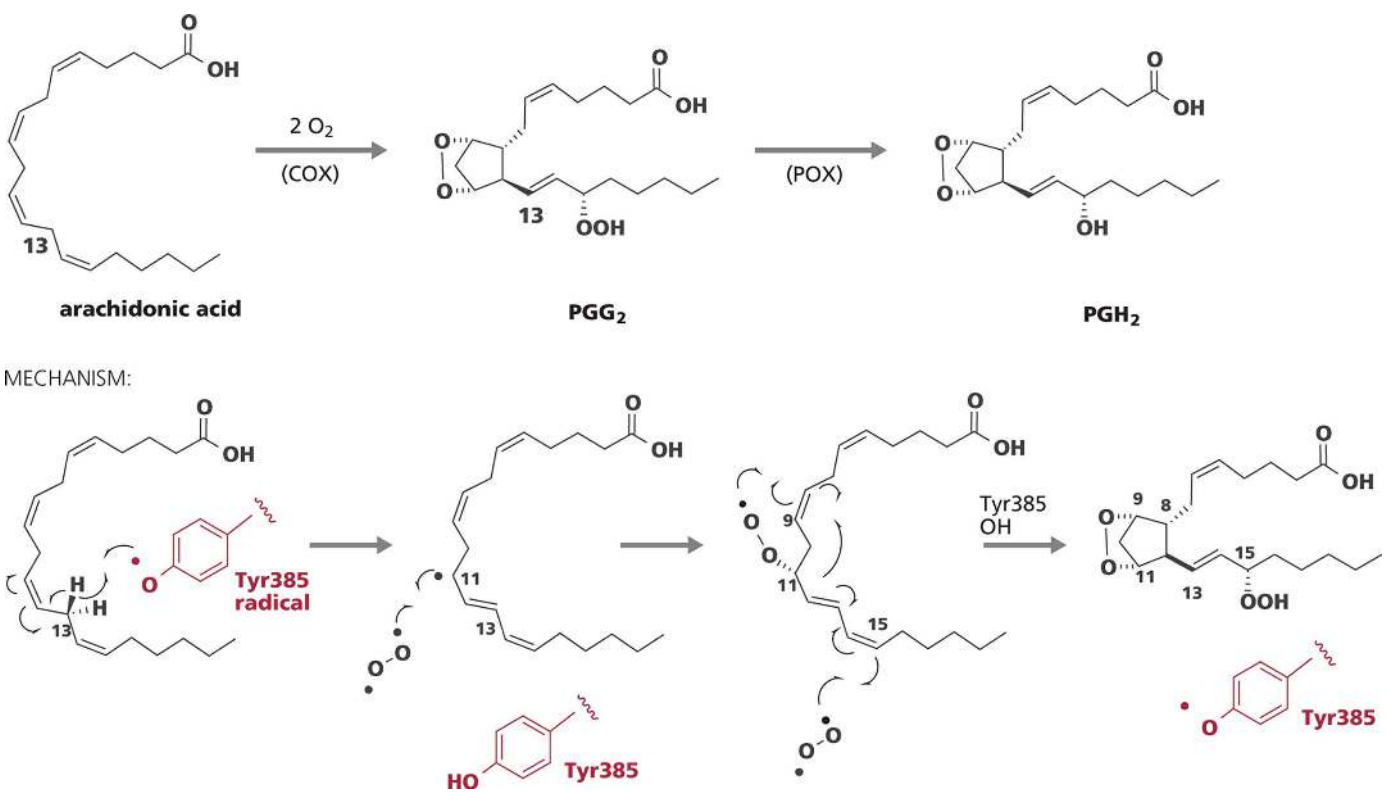


Figure 7.39 Cyclooxygenase-catalyzed conversion of arachidonic acid to PGH_2. Cyclooxygenase cyclizes and oxidizes arachidonic acid to form PGG_2, then POX mediates its reduction to generate PGH_2, a precursor to prostaglandins and thromboxane. The oxidation and cyclization mechanism involves the generation of a tyrosyl radical (red), which abstracts a specific H-atom from the substrate, followed by the addition of oxygen and then cyclization to an endoperoxide. Addition of oxygen at the C-15 position produces PGG_2 and regenerates the tyrosyl radical.

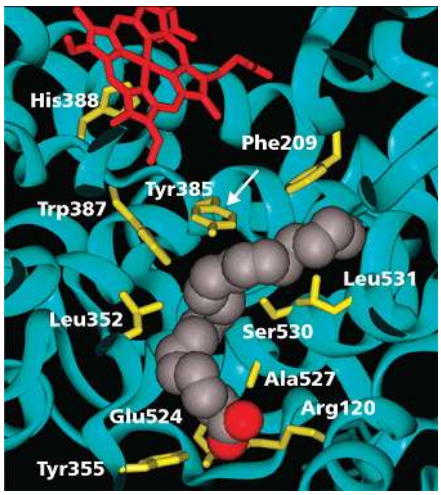


Figure 7.40 X-ray crystal structure of arachidonic acid bound in the active site of COX-1. Arachidonic acid is shown in gray, with the carboxylate oxygens in red; key enzyme residues involved in substrate binding are in yellow; and the heme unit that generates the initial tyrosyl radical is in red. The carboxylate interactions with Arg120 and Tyr355 are seen at the bottom of the figure. Numerous hydrophobic residues (such as Tyr and Leu) line the binding pocket to accommodate the hydrophobic arachidonic acid chain. Tyr385 (white arrow) is adjacent to the heme moiety and in proximity to the C13 proton.

Source: Adapted from Blobaum, A. L., and Marnett, L. J. (2007). *J Med Chem* 50:1,425–1,441. With permission from the American Chemical Society.

The membrane-bound enzyme is a homodimer of two 70 kDa subunits, and the catalytic site is at the end of a deep hydrophobic binding pocket. Important binding interactions between the protein and the substrate include ionic interactions between the carboxylic acid of the substrate and Arg120 and hydrogen bonding between the acid and Tyr355. The remaining interactions are hydrophobic, with the substrate aligning in the hydrophobic pocket to orient the C-13 hydrogen near the Tyr385 radical, as shown in **Figure 7.40**.

Inhibitors of COX (both isoforms 1 and 2) are used as anti-inflammatory drugs and comprise the family known as nonsteroidal anti-inflammatory drugs (NSAIDs). The oldest member of this family is aspirin; it has been in use as a single agent since the late 1800s, but its medicinal properties have been known for thousands of years as a constituent of willow bark (see Chapter 1). Others include arylacetic acids, such as flufenamic acid, and arylpropionic acids, such as ibuprofen and naproxen (**Figure 7.41a**). All these inhibit both isoforms of the enzyme and are reversible competitive inhibitors, with the exception of aspirin. Aspirin irreversibly acetylates Ser530 in the enzyme's binding pocket. This mechanism leads to the use of low doses of aspirin as an anti-thrombotic due to the irreversible inhibition of COX in platelets. Since platelets cannot form new enzymes, they are inhibited for the lifetime of the platelet, resulting in a lowered production of thromboxane A2, which then prevents platelet aggregation.

Some prostaglandins have a protective role in the stomach lining, and side effects of the nonselective NSAIDS, since they reduce the production of prostaglandins, include increased incidence of ulcers. Because of this, when the COX-2 isoform was identified in 1991, there was significant effort put into developing selective COX-2 inhibitors based on the rationale that the inhibition of the widely distributed COX-1 by the nonselective inhibitors was responsible for their gastrointestinal side effects. One key difference in the two isoforms is a change from Ile523 in COX-1 to the smaller Val523 in COX-2, as revealed in the X-ray crystal structures of the two isoforms (shown in **Figure 7.41b**). This small change reveals an additional binding pocket (blue), allowing for selectivity in drugs that contain a group that can occupy it.

The first selective COX-2 inhibitor reported was an experimental compound from DuPont Merck, known as DuP-697. This compound had a lower incidence of formation of stomach ulcers in animal models and lent credence to the value of selective inhibition. This diaryl heterocycle was a derivative of two older anti-inflammatory compounds, phenylbutazone and oxaprozin. Other compounds in this class include the first COX-2 selective inhibitor to be marketed, rofecoxib (Vioxx), approved by the FDA in 1999, soon followed by celecoxib and valdecoxib. All have in common two aromatic rings on adjacent sites on a central scaffold and a sulfonamide or methylsulfone substituent on one of the phenyl rings that extends into the selectivity pocket. While effective, and with a lower incidence of gastrointestinal side effects, a higher incidence of cardiovascular events, such as heart attack and stroke, was observed. This side effect has led to the withdrawal of most of the COX-2 selective inhibitors, with only celecoxib still on the market.

While there are other examples of drugs acting at oxidative and reductive enzymes, the two classes described here comprise two of the most successful examples of drug design. As stated previously, the mechanisms of the specific enzymes within this class vary more widely than in other classes, like the proteases and kinases. As such, general strategies for drug design are not possible and most often rely on understanding the specific action of the target enzyme.

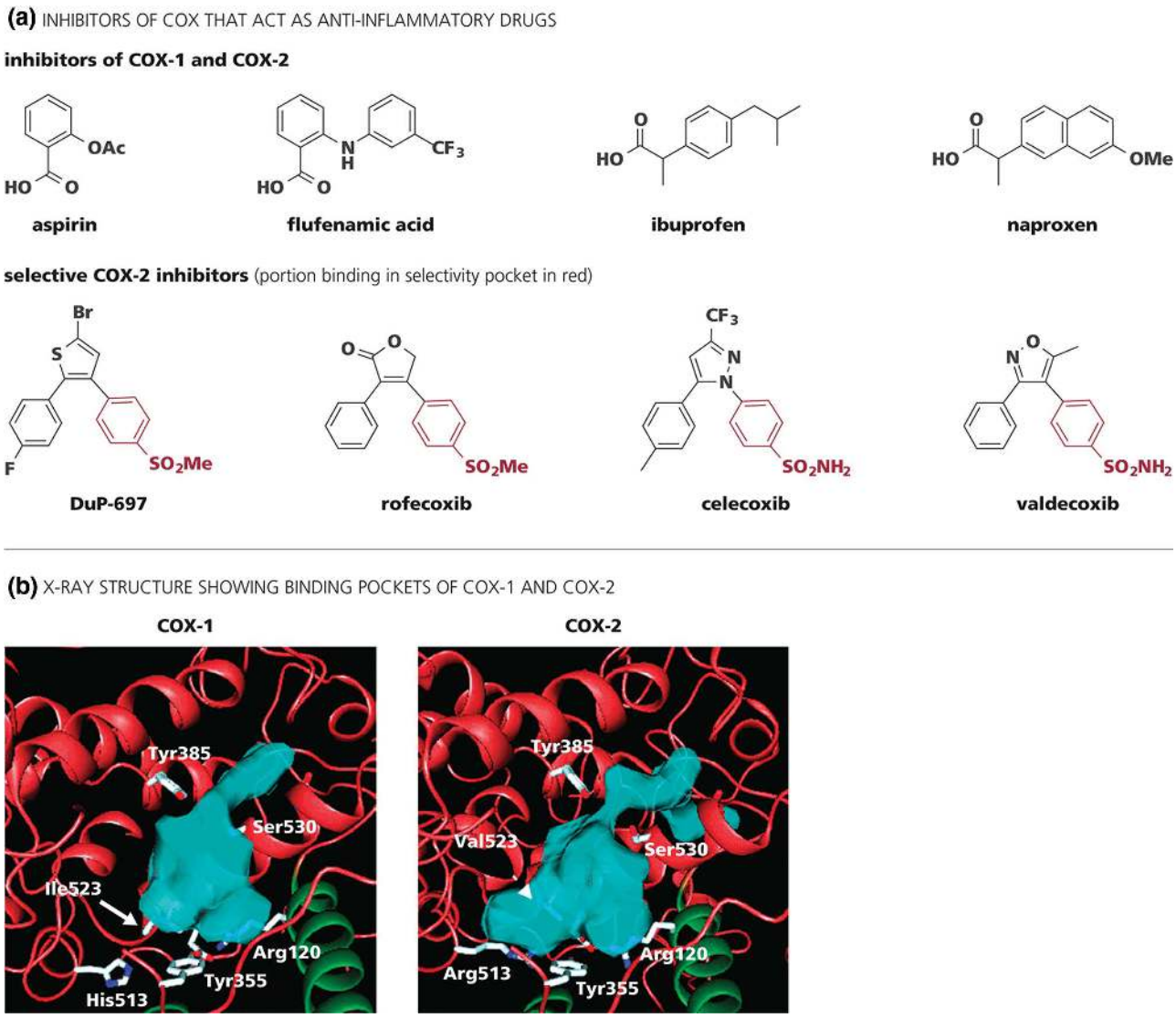


Figure 7.41 Structures of NSAIDs that are COX inhibitors and X-ray structures of COX-1 and COX-2. (a) Structures of NSAIDs that act as COX inhibitors. Aspirin, flufenamic acid, ibuprofen, and naproxen act at both COX-1 and COX-2; DuP-697, rofecoxib, celecoxib, and valdecoxib are representatives of selective COX-2 inhibitors. Substituted phenyl groups that instill selectivity are highlighted in red. (b) X-ray structures of active sites of COX-1 and COX-2. Binding pockets are shown in blue. The change of an Ile523 in COX-1 to Val523 in COX-2 opens the selectivity pocket (to left of Val523), where the aryl groups (red) in the selective inhibitors bind.

Source: (b) Adapted from Blobaum, A. L., and Marnett, L. J. (2007). *J Med Chem* 50:1,425–1,441. With permission from the American Chemical Society.

7.8 SUMMARY

There are numerous enzymes, each carrying out specific reactions through unique mechanisms, that are implicated in diseases. Likewise, there are numerous enzyme modulators that are effective drugs that work via various mechanisms and target multiple therapeutic indications. In fact, approximately 10% of all drugs act as kinase inhibitors, with the percentage that inhibit any enzyme likely reaching the 30–40% range. While the scope of all these possibilities cannot be covered in this text, there are several recurrent themes. First, enzymes are considered druggable targets. Their mechanisms are usually well-understood and can be studied by various methods. Small-molecule modulators of enzymes can be readily characterized by kinetics as well as by biophysical measurements. While a number of important drugs that work through enzyme inhibition were discovered via phenotypic screening, more modern approaches rely on a detailed knowledge of the enzyme's mechanism, using the principles of medicinal chemistry and structure-based drug design. The importance of enzymes as drug targets, and the success at developing drugs targeting them, is reflected in the observation that, as of 2017, kinases and proteases make up approximately 10% of the drug targets, whereas in 2016, that number was a mere 2%.

CASE STUDY 1 DISCOVERY OF THE PROTEIN TYROSINE KINASE INHIBITOR IMATINIB

Imatinib, an inhibitor of the kinase Bcr-Abl, is considered a breakthrough therapy. After its introduction, survival from chronic myelogenous leukemia (CML) reached normal life expectancy, while prior treatment was typically a bone marrow transplant for the few eligible patients and a 30% 5-year survival rate. The example of imatinib highlights the power of basic scientific discoveries in cancer biology coupled with medicinal chemistry and iterative rational drug design.

In the 1970s, researchers identified a specific genetic abnormality in patients diagnosed with CML: a reciprocal translocation between two chromosomes that results in the Philadelphia chromosome. This finding was the first that showed a direct connection between genetic changes and cancer. The protein product of the abnormal gene, Bcr-Abl, is a tyrosine kinase with constitutive, increased activity compared to the non-transformed gene product and was shown to be essential for cellular transformation in CML. By inhibiting Bcr-Abl, a kinase that is not present in untransformed cells, a drug would be possible that was highly specific for cancer cells containing the Philadelphia chromosome.

The starting point for the drug discovery program was the phenylaminopyrimidine core (**Figure 7.42**), identified in a screen for inhibitors of a related kinase, protein kinase C (PKC). This scaffold exhibited *in vitro* activity against several different kinases including PKC, possessed good lead-like properties, and was amenable to further synthetic chemistry elaboration.

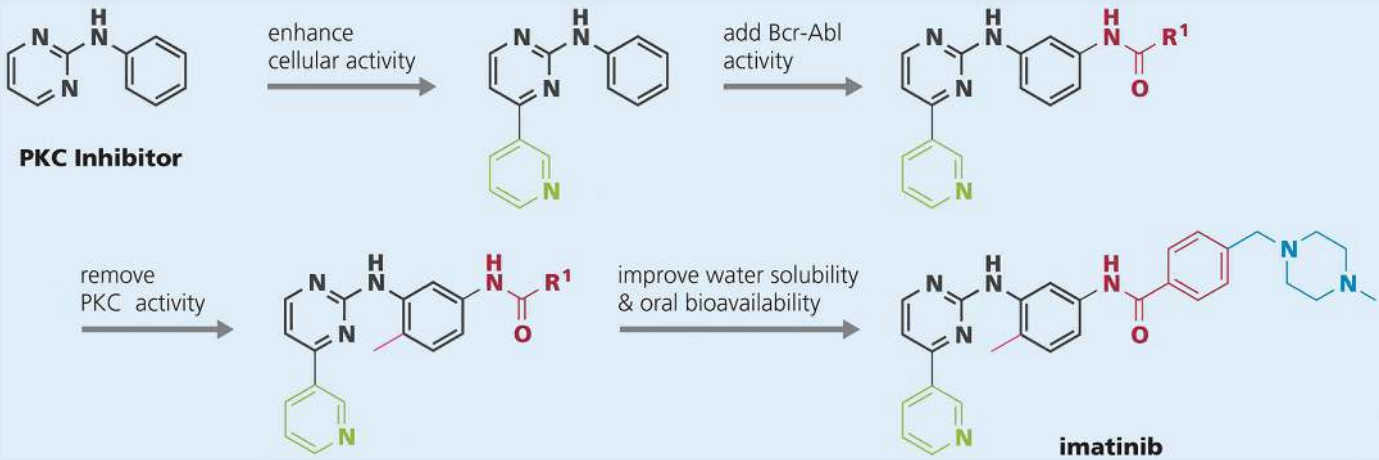


Figure 7.42 Structures in the development of imatinib. The PKC inhibitor (black) was elaborated by adding a 3-pyridyl group (green), which enhanced cellular activity. The addition of an amide group (red) broadened the activity to include Bcr-Abl, while selectivity for Bcr-Abl over PKC was achieved by the addition of a 6-methyl group (pink). The addition of the *N*-methylpiperazine moiety (blue) improved bioavailability and led to imatinib.

The incorporation of a 3-pyridyl group (green) onto the scaffold enhanced the cellular PKC activity. A key finding was that further elaboration to include an amide group (red) extended the kinase inhibition profile to Bcr-Abl kinase, while the addition of a methyl group (pink) at the 6-position of the diaminophenyl group abolished the PKC activity but retained Bcr-Abl potency. At this stage, potent selective inhibitors were in hand; however, pharmaceutical properties such as solubility and toxicity were of concern. To address those issues, a methylene-*N*-methylpiperazine moiety (shown in blue) was incorporated. The rationale for the addition of this group included using the polar *N*-methylpiperazine to improve solubility and incorporating a methylene between the piperazine and phenyl group to avoid a potentially mutagenic aniline functionality.

The imatinib–Bcr-Abl co-crystal structure shows that the molecule binds to the conserved nucleotide-binding pocket when the protein is in the inactive DFG-out

conformation (a type II inhibitor), thereby occluding access of the substrate to the active site. Some of the key interactions between Bcr-Abl and imatinib are shown in Figure 7.43. Some essential interactions that explain potency and selectivity include the following:

- The nitrogen of the pyridine ring makes an H-bond to Met318. This is the same interaction made by adenine N1 in the ATP substrate.
- The secondary amino group makes an H-bond to Thr315; Thr315 is relatively unique to Bcr-Abl. Most other kinases contain a methionine at this site, which is not capable of H-bonding. This interaction provides potency and selectivity for Bcr-Abl.
- Numerous snug van der Waals (hydrophobic) interactions between hydrophobic residues and the aromatic rings of the inhibitor are evident.
- The amide makes H-bonds in the DFG loop to Asp381 and also to Glu286 on helix C.

Since its introduction, imatinib has been shown to inhibit other kinases, despite its original design for selectivity. In addition, resistance toward imatinib has emerged, and this has motivated the design and synthesis of other Bcr-Abl inhibitors.

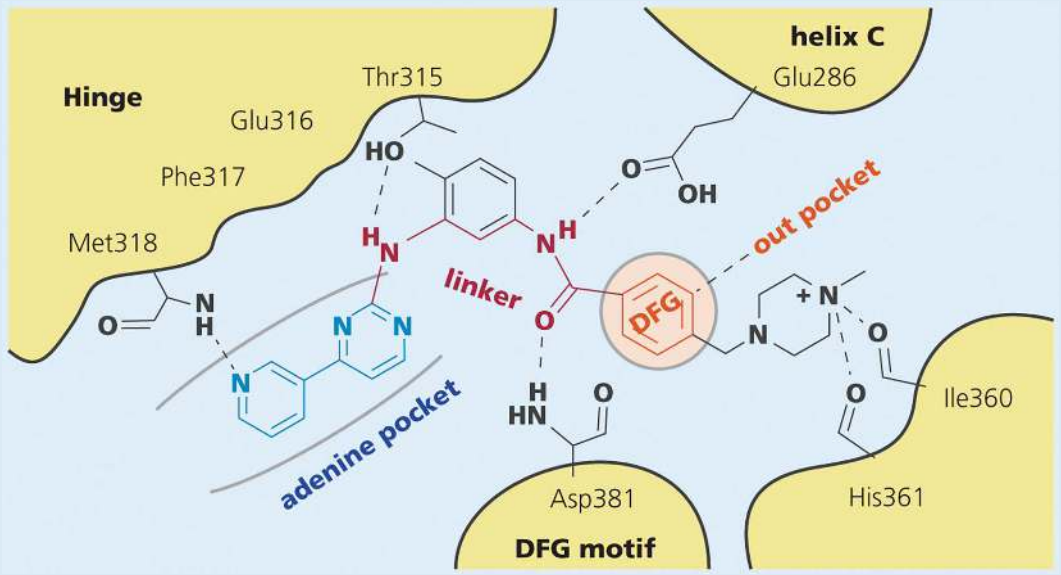


Figure 7.43 Key binding interactions of imatinib. These include an H-bond from the 3-pyridyl nitrogen of imatinib to Met318 NH, an H-bond from the secondary amine of imatinib to the side chain of Thr315 (unique to Bcr-Abl), and H-bonds from the amide carbonyl of imatinib to the backbone of Asp381 and from the amide NH of imatinib to Glu286.

Source: Adapted from Pan, X.; Dong, J.; Shi, Y.; et al. (2015). *Org Biomol Chem* 13:7,050–7,066. With permission from the Royal Society of Chemistry.

CASE STUDY 2 DISCOVERY OF GRAZOPREVIR, AN HCV NS3/4A PROTEASE INHIBITOR

Hepatitis C (HCV) is a chronic viral disease of the liver, which affects nearly 200 million people worldwide and can lead to death by cirrhosis or liver cancer. Until the early 2010s, the only treatment was a combination of ribavirin and pegylated interferon-α taken for 48 weeks. Ribavirin is a nucleoside polymerase inhibitor, while interferon-α is a peptide cytokine that prevents viral replication. However, that combination therapy was only effective in about 50% of patients and had serious side effects, such as anemia, neutropenia, and thrombocytopenia, that made adherence to the long treatment regimen difficult.

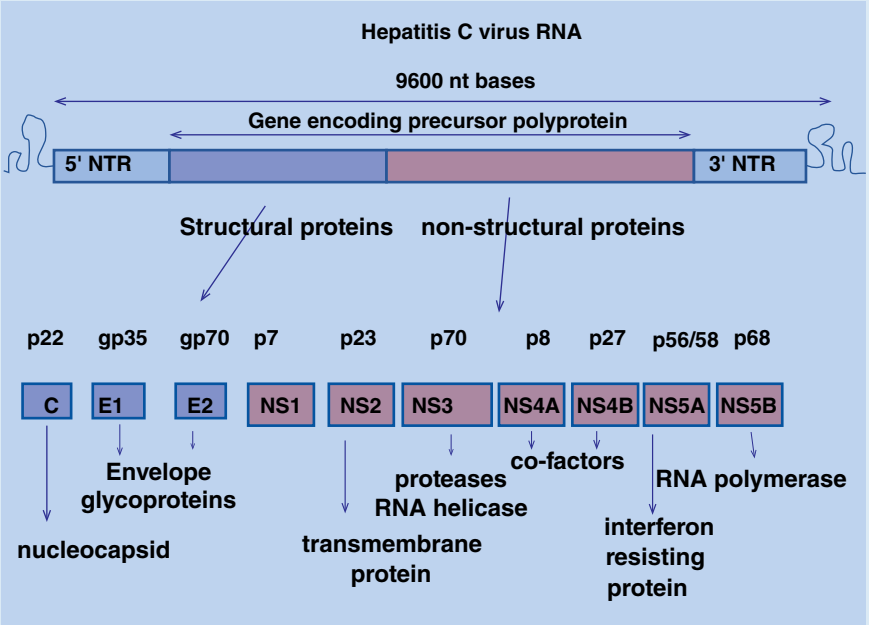


Figure 7.44 **Hepatitis C genome and encoded proteins (NS proteins are non-structural).**

Source: Wikipedia.

The search for more effective drugs with fewer side effects began in earnest after 1989, when HCV's genome was sequenced and found to encode a number of potentially druggable enzymatic viral targets (**Figure 7.44**). One of the most promising targets was named HCV NS3/4A. The NS3 protein is a chymotrypsin/trypsin-like serine protease, and NS4A is its co-factor. This protease cleaves the initially translated polypeptide at multiple sites to produce four different viral enzymes, including other proteases, a helicase, and a polymerase, all of which are required for HCV viral replication. A number of organizations successfully initiated programs to target this protease relying on structure-based design. This case study describes one such example, the development of grazoprevir by Merck, which was approved by the FDA in 2016 for use in combination with elbasvir, an NS5A protein inhibitor.

Design Strategy

A substrate for NS3/4A is a dodecapeptide, and an early finding was that the N-terminal hexapeptide cleavage product served as an inhibitor of the NS3/4A protease. The N-acetyl hexapeptide (Ac-DDIVPC-OH) was the basis for drug design at Boehringer Ingelheim, which resulted in the development of clinical candidate BILN-2061 (**Figure 7.45**). This macrocyclic inhibitor contains a hydrophobic linkage between the P1 site (at the C-terminal site) and the P3 site that mimics the valine in the hexapeptide, and a large heterocyclic ether substituent on the 4 position of the proline at P2 (red in **Figure 7.45**). This compound progressed to Phase II clinical trials but was discontinued due to cardiotoxicity.

At Merck, scientists used molecular modeling and noticed that when docked to HCV NS3 alone, the thiazolyl quinoline in BILN-3061 did not make any productive biomolecular interactions (**Figure 7.46**). This moiety evolved early on in Boehringer Ingelheim's design of BILN-2061 from their finding that an aryl substitution at that position conferred a 400-fold increase in binding affinity. However, modeling of BILN-2061 in the HCV NS3/4A complex, rather than just NS3, showed a cation–π interaction of the quinoline with Gln526, and an H-bonding interaction between the carbamate carbonyl at P4 and His528 of the helicase region of the co-factor NS4A, as well as a pocket that

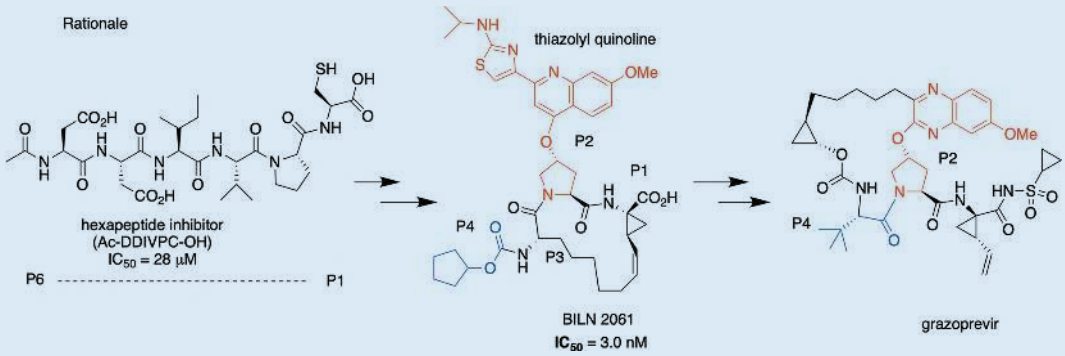


Figure 7.45 Compounds leading to the development of grazoprevir. The N-acetyl-hexapeptide was the earliest-identified inhibitor of HCV NS3/4A. Adding a constraint (shown in black) from P1 to P3 and a large P2 proline aryl group (in red) led to the macrocyclic clinical candidate BILN 2061. Changing the position of constraint from P1–P3 to P2–P4 (Ile mimic in blue) led to grazoprevir.

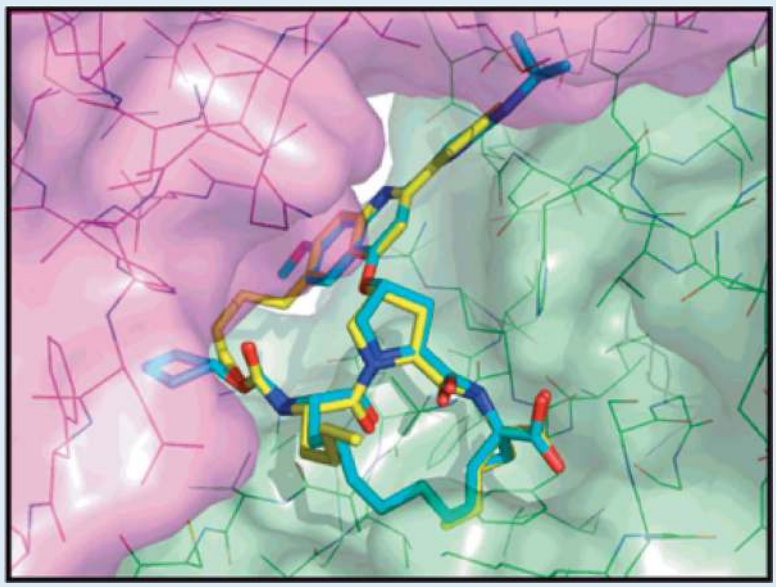


Figure 7.46 Model of BILN-2061 (blue) overlaid with a proposed P2- to P4-linked macrocycle (yellow), bound to the NS3/4A protein. The green region is the protease domain of NS3, and purple is the helicase region of 4A.

Source: Liverton et. al. *J. Am. Chem. Soc.* 2008, 130, 4,607–4,609.

could accommodate a chain linking the P2 and P4 sites. They proposed a macrocyclic bridge between these two features, which would provide a conformationally low-energy state that would reach these key features and also provide novel compounds.

Synthesis and investigation of early analogs with different-sized carbamate macrocycles found that a 5-carbon link between the quinoline and carbamate was most potent. In an enzyme inhibition assay, the compound where $n = 1$ showed relatively poor activity, with $n = 3$ showing low nM potency (compounds **1a–1c** in Figure 7.47). A key finding was that the point of attachment to the quinoline ring at C6 was found to be important for potency, with an analog of **1c** linked at C5 (compound **2**) much less active.

Other changes improved activity further, resulting in grazoprevir: replacing the n-butyl chain with *tert*-butyl, and bioisosteric replacement of the acid (corresponding to P1) with a cyclopropyl acyl sulfonamide.

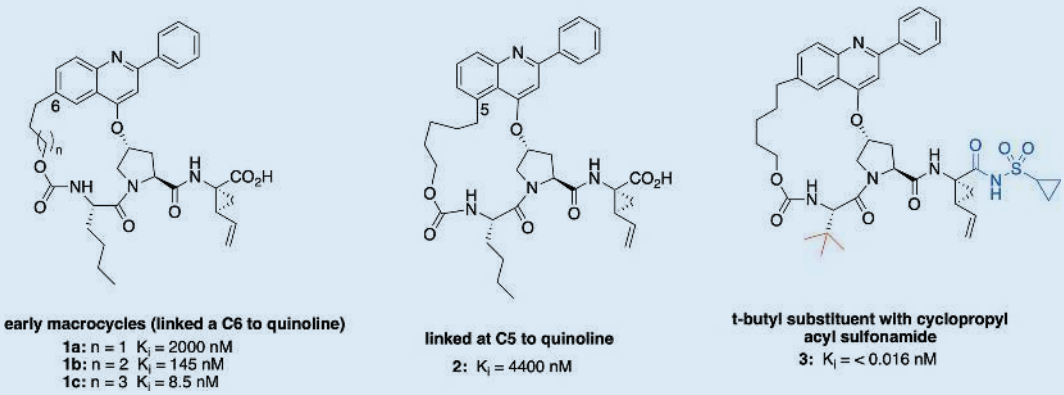


Figure 7.47 Structures of early leads. Investigation of chain length of C6-carbamate shows *n* = 3 is optimal; linkage to C5 (compound 2) is much less active than the C6-linked compound, and the analog with a *tert*-butyl (red) vs. n-butyl and cyclopropyl acyl sulfonamide (blue) replacing the acid was optimal.

Final modifications of this advanced lead included the introduction of further conformational constraint using a cyclopropyl ring in the P2–P4 linkage chain, which results in broad activity against various genotypes. The addition of a second nitrogen atom to the quinoline afforded a less-basic quinoxaline heterocycle, which improved the aqueous solubility of its zwitterionic form. In addition to enzyme inhibition, compounds were assayed in a cell-based HCV replicon assay and showed potent effects. Pharmacokinetics in rats and dogs showed good oral exposure, as well as liver exposure, with half-life of 1.4 hours (rat) and 3 hours (dog).

Process Synthesis of Grazoprevir

The initial synthesis of grazoprevir relied on ring-closing metathesis reaction for the formation of the 18-membered ring macrocycle, but there were some challenges on scaling up, including instability of the bis-alkene precursor. The introduction of the cyclopropyl onto the linking chain was also problematic. The process synthesis, outlined in Figure 7.48, took a different approach and used macrolactamization to form the ring, which solved those early problems.

The synthesis developed was modular, also contributing to efficiency. Synthesis of the prolyl heteroaryl ether relied on an S_NAr substitution between a protected hydroxyproline and dichloroquinoxaline to afford the coupled product **4** with very high selectivity. Coupling of t-butyl glycine with the elaborated cyclopropyl alcohol afforded the alkyne **5**. Coupling of the alkyne **5** to the aryl chloride **4** was done using a Sonagashira reaction to give the macrocylization substrate **6**. Hydrogenation of the alkyne, and Boc deprotection, was followed by macrolactamization to afford the macrocycle **7**. The final steps were saponification of the methyl ester of **7** and then coupling with the cyclopropyl acyl sulfonamide to afford grazoprevir with 99.7% purity and in 52% overall yield.

Since the introduction of the HCV NS3/4A protease and NS5A protein inhibitors, treatment of hepatitis C has changed dramatically. These drugs, also known as DAAs (direct-acting antivirals), are typically used as combinations and are now the most common treatment for hepatitis C infections, with cure of the disease in about 95% of all cases. Treatment regimens are much shorter (12–16 weeks) and produce less-severe side effects than the ribavirin/interferon combination. There are now five drug combinations commonly used, with one being grazoprevir and elbasvir (Zepatier), approved in 2016.

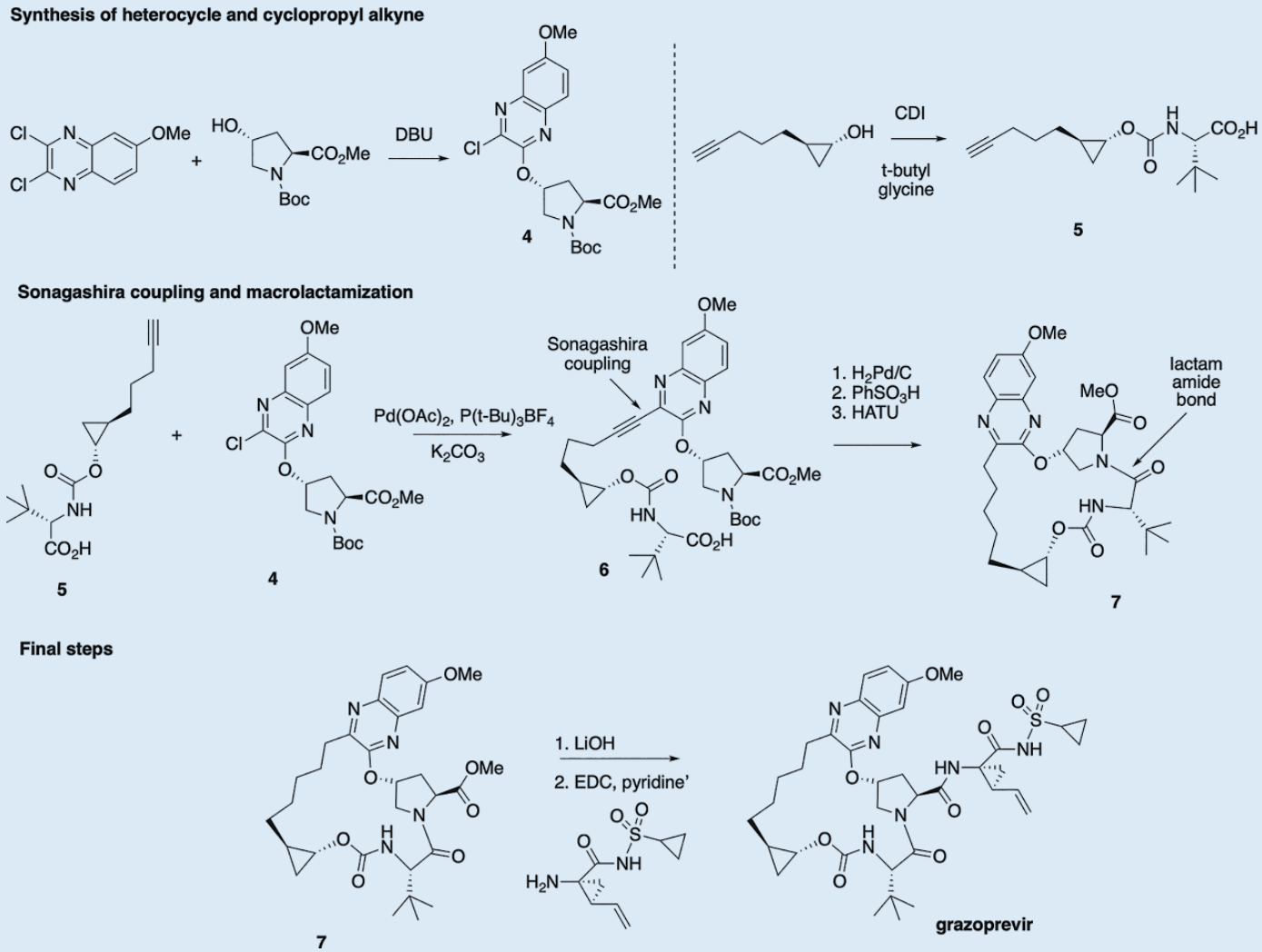
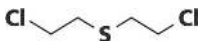


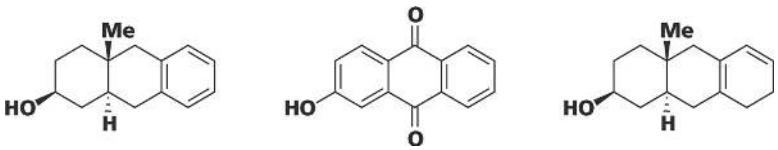
Figure 7.48 **Key steps in the process synthesis of grazoprevir.**

7.9 REVIEW QUESTIONS

1. Describe the difference between an allosteric inhibitor and an active site inhibitor and the advantages and disadvantages of each.
2. Explain why transition state mimetics are a viable strategy to inhibit aspartyl proteases.
3. Explain why transition state mimetics are not often used as a strategy for kinase inhibitor design.
4. Draw an abbreviated mechanism of inhibition of a serine protease by an aldehyde.
5. List four reasons that enzymes are considered druggable targets.
6. Explain the lock and key model, and compare it to the induced-fit model. What observations have been made to favor the induced-fit model?
7. Describe the features of the ATP binding site of a kinase and the different types of kinase inhibitors.

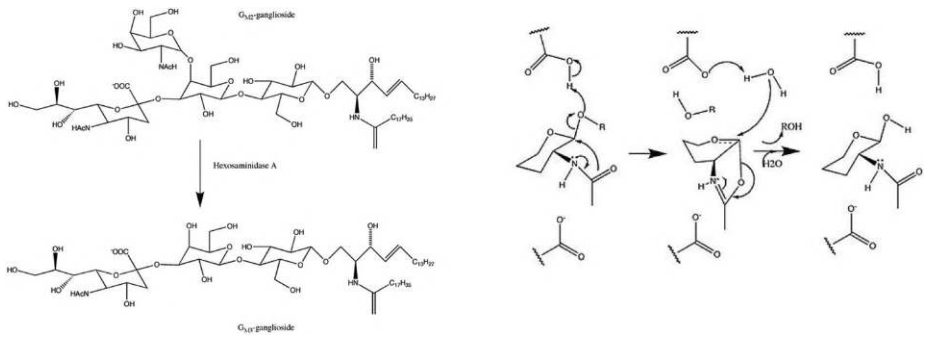


8. Explain proximity and orientation effects in the binding of an inhibitor to an enzyme.
9. How does the measured inhibition of a competitive inhibitor change with increasing concentration of a substrate?
10. Of the compounds described in this chapter, find one example each of incorporating solubilizing groups, isosteres, conformational constraint, and privileged structures.



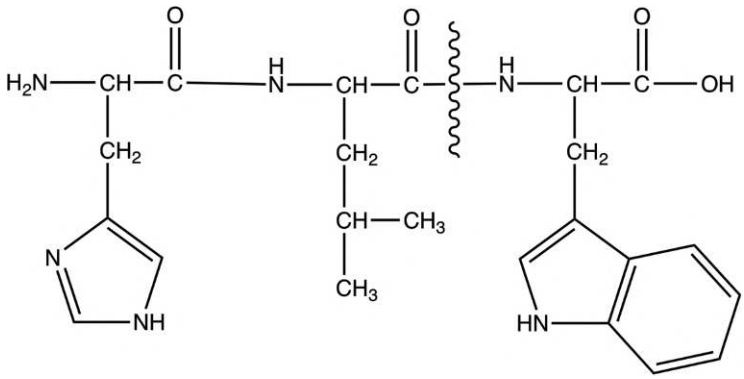
7.10 APPLICATION QUESTIONS

1. Hexosaminidase A is an enzyme that hydrolyzes certain carbohydrate residues. The mechanism of the reaction it catalyzes is shown here. Design an inhibitor of this enzyme based on the reaction mechanism. Explain the rationale for your design.

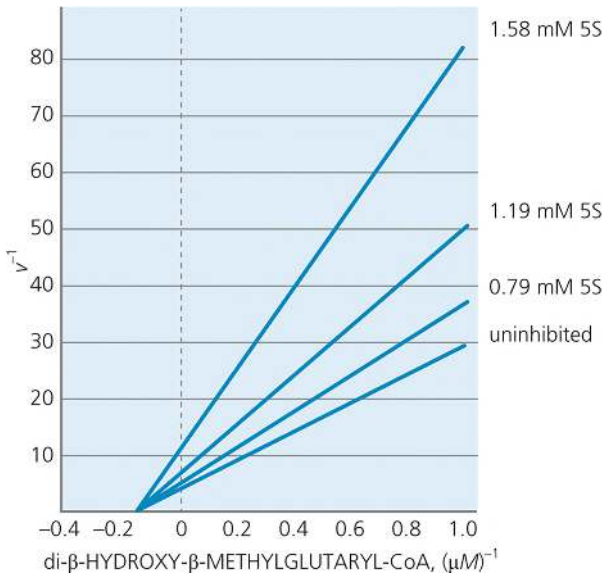


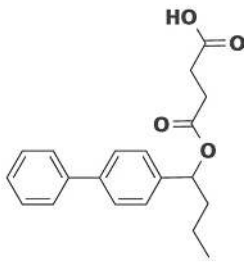
2. Look up the structure of an HIV reverse transcriptase phosphate prodrug. Draw its structure, and explain its mechanism and how it differs from those shown in **Figure 7.28**.
3. What hit identification strategies can usually be applied to enzyme inhibitor drug discovery? Explain why these strategies can be particularly successful. What are the similarities and differences in approaches to hit identification when working with enzymes vs. receptors?

4. A theoretical aspartyl protease cleaves this substrate at the scissile bond shown with the wavy line. Design three inhibitors based on this substrate.



5. Research the structure of staurosporine. What type of enzyme does it inhibit? What are the binding interactions it makes with the target? And compare compounds described in this chapter.
6. Using the PDE3 pharmacophore model in **Figure 7.32** and the table of bioisosteres in Chapter 4 (see Tables 4.3 and 4.4), design three compounds that may be PDE3 inhibitors.
7. This plot shows the effects of compound **5S** on the enzyme HMG-CoA reductase. Compound **5S** was one of a series of early compounds prepared as inhibitors of the enzyme. (From Boots, M. R.; Boots, S. G.; Noble, C. M.; and Guyer, K. E. (1973). *J Pharm Sci* 62:952–957. With permission from the Wiley Periodicals Inc. and the American Pharmacists Association.)





compound "5S"
note: S = succinate,
not stereochemistry
(compound is racemic)

8. What is the name of this type of plot? Is the inhibitor competitive or noncompetitive? What portions of **5S** correspond to the model for synthetic analogs of statins (see **Figure 7.38**)?
9. A number of the drugs described in this chapter contain fluorine atoms. Identify three of them, and on the basis of the concepts in Chapter 4, hypothesize what role the fluorine may be playing in these molecules.
10. Apply Lipinski rules to Lipitor, indinavir, Gleevec, and sildenafil. Do they comply? (Online tools can be used to calculate log*P*.)

FURTHER READING

General References

Koshland DE (1994) The key-lock theory and the induced fit theory. *Angew Chem Int Ed* 33:2375–2378.

Robertson JG (2005) Mechanistic basis of enzyme-targeted drugs. *Biochemistry* 44:5561–5571.

Kinases

Arter C, Trask L, Ward S, Yeoh S & Bayliss R (2022) Structural features of the protein kinase domain and targeted binding by small-molecule inhibitors. *J Biol Chem* 298(8):102247.

Attwood MM, Fabbro D, Sokolov AV et al. (2021) Trends in kinase drug discovery: targets, indications and inhibitor design. *Nat Rev Drug Disc* 20:839–861.

Clark JD, Flanagan ME & Telliez J-B (2014) Discovery and development of Janus kinase (JAK) inhibitors for inflammatory diseases. *J Med Chem* 57:5023–5038.

Nobele MEM, Endicott JA & Jonson LN (2004) Protein kinase inhibitors: insights into drug design from structure. *Science* 303:1800–1805.

Roskoski Jr R (2024) Properties of FDA-approved small molecule protein kinase inhibitors: a 2024 update. *Pharmacol Res* 200:107059.

Zhang J, Yang PL & Gray NS (2009) Targeting cancer with small molecule kinase inhibitors. *Nat Rev Cancer* 9:28–39.

Proteases

Erez E, Fass D & Bibi E (2009) How intramembrane proteases bury hydrolytic reactions in the membrane. *Nature* 459:371–378.

Hernandez AA & Roush WR (2002) Recent advances in the design, synthesis and selection of cysteine protease inhibitors. *Curr Opin Chem Biol* 6:459–465.

Njoroge FG, Chen KS, Shih N-Y & Piwinski JJ (2008) Challenges in modern drug discovery: a case study of boceprevir, an HCV protease inhibitor for the treatment of hepatitis C virus infection. *Acc Chem Res* 41:50–59.

Polymerases

Esposito F, Corona A & Tramontano E (2012) HIV-1 reverse transcriptase still remains a new drug target: structure, function, classical inhibitors, and new inhibitors with innovative mechanisms, of actions. *Mol Biol Int* 2012:586401.

Steitz TA (1998) A mechanism for all polymerases. *Nature* 391:231–232.

Imatinib

Capdeville R, Buchdunger W, Zimmermann J & Matter A (2002) Glivec (STI571, Imatinib), a rationally developed, targeted anticancer drug. *Nat Rev Drug Discov* 1:493.

Schindler T, Bornmann W, Pellicena P et al. (2000) Structural mechanism for STI-571 inhibition of Abelson tyrosine kinase. *Science* 289:1938–1942.

Grazoprevir

Harper S, McCauley JA, Rudd MT et al. (2012) Discovery of MK-5172, a macrocyclic hepatitis C virus NS3/4a protease inhibitor. *ACS Med Chem Lett* 3:332–336.

Kuethe J, Zhong Y-L, Yasuda N et al. (2013) Development of a practical, asymmetric synthesis of the hepatitis C virus protease inhibitor MK-5172. *Org Lett* 14:4174–4177.

Liverton NJ, Holloway MK, McCauley JA et al. (2008) Molecular modeling based approach to potent P2-P4 macrocyclic inhibitors of hepatits C NS3/4A protease. *J Am Chem Soc* 130:4607–4609.

Phosphodiesterases

Manallack DT, Hughes RA & Thompson PE (2005) The next generation of phosphodiesterase inhibitors: structural clues to ligand and substrate selectivity of phosphodiesterases. *J Med Chem* 48:3449–3462.

Rotella D (2002) Phosphodiesterase 5 inhibitors: current status and potential applications. *Nat Rev Drug Discov* 1:674–685.

Spina D (2008) PDE4 inhibitors: current status. *Br J Pharmacol* 155:308–315.

Oxidoreductases

Blobaum AL & Marnett LJ (2007) Structural and functional basis of cyclooxygenase inhibition. *J Med Chem* 50:1425–1441.

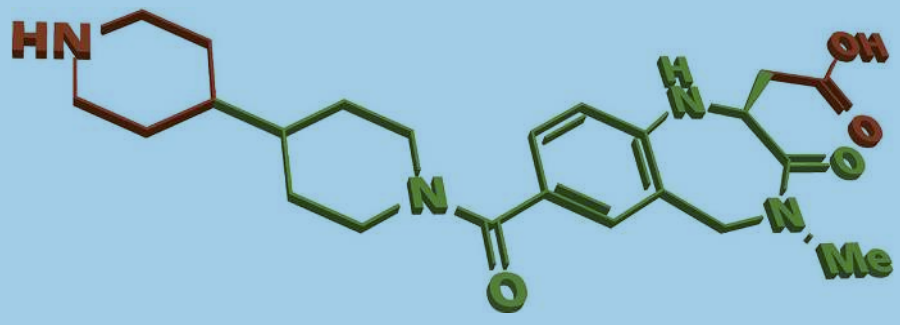
Flower RJ (2003) The development of COX2 inhibitors. *Nat Rev Drug Discov* 2:179–191.

Roth BD (2002) The discovery and development of atorvastatin, a potent novel hypolipidemic agent. *Prog Med Chem* 40:1–22.

Tolbert JA (2003) Lovastatin and beyond: the history of the HMG-CoA reductase inhibitors. *Nat Rev Drug Discov* 2:517–526.

Media Links

U.S. FDA approved protein kinase inhibitors. http://www.brimr.org/PKI/PKIs.htm; https://brimr.org/protein-kinase-inhibitors/



Protein–Protein and Lipid Structure Interactions as Drug Targets

8

LEARNING OBJECTIVES

- Understand why atypical drug targets, such as protein–protein interactions and lipid structures, are more challenging than traditional receptor and enzyme targets.
- Gain familiarity with the general strategies used for hit identification for these targets.
- Identify several different mechanisms by which small molecules can modulate protein–protein interactions.
- Develop an understanding of how molecular glues and bifunctional degraders work.
- Learn examples of drugs that work by targeting protein–protein interactions and by interacting with lipids and cell membranes.

8.1 INTRODUCTION

Receptors and enzymes constitute the majority of small-molecule drug targets. This focus may, in part, be due to the fact that many of these proteins' functions rely on the binding of small-molecule natural ligands and substrates. In these cases, drug design and discovery efforts can model drugs after those small molecules. Additionally, as discussed in Chapters 6 and 7, the three-dimensional structures of receptors and enzymes have inherent features that make their function amenable to modulation by small molecules. However, receptors and enzymes are responsible for only a fraction of the physiological processes that can be modulated in order to mediate disease states. Many others are mediated by interactions between proteins, often two, but sometimes multiple proteins that interact to initiate or halt a biological process. These interactions offer an opportunity for intervention by small molecules, whether by inhibiting them, stabilizing them, de-stabilizing, or promoting novel protein–protein interactions that do not occur naturally or only occur rarely. Lipids perform an essential function in maintaining cell and membrane-bound protein structure and can also act as signaling molecules. Although few drugs act directly on lipids, the interruption of their structure or interactions is also a drug target. However, until recently, neither targeting protein–protein interactions nor lipid interactions has been common despite the fact that both play a pivotal role in virtually all physiological processes, such as cell adhesion, cell division, transport of molecules, maintenance of

DOI: 10.1201/ 9781003465607-10

Table 8.1 Examples of Drugs That Affect Protein–Protein and Lipid Structure Interactions, Their Specific Targets, Mechanism of Action, and Therapeutic Use

Macromolecular Target	Target Type	Mechanism of Action	Drug	Indication
Transthyretin	Protein	Stabilizes multimeric protein form	Tafamidis, acoramidis	Transthyretin-mediated amyloidosis
Tubulin	Protein	Destabilizes protein oligomer	Colchicine, vincristine	Cancer
Tubulin	Protein	Stabilizes protein oligomer	Paclitaxel, ixabepilone	Cancer
Calcineurin	Protein	Molecular glue	Cyclosporin	Anti-rejection therapy
FKBP + mTOR	Protein	Molecular glue	Rapamycin	Anti-rejection therapy
Cereblon	Protein	Molecular glue degrader	Lenalidomide/thalidomide	Leprosy, multiple myeloma, other cancers
Androgen receptor	Protein	Bifunctional degrader PROTAC	Bavdegalutamide (ARV-110)*	Cancer
Glycoprotein IIb/IIIa	Protein	Inhibits protein–protein interaction	Tirofiban	Cardiovascular disease
CCR5	Protein receptor	Inhibits protein–protein interaction	Maraviroc	HIV
Fungal ergosterol	Lipid	Sequesters lipids and/or forms pores	Amphotericin, nystatin	Fungal infection
Bacterial cell membranes	Lipids	Disrupts lipid membrane	Daptomycin	Bacterial infection

*Not yet FDA-approved.

cell structure and morphology, and signaling. In fact, these types of targets are often considered undruggable due to several factors:

- Binding between macromolecules, including those that mediate physiological effects, often occur at interfaces that are relatively flat and have large surface areas, which do not favor specific binding of a small, druglike molecule.
- There is poor understanding of the interactions that promote binary (and therefore selective) vs. oligomeric (and likely promiscuous) macromolecular structures and, consequently, a lack of understanding of how to promote or disrupt them.
- The hydrophobic nature of some of these targets, such as lipids, disfavors specific binding interactions.

Historically, of the small molecules that exert their effects through protein interactions or interactions with lipids in cell membranes, many have been natural products or their derivatives, and their mechanisms were determined only after their downstream biological effects were characterized. However, modern medicinal chemistry efforts are now actively focusing on this underexplored area, and there are a number of notable successes. Treating these atypical targets as one class is not possible, since they constitute diverse macromolecular structures, such as monomeric proteins involved in signaling, oligomeric protein assemblies that are part of specific cellular structures, and lipids that constitute cell membranes. Accordingly, the mechanisms by which drugs affect these systems are equally varied and include small molecules that interfere with or mimic binding between two proteins, stabilize or destabilize a multicomponent structure, promote binding between two different proteins that do not typically interact, and mediate pore formation in membranes. **Table 8.1** includes representative examples of such drugs.

8.2 PROTEIN–PROTEIN INTERACTIONS AS DRUG TARGETS

Numerous physiological processes are mediated by interactions between two or more proteins, termed **protein–protein interactions** (PPIs). For example, enzymes and nuclear receptors are regulated by their interactions with other proteins, and GPCRs

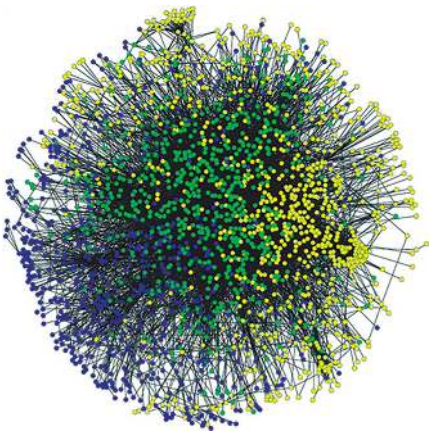


Figure 8.1 Representation of a portion of the human interactome. This map indicates known interactions between certain proteins. Each colored circle represents one protein; lines between circles represent interactions. In this map, the yellow region represents protein interactions associated with long QT syndrome, and the blue region represents protein interactions associated with epilepsy, with the green region where they overlap.

Source: Courtesy of the National Institute of General Medical Science, NIH, USA.

bind to G proteins and arrestins as part of the signal transduction process. In addition, many proteins that are not classified as either enzymes or receptors interact with themselves (homodimers and higher species) or with other proteins (heterodimers and higher species) to initiate or terminate a biological response or even to form a large macromolecular complex that exerts a distinct physiological effect. This process is further complicated by the multiple roles and interactions that can often be attributed to just one protein and by the fact that many of these interactions are transient and context-dependent, occurring only under specific conditions and in specific cellular locations. Efforts to delineate PPIs have led to databases of the **interactome**, a map of the network of molecular interactions within a particular cell or organism. Figure 8.1 is one representation of a portion of the human interactome, where each colored circle represents a specific protein and each line indicates an interaction between two proteins.

In Figure 8.1, one can see a vast number of known proteins and an enormous number of interactions, as exemplified by the many overlapping lines. In fact, it has been estimated that, among human proteins, between >50,000 and 650,000 binary protein interactions are involved in normal physiological events, but only a fraction of these have been carefully documented in such interactome maps. Accordingly, modulating these PPIs has the potential to be a powerful therapeutic strategy when those interactions mediate pathological processes. Figure 8.1 highlights this complexity by its representation of interactions associated with just two disorders, long QT syndrome and epilepsy. As another example, PPIs are involved in many cellular events associated with cancer, including regulating transcription factors as well as maintaining structural proteins, such as the microtubules that are necessary for cell division. If even half of the mapped PPIs are involved in disease states, then the opportunities for therapeutic intervention by disrupting or stabilizing the interaction significantly dwarf the number of more traditional enzyme and receptor drug targets. Furthermore, some of those PPIs can be "hijacked" by small molecules and re-directed from their typical, physiological (or pathophysiological) processes to interact in new ways that can have beneficial effects in disease states. This **chemically induced proximity (CIP)** expands the targetable proteome even further.

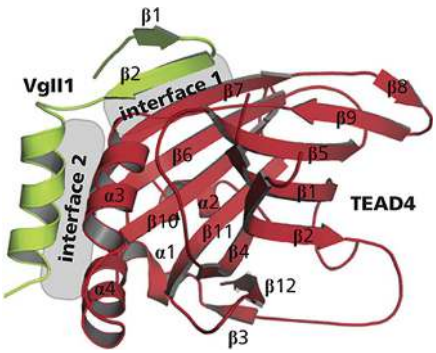


Figure 8.2 Ribbon diagram representation of the interface between two proteins, highlighting a typical flat and extended PPI binding surface. Transcription co-factor Vgll1 is shown in green, transcription factor TEAD4 is shown in red, and the interfaces between the two are depicted as gray cylinders. Interface 1 occurs between two β-sheets (β2 of Vgll1 and β7 of TEAD4), while interface 2 forms between the α-helix of Vgll1 and two α-helices (α3 and α4) of TEAD4.

Source: From Pobbati, A. V.; Chan, S. W.; Lee, I.; et al. (2012). *Structure* 20:1,135–1,140. With permission from Elsevier.

8.3 MODULATING PROTEIN–PROTEIN INTERACTIONS

One strategy for targeting protein–protein interactions is similar to more classical drug targets and involves either inhibiting the interaction between two protein partners (related to antagonism) or mimicking one of the protein partners (related to agonism). However, one of the more difficult aspects in modulating PPIs with small molecules is the need to cover a large, flat binding interface, as opposed to a small binding pocket typically found in more traditional targets, such as receptors or enzymes. In Figure 8.2, a ribbon diagram shows the transcription factor TEAD4 (red) bound to the transcription co-factor Vgll1 (green) and the large, flat extended surface area typical of PPI interfaces (depicted as gray cylinders) and highlights the challenges of using a small molecule to disrupt or mimic that interface.

In general, interactions between two proteins are a culmination of many weak interactions over a large surface, with an average area of ~1,500 Å^2. In contrast, the high-affinity binding of a small molecule to an enzyme or receptor relies on several strong interactions in a small, localized binding pocket. That three-dimensional binding pocket allows for the occupying ligand to make contacts in multiple directions, whereas a PPI interface is often only along one dimension. Consequently, molecules that interfere with PPIs usually have properties distinct from more classical inhibitors of enzymes or receptors. They are typically of higher molecular weight, a characteristic that is not usually considered druglike, and counterintuitively, they possess a greater degree of three-dimensionality. An alternate mechanism by which small molecules modulate PPIs is through binding to allosteric sites. Since protein conformations are dynamic and context-dependent, binding of a small molecule to only one conformation, or under just one set of conditions, may affect the protein's ability to adopt an active conformation and prevent binding to its partner. The protein's dynamic nature

may reveal allosteric sites for small-molecule binding that may not be apparent otherwise.

Methods to measure small molecules that interact with atypical protein targets are like those used for receptors (see Chapter 6) and rely on the evaluation of binding as a primary assay. When the macromolecule is isolable, common methods to detect binding of the small molecule, or a tagged derivative, include those that measure effects related to differences in size or other physical characteristics between the bound complex and the unbound constituents. Fluorescence-based assays, such as fluorescence polarization (FP) and FRET, and enzyme-linked immunosorbent assays (ELISA) detect changes in luminescence or fluorescence upon binding of the small molecule to a tagged protein; surface plasmon resonance (SPR) measures differences in refractive index between an immobilized protein and one to which a small molecule is bound. Other methods that do not require significant manipulation of one of the binding partners include thermal shift assays, isothermal calorimetry, and nuclear magnetic resonance spectroscopy (NMR). Data from these methods can provide dissociation constants (K_d) to measure binding affinity. In some cases, structure determination of the small molecule–macromolecule complex to identify the binding site and specific biomolecular interactions can be accomplished by NMR, X-ray crystallography, and cryo-electron microscopy. As with receptors, binding does not always translate to biological effects, and additional assays to evaluate downstream consequences are also necessary.

Until recently, high-throughput screening was only modestly successful in identifying hits for this class. The reasons for this vary, but it has been speculated that compound screening libraries are populated with molecules that are more appropriate for traditional enzyme and receptor targets and therefore contain few structures that might inhibit PPIs. In order to address this issue, an active effort to design and synthesize small-molecule libraries that exhibit the known characteristics of PPI modulators (three-dimensionality, for example) for inclusion in screening decks has been a priority. Additionally, high-throughput screening of natural products or natural product–inspired analogs has experienced renewed interest. One PPI hit identification strategy that has seen particular success is fragment-based screening (see Chapter 2). Direct binding of a small low-affinity fragment to one of the protein partners is detected by methods such as SPR, NMR, MS, or X-ray crystallography. Further optimization can generate small molecules that modulate protein–protein interactions with very high affinity. When feasible, structure-based design has also been a successful approach. As a reminder, PPIs and their inhibitors are not a single category in the sense that enzymes, GPCRs, ion channels, and nuclear receptors are; therefore, generalizing strategies for hit identification is inappropriate. The following sections highlight some specific examples of small molecules that either inhibit or mimic a PPI.

8.3.1 One Strategy to Prevent Binding between Two Proteins Is to Mimic One of the Protein Partners

When binding between two proteins contributes to a disease state, inhibitors of the PPI are sought. A small molecule that binds to either protein and prevents its interaction with its partner is one common approach. Examples of this approach have been developed by the use of several different strategies, including generating peptide mimetics, identifying hits from high-throughput screening, and using structure-based design that relies on knowledge of the co-structure and interactions and conformations that occur at the binding interface.

The α-helix is one common protein secondary structure motif, accounting for over 30% of protein secondary structures, and it is frequently observed at PPI binding interfaces. The precise nature of the interactions varies widely. In the simplest cases, only one face of a relatively short span of the α-helix is involved in recognition; in others, the interface involves multiple faces of the helix over longer stretches, and still others rely on multiple helices interacting with the single binding partner. An α-helix is a right-handed coiled conformation where every peptide backbone N-H makes a hydrogen bond with the carbonyl in the backbone four residues away from it. This arrangement results in every third or fourth amino acid side chain being stacked above one

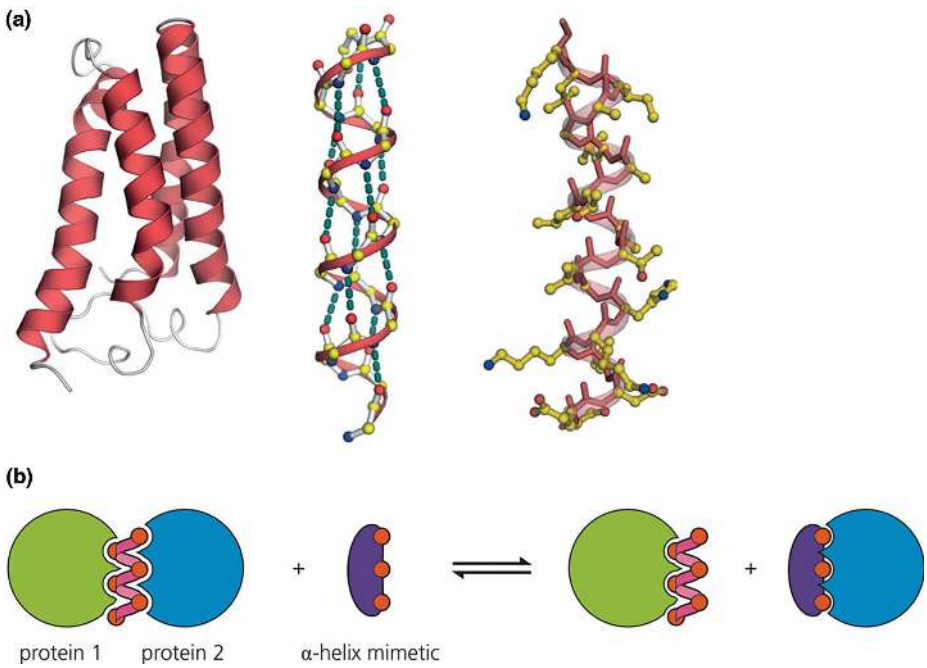


Figure 8.3 Representation of α-helix secondary structure motif. (a) *Left*, ribbon diagram of several α-helices of a protein in close proximity. *Center*, ball-and-stick model highlighting H-bonds (shown as dotted lines) between backbone N-H and carbonyl groups (side chains not shown). *Right*, helix showing side chains (yellow) aligned and extending from helix. (b) Model of binding between two proteins mediated by an α-helix interface, and schematic of a small-molecule mimic of an α-helix, mimicking the binding of protein 1 to protein 2.

Source: (b) Adapted from Azzarito, V.; Long, K.; Murphy, N. S.; and Wilson, A. J. (2013). *Nat Chem* 5:161–173. With permission from Macmillan Publishers Ltd.

another so that the α-helix can be divided into distinct faces on each turn of the helix (Figure 8.3a). Since the overall shape of the scaffold is the same and differs only in the identity of the side chains displayed, general strategies to mimic α-helices that can compete with the native α-helix partner for binding may be broadly applicable for inhibitor or even agonist design for interactions mediated by α-helices (Figure 8.3b).

The earliest and simplest efforts to design mimetics of α-helices relied on modifying an appropriate peptide fragment so that it adopts and maintains the active α-helical secondary structure. To do so, the general concept of conformational constraint (Chapter 4.5.5) can be applied by cyclizing the peptide termini or by cyclizing side chains, forming disulfides and lactams. Not unexpectedly, these compounds suffered from the poor druglike properties common to peptides, such as metabolic instability and poor absorption. More recently, stapled peptides, which are short fragments of α-helical peptides constrained by a hydrocarbon bridge between the ends of the helix (see Chapter 4), have seen some success. In many cases, a ring-closing metathesis reaction is applied to cyclize the peptide (Figure 8.4). Incorporation of non-native amino acids, particularly those containing an α,α-disubstitution pattern, and further optimization can provide enhanced helicity and protease stability as well as improved cellular uptake compared to native peptides.

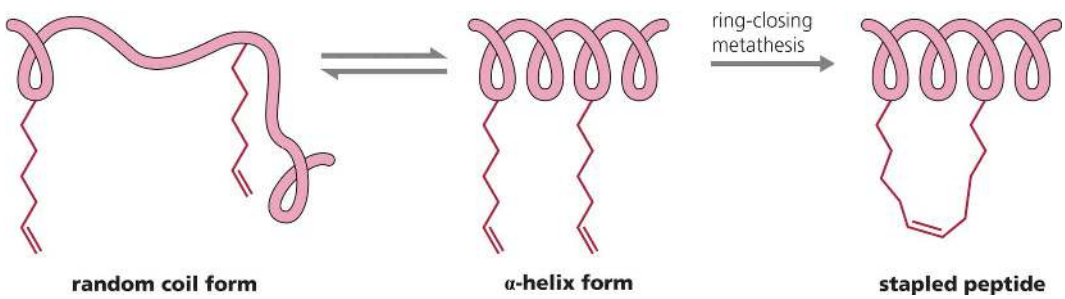


Figure 8.4 Representation of stapled peptide formation. A properly designed linear peptide containing olefinic side chains can adopt an α-helical conformation, which brings the olefinic side chains in close enough proximity that a ring-closing metathesis reaction will generate the conformationally constrained α-helix stapled peptide.

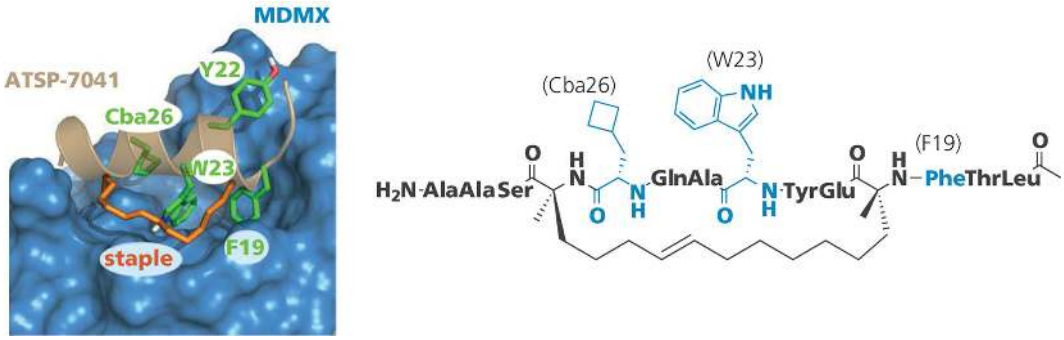


Figure 8.5 Structure of stapled peptide ATSP-7041. *Left*, X-ray structure of ATSP-7041 bound to MDMX. Helix is shown in brown, the staple is depicted in orange, and side chain residues (F, W, and Cba) making key contacts are shown in green. Tyrosine 22 also makes significant contacts to the protein. *Right*, chemical structure of ATSP-7041, with key residues F19, W23, and Cba26 (a cyclobutane amino acid) shown in blue.

Source: Left, from Chang, Y. S.; Graves, B.; Guerlavais, V.; et al. (2013). *Proc Natl Acad Sci USA* 110:E3,445–E3,454. With permission from the National Academy of Sciences.

In one example, the binding between an α-helix of the tumor suppressor p53 and its negative regulators, murine double minute-2 and X (MDM2 and MDMX), was targeted. This tumor suppressor has been called "the guardian of the human genome" and controls cellular response to DNA damage. In certain cancers, MDM2 and MDMX (the human versions) are overexpressed and suppress p53's activity, thereby allowing uncontrolled tumor growth; compounds that inhibit this interaction should restore the repressive activity of p53 and slow cancer growth. Stapled peptides were designed which mimicked p53's binding interface with MDM2/MDMX; they bound to them with high affinity and prevented MDM2/MDMX from binding to p53, thereby restoring and enhancing p53 tumor suppressor activity. The stapled peptide ATSP-7041 (**Figure 8.5**) was the result of an effort based first on optimizing linear and cyclic peptides that contained key amino acid residues (Phe, Trp, and Leu) found on the p53 binding surface, and then ultimately "stapling" them. ATSP-7041 retains the Phe and Trp residues (shown in blue) and replaces the leucine residue with a cyclobutane amino acid (Cba, also shown in blue).

ATSP-7041 exhibits high binding affinity to both MDM2 and MDMX (K_i = 0.9 and 7 nM, respectively). The stapled peptide was found to be highly helical, as designed. It exhibited favorable PK properties, was efficacious in *in vivo* models, and served as the basis of additional compounds, such as ALRN-6924 (sulanemadlin), which progressed to clinical trials as an anti-tumor drug. X-ray crystallography of ATSP-7041 bound to MDMX (see **Figure 8.5**) indicated that the essential Phe, Trp, and Cba residues were indeed bound as expected; however, additional binding interactions from a tyrosine residue side chain and from the hydrocarbon staple were also evident and probably contributed to the high binding affinity.

The stapled peptide strategy has numerous potential applications: For example, stapled peptides can be designed that inhibit a PPI or act as an agonist by binding to a protein partner and propagating downstream events. Reports of effective stapled peptides span therapeutic applications in cancer, infectious diseases, metabolic diseases, and neurology. In addition, new chemistries beyond ring-closing metathesis, such as click chemistry and C-H activation, among others, are also an active area of research. Multiple stapled peptides are undergoing clinical trials, though none have yet been approved by regulatory agencies.

More recently, it has been appreciated that while PPI interfaces may be large, a small subset of residues, or **hot spots**, often contributes disproportionately to binding energies. This is indeed the case for binding of p53 to its negative regulators: The sites on MDM2 and MDMX occupied by the Phe19, Trp23, and Leu26 residues of p53 are considered hot spots. When hot spots are part of the binding interface, there is an increased chance of success in identifying small-molecule modulators of PPIs, since the competing ligand does not need to span such a large area. In fact, a high-throughput

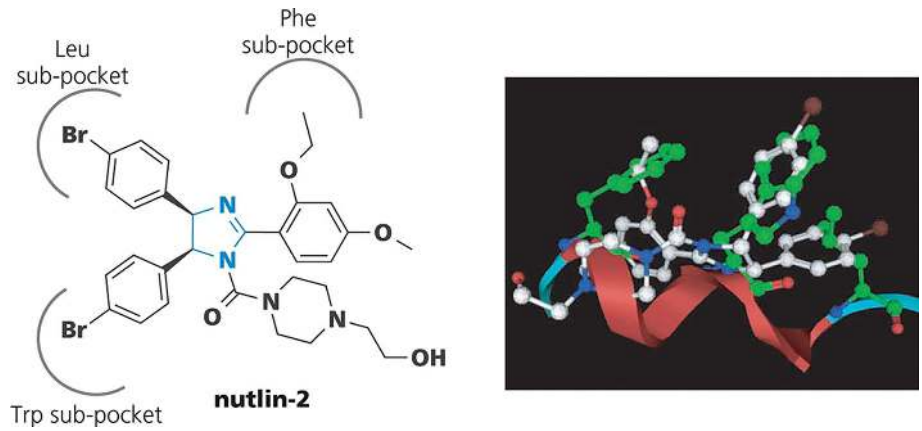


Figure 8.6 Structure of nutlin-2. *Left*, chemical structure of nutlin-2, with cis-imidazoline scaffold in blue. The substituents occupy the same sub-pockets as the hot spot residues in p53: One bromophenyl occupies the Trp subsite, another the Leu subsite, and the ethyl ether side chain occupies the Phe subsite. *Right*, overlay of p53 peptide backbone (red) containing Phe, Trp, Leu triad (green), and nutlin-2 (carbon = white; nitrogen = blue; oxygen = red; bromine = brown). The ethyl ether and Phe side chain occupy the same space; the bromophenyl groups overlay with Trp and Leu side chains, respectively.

Source: From Vassilev, L. T.; Vu, B. T.; Graves, B.; et al. (2004) *Science* 303:844–846. With permission from AAAS.

screen to identify inhibitors of the interaction between p53 and MDM2 led to one of the early examples of potent small-molecule PPI inhibitors.

The *nutlins* (named for Nutley inhibitor), reported by workers at Hoffmann-LaRoche, are a series of small-molecule inhibitors of MDM2's binding to p53. These *cis*-imidazolines were the result of an optimization effort that started from a hit identified from a high-throughput screen, and they are considered one of the first examples of highly potent PPI inhibitors developed from traditional screening. X-ray crystallography showed that, when bound to MDM2, the *cis*-imidazoline scaffold (blue) mimics the α-helical backbone of the p53 peptide by orienting the substituents so they occupy the binding sites normally filled by the key Phe, Trp, and Leu triad. In the overlay of nutlin-2 (multicolor) and the p53 peptide (side chains = green) containing the hot spot residues shown in **Figure 8.6**, one bromophenyl of nutlin-2 mimics the Leu of p53, another overlaps with the Trp side chain, and the ethyl ether occupies the same site as Phe.

Since the report of the nutlins, a number of additional small-molecule inhibitors of MDM2 that work by mimicking the p53 helix have been reported. **Figure 8.7** illustrates a few representative scaffolds of MDM2 inhibitors that have advanced to clinical trials. Research on the nutlins led to idasanutlin, which showed promising results against solid tumors and leukemias. AMG 232 (navtemadlin), which advanced

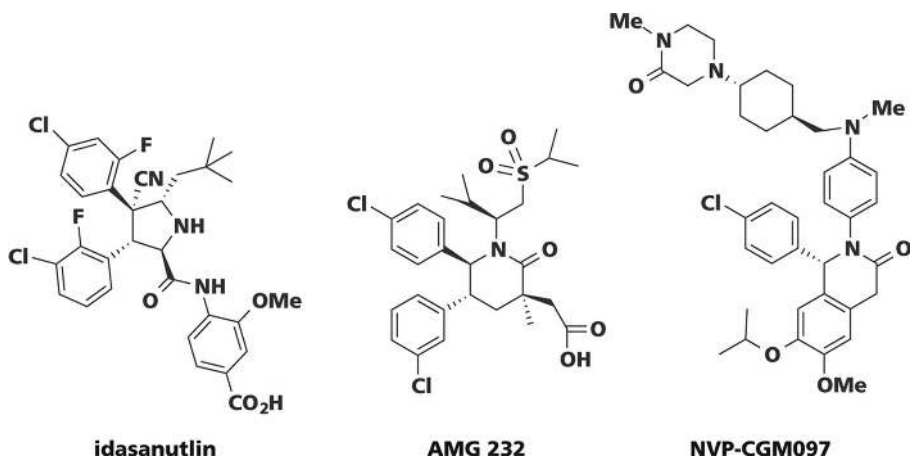


Figure 8.7 Structures of small-molecule inhibitors of p53 binding to MDM2 that have advanced to clinical trials. Idasanutlin was developed from the nutlin leads, while AMG 232 was based on a high-throughput screening hit followed by rational design. NVP-CGM097 was developed from a virtual screening program. Despite containing different scaffolds, all orient their substituents for optimal binding to the hot spot regions of MDM2; all three compounds advanced to clinical trials.

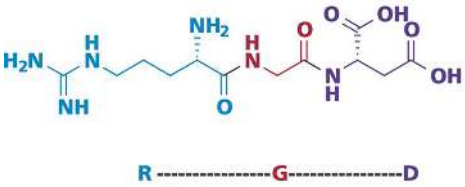
to Phase III clinical trials, is a compound developed at Amgen using a combination of high-throughput screening to identify a lead and rational design based on other known inhibitors for optimization. A third example, NVP-CGM097, was developed at Novartis from a hit identified using a virtual screening approach. This compound was evaluated for its effects against solid tumors. Despite their diverse scaffolds, all these structures correctly orient the side chains for optimal binding to MDM2. These small-molecule inhibitors of MDM2, many optimized by rational drug design approaches, illustrate that large peptides can indeed be mimicked by small druglike molecules.

Another example of a PPI target that is mediated by only a few amino acids can be seen in the interaction between the proteins fibrinogen and glycoprotein IIb/IIIa (GPIIb/IIIa). This binding event is essential to platelet aggregation, and its inhibition is a viable therapeutic strategy to treat myocardial infarction and stroke. GPIIb/IIIa recognizes the short amino acid sequence arginine–glycine–aspartic acid (Arg-Gly-Asp, or RGD) on fibrinogen. By mimicking this sequence of amino acids, compounds that inhibit the binding of fibrinogen to GPIIb/IIIa were developed and have found use as anti-platelet agents.

Eptifibatide is a drug that was developed on the basis of inhibiting this PPI. A series of peptides isolated from snake venom, called disintegrins, were known to inhibit platelet aggregation, and most contained the RGD sequence. Screening of 62 snake venoms yielded a 73–amino acid peptide called barbourin that selectively inhibited the binding of GPIIb/IIIa to fibrinogen. Instead of RGD, barbourin contained a KGD sequence, with the basic lysine (K) replacing the arginine (R). Optimization by shortening the peptide and constraining its conformation via cyclization led to eptifibatide (**Figure 8.8**), a heptapeptide containing a disulfide bridge. It is a potent inhibitor of the binding of GPIIb/IIIa to fibrinogen and is approved as an injection to prevent blood clots from forming after certain types of chest pain and heart attack, as well as for patients undergoing coronary angioplasty.

A significant effort to convert the RGD peptide sequence to peptide mimetics with improved druglike characteristics was undertaken by a number of research groups. The isosteric replacement of the polar guanidine and carboxylic acid groups was a common strategy, as was replacement of the labile peptide bonds. The considerable SAR available indicated that the optimal distance between the acidic and basic functions that correspond to Arg and Asp, respectively, was between 10 and 15 Å, and this information was used to design small-molecule inhibitors, as well as to screen for compounds that adhered to this pharmacophore. Of the many small-molecule RGD mimetics that entered clinical trials, only tirofiban (**Figure 8.9**) was approved. The piperidine nitrogen of tirofiban mimics the guanidino group of the RGD sequence, and the tyrosine carboxylic acid replaces the aspartic acid residue. It is proposed that the sulfonamide makes additional interactions with the GPIIb/IIIa protein, not observed in the original disintegrins.

Figure 8.9 also shows a few of the RGD mimetics that advanced to clinical trials but ultimately failed due to a lack of efficacy as well as slight increases in mortality as compared to placebo. Along with isosteric replacements, design strategies involved



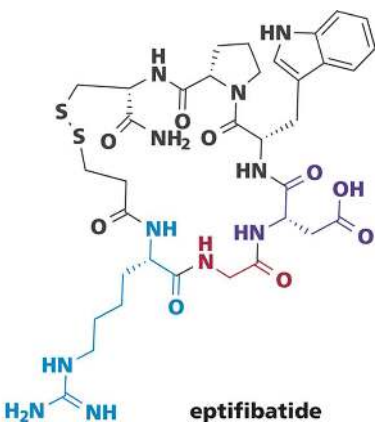


Figure 8.8 Structure of RGD and eptifibatide. *Left*, RGD, the Arg-Gly-Asp tripeptide present in fibrinogen, is recognized by GPIIb/IIIa. *Right*, eptifibatide is a constrained macrocyclic peptide that contains the RGD sequence. Blue, red, and purple correspond to R, G, and D residues, respectively.

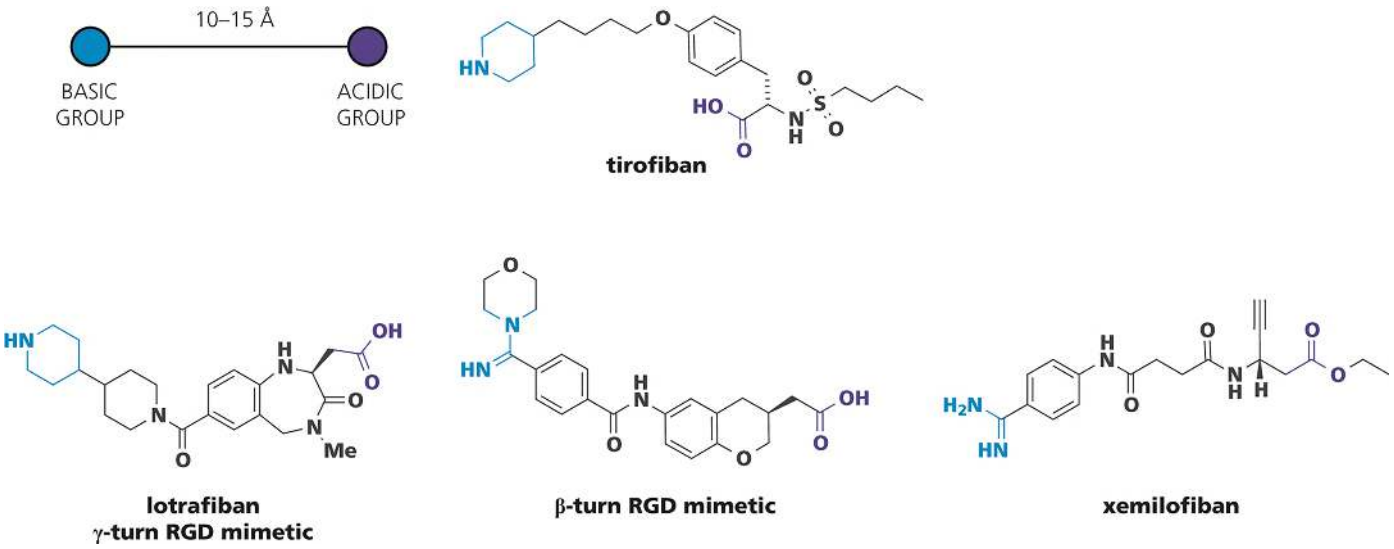


Figure 8.9 Pharmacophore model for GPIIb/IIIa antagonists. Blue denotes basic functionality, and purple denotes acidic functionality. Tirofiban is the only approved peptide mimetic drug in this class. The piperidine (blue) and acid (purple) correspond to the Arg and Asp side chains in fibrinogen. Other representative structures of inhibitors that are RGD mimetics include lotrafiban, a beta-turn mimic, and the prodrug ethyl ester xemilofiban. In each of these, blue indicates the basic group, and purple the acidic group, of the pharmacophore.

mimicking β- or γ-turn secondary structures observed in potent acyclic and cyclic RGD peptides, respectively. Examples include the benzodiazepine analog lotrafiban and the β-turn mimic shown. Efforts to use a prodrug strategy in order to generate orally bioavailable molecules have also been described, with the ethyl ester xemilofiban as one example that progressed to clinical trials but did not receive approval.

The development of compounds that bind to BCL-X_L and BCL-2 represents a third example of small molecules blocking protein–protein interactions at their interface. These two proteins are anti-apoptotic and important for the survival of certain cancer cells. It was rationalized that small molecules that interacted with BCL-X_L and its homologs and prevented their binding to their natural partners would inhibit their pro-survival effects and be potential anti-cancer agents. In a fragment-based screen, researchers at Abbott used NMR to identify low-molecular-weight compounds that bound to BCL-X_L at two distinct sites, later identified as the P2 and P4 sites, albeit with very low affinity. Linking of these fragments, a biphenyl acid and a biphenyl alcohol, together and further optimization eventually led to navitoclax, an inhibitor of both BCL-X_L and BCL-2 binding to their protein partners. The thiophenyl fragment occupies the same site the original biphenyl alcohol hit did (P4), and the chloro-cyclohexenylphenyl occupies the same site as the biphenyl acid hit (P2). Navitoclax advanced to clinical trials for small-cell lung cancers (**Figure 8.10a**) but was dropped due to an unforeseen side effect of thrombocytopenia, linked to its inhibition of BCL-X_L. The similar drug venetoclax, selective for BCL-2, was then advanced and received FDA approval for leukemia in 2016 as the first BCL-2 inhibitor (**Figure 8.10c**). The success of this project has been partially attributed to the fact that the protein interface between BCL-X_L and its partners is smaller than other protein–protein interfaces (~600 Å^2) and consists of a deep hydrophobic groove of about 20 Å where small molecules can bind. This groove has two sites, P2 and P4, that bind its protein partners. An X-ray structure of navitoclax shows that two portions of the molecule bind in those P2 and P4 hydrophobic sites (**Figure 8.10b**).

These examples highlight two general strategies for identifying inhibitors of PPIs. First, by understanding the local conformation and structure of the binding interface, scaffolds can be designed that correctly orient side chains for efficient binding with the protein partner. This approach is applicable not only to α-helices but also to other types of protein secondary structures, such as β-turns. Second, the concept of hot spots explains that, despite a large binding surface, there are certain segments of some PPI interfaces that contribute to binding more so than others. In these circumstances, identifying small-molecule inhibitors of the PPI is much more likely. These interfaces are characterized by hydrophobicity as well as size, shape, and depth. Considerable efforts are underway to accurately identify hot spots by computational methods in

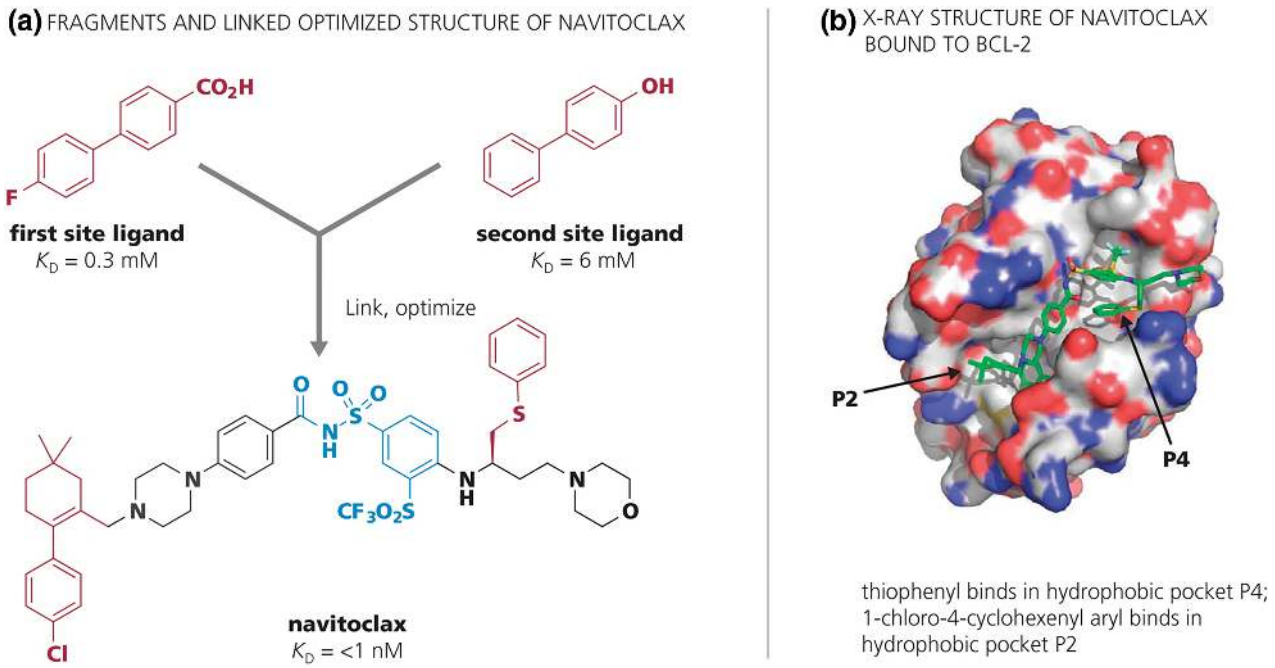


Figure 8.10 PPI inhibitors of BCL-X_L identified by fragment-based screening. (a) Two low-affinity fragments, a biphenyl acid and a biphenyl alcohol, that bind to distinct sites of BCL-X_L were identified by NMR. Linking the two and extensive optimization–generated navitoclax (portions derived from fragments depicted in red, and linker in blue) are shown. (b) X-ray crystal structure of navitoclax bound in hydrophobic cleft; thiophenyl binds in pocket P4 and cyclohexenyl aryl chloride binds in pocket P2. (c) Structure of FDA-approved BCL-2 inhibitor venetoclax.

Source: (b) Adapted from Souers, A. J.; Leverson, J. D.; Boghaert, E. R.; et al. (2013). *Nat Med* 19:202–210. With permission from Macmillan Publishers Ltd.

order to expedite PPI drug design. Given the diverse nature of PPIs and the interactions that contribute to disease states, a variety of different strategies to modulate these interactions are possible. Some of these strategies have led to approved drugs, while some are still in the earliest stages of discovery.

8.3.2 Protein–Protein Interactions Can Be Modulated by Chemically Induced Proximity

Chemically induced proximity (CIP) drugs work through mechanisms that are fundamentally different from the more conventional small-molecule drugs in that they induce proteins to form ternary (or even higher) complexes, and then these complexes exert their pharmacological effects. These molecules can work via several mechanisms; the two most common are broadly classified as **bifunctional degraders** and **molecular glues**. Bifunctional degraders have three components: a binding element to one protein, a second component with a binding element to another protein, and a linker connecting the two. Molecular glues bind to one protein, forming a "neoprotein" that may then bind another protein or change the stability of the original protein. These modalities are distinct from conventional drug molecules, where a 1:1 complex between drug and target is more typical.

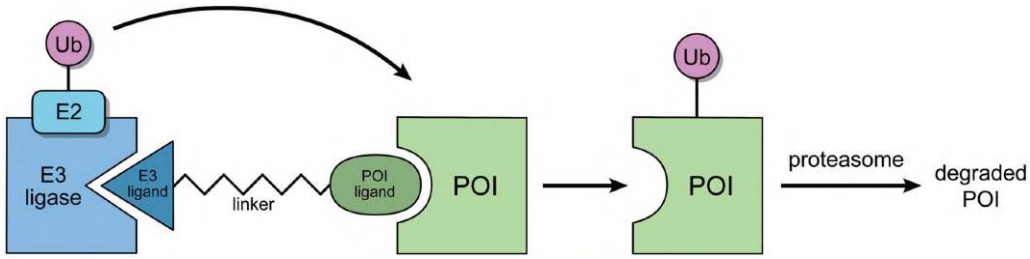


Figure 8.11 The PROTAC, made up of an E3 ligase ligand and a POI ligand connected through a linker, binds both to the E3 ligase and the POI. The E3 ligand complex, containing E2 ligase and ubiquitin (Ub), transfers ubiquitin to the POI, marking it for degradation by the proteasome. The specific binding component on the E3 ligase may be cereblon (CRBN), von Hippel Landau (VHL), inhibitor of apoptosis (IAP), or MDM2.

8.3.2.1 *Bifunctional Degraders*

Heterobifunctional molecules are designed to bring two specific proteins in proximity by linking small molecules that are known binders to each. In these cases, one of the protein components is part of an E3 ligase complex. E3 ligases are multicomponent enzymes that, together with other proteins, such as E2 ligases, tag target proteins with ubiquitin, which ultimately results in the target protein's degradation. The earliest small molecular example of these are termed **PROTACS (proteolysis-targeting chimeras)** proposed by Craig Crews and Raymond Deshaies in the early 2000s. These compounds promote the formation of a ternary complex (TC) bringing an E3 ligase in proximity to a protein of interest (POI), resulting in the POI being targeted for degradation by the cell's ubiquitin proteasome system (UPS). Most PROTACS contain a ligand for a component of an E3 ligase complex, often cereblon (CRBN), but others include von Hippel Landau protein (VHL), inhibitor of apoptosis (IAP), or MDM2. The E3 ligase ligand is connected through a linker to a ligand of the protein of interest, as shown in Figure 8.11. The result is that the POI is effectively inhibited because it is degraded in a catalytic manner. The consequence of this mechanism can be profound, since PROTACS are essentially catalytic in their effects, and once a protein molecule is degraded, the PROTAC can continue to act on remaining stores of the POI.

To effectively design a PROTAC, extensive knowledge of the pharmacophore for the POI and ubiquitin ligase moiety (ULM) is required in order to understand what features of the molecule must be presented for binding and where linkers can be incorporated. In the case of the most frequently targeted E3 ligase components CRBN, VHL, IAP, and MDM2, the pharmacophores are well understood. Structures for E3 ligase ligands with validated points for linker attachment are shown in Figure 8.12.

The E3 ligase ligands for the specific components of the complex, shown in Figure 8.12, are based on known ligands. Early PROTACS were based on the VHL peptide ligand HIF-1α and require a hydroxyproline core. The next generation of PROTACS used MDM2 ligands that are based on idasanutlin. Currently, the most widely used are the CRBN ligands, based on the "IMiDs," such as lenalinomide, thalidomide, or pomalidomide. PROTACS using IAP ligands mostly commonly are based on LCL-161, and these PROTACS are also termed "SNIPERS" (specific and nongenetic IAP-dependent protein erasers).

As this field has developed, it has become clear that, in addition to the POI ligand and ULM, the identity of the linker also contributes to potency and efficacy of degradation and can also be used as a site to optimize physical and pharmaceutical properties, such as solubility and permeability. Given the need to generate a productive ternary complex that can transfer ubiquitin to the POI, the length and orientation of the linker are essential. While early PROTACS contained very flexible linkers, more recently, conformationally constrained linkers are more prevalent. The inclusion of heteroatoms (and heterocyles) into linkers also contributes to the optimization of physical and pharmaceutical properties. In some cases, linkers also contribute to binding to one or both targets. The considerable effort devoted to linker design is termed "linkerology." Linkerology can rely on structure-based design when the co-structures of the ligand-bound POI is available to ensure optimal distance and orientation, but more frequently, a more empirical approach is applied, typically through the synthesis of arrays of PROTACS in which the linker and points of attachments are varied.

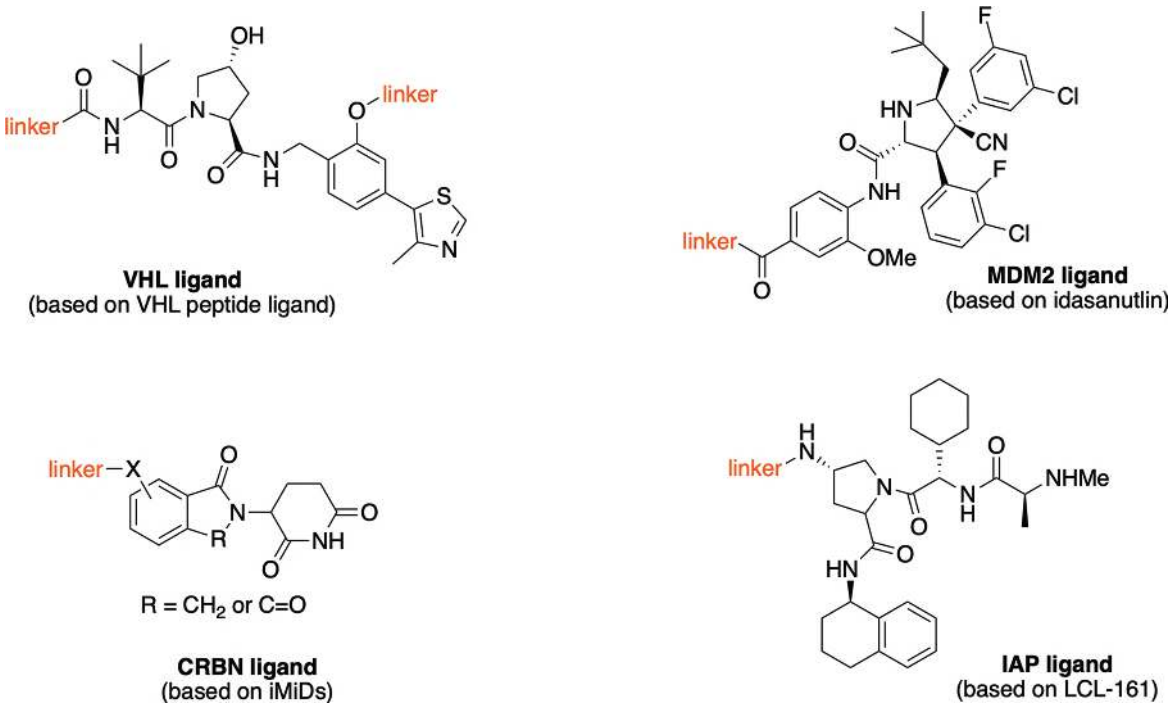


Figure 8.12 Structures of common E3 ligase ligands, with the most common points of linker attachment in red.

While PROTACS provide an innovative approach to inhibit certain proteins, this modality presents some challenges related to development, particularly oral bioavailability, that are unique to these large molecules. With molecular weights in the 700–1,500 Da range and high lipophilicity, PROTACS fall well beyond the rule of five parameters (called beyond Ro5, or bRo5). One consequence is that, currently, only PROTACS that rely on the relatively compact lenalidomide-derived E3 ligase ligands are considered compatible with oral bioavailability. Furthermore, it has been observed that traditional parameters used to predict oral bioavailability are not effective when applied to PROTACS, perhaps due to three-dimensional folding effects that affect polarity and permeability. PROTACS and other bRo5 compounds may be "chameleonic" and require novel parameters for the prediction of druglike properties. In addition to tPSA (topological polar surface area), experimental measurements of exposure polarity (EPSA), and the ratio of the two (ETR: EPSA to tPSA ratio), as well as multi-parameter scoring metrics, such as those developed by AbbVie, that incorporate clogD, number of rotatable bonds, and number of aromatics rings are being used to predict *in vivo* PK properties of not only PROTACS but also other bRo5 compounds.

Evaluating the biological activity of PROTACS requires a series of assays that measure binding of the molecules to both the E3 ligase and target protein, as well as its functional effects. Binding assays employed are identical to those used to detect binding for more traditional modalities, such as SPR or displacement assays detected by fluorescence polarization. Secondary assays can detect ternary complex formation between the two proteins and a PROTAC using traditional biophysical measurements, such as SPR or ITC (isothermal titration calorimetry). To monitor target protein degradation in cells, Western blotting, which detects reduction/loss of protein, is common; alternatively, assays that use overexpressing and tagged version of target proteins can be used to evaluate changes in protein level. An important value generated is a DC_{50}, which reflects the concentration of compound that results in the degradation of 50% of the POI.

The first PROTACs entered clinical trials in 2019 using well-established cancer targets, the estrogen and androgen receptors, as the POIs. **Figure 8.13** illustrates the structures of two specific PROTACS: ARV-471 (vepdegestrant) targets the estrogen receptor and ARV-110 (bavdegalutamide) targets the androgen receptor. For both compounds, the E3 ligase moiety in blue targets CRBN, the target receptor ligand is in green, and the linker is the red piperazinyl piperidine. As of 2025, vepdegestrant was in Phase III trials and received FDA Fast Track designation for patients with ER-positive/HER2-negative metastatic breast cancer.

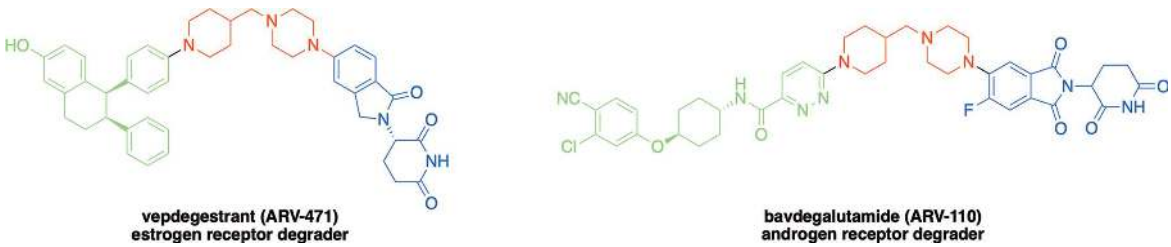
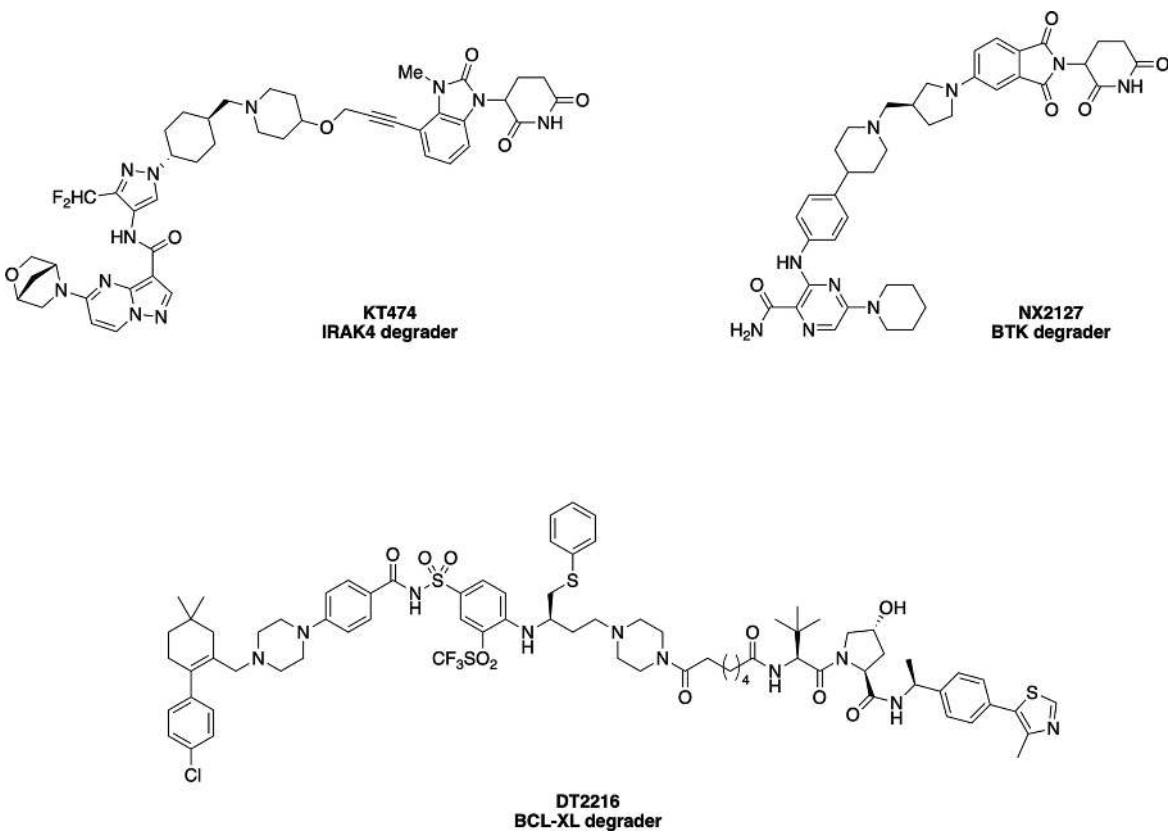


Figure 8.13 Examples of early PROTACS that advanced into clinical trials. E3 ligase binding region in blue, linker in red, target protein binding region in green.

While the most advanced PROTACS to date target POIs involved in cancer, PROTAC strategies are being widely applied to other targets, with a number in early clinical trials. These POIs include druggable targets, such as kinases, as well as "undruggable" targets, such as the misfolded proteins tau and synuclein. As described earlier and exemplified by ARV-471 and ARV-110, most PROTACS target CRBN as the E3 ligase, although a few target VHL. Some examples of other PROTACS in early clinical trials as of 2025 are shown in **Figure 8.14**. Both KT474 and NX2127 bind to CRBN as the E3 ligase and target kinases. KT474 targets and degrades IRAK4 (interleukin-1 receptor–associated kinase), which is involved in a signaling cascade that results in the release of proinflammatory cytokines and chemokines. It advanced to Phase II clinical trials for the treatment of immune-inflammatory diseases, such as atopic dermatitis, but was discontinued. NX2127 targets BTK (Bruton's tyrosine kinase) and is in Phase I trials for B-cell malignancies. One example of a PROTAC that binds to VHL as the E3 ligase is DT2216, which degrades BCL-X_L. The VHL-binding region is linked to venetoclax, discussed previously (**Figure 8.10**). This PROTAC is in clinical trials for the treatment of B-cell malignancies and received FDA Fast Track designation.

Figure 8.14 Examples of PROTACS that advanced to clinical trials. Both NX2127 and KT474 target kinases for degradation and bind CRBN as the E3 ligase. DT2216 targets BCL-X_L for degradation and binds VHL as the E3 ligase.

The advancement of PROTACS has led to other **targeted protein degradation** (TPD) strategies. Hydrophobic tag–based protein degradation (HyT-PD) relies on the incorporation of a hydrophobic group, such as adamantane, in place of the E3 ligase ligand, and like a PROTAC, this is linked to a ligand of a POI. The hydrophobic tag is thought to resemble misfolded or damaged proteins and triggers the cell's protein quality control mechanisms to degrade the POI. Lysosome-targeting chimeras



(LYTACs) depend on the lysosome degradation pathway and, unlike PROTACs, can be used to degrade extracellular and membrane-bound POIs, and RiboTACs degrade RNA. These and other TPD approaches will likely see considerable attention in the future.

8.3.2.2 Molecular Glues

Molecular glues are small molecules that, by way of their binding to one protein, promote a new protein–protein interaction. By binding to one protein, a new structure is formed, called a **neoprotein**, which, in essence, changes the interactome. The neoprotein will have different properties compared to the native protein and can bind to a second protein (called a neosubstrate), and that ternary complex can result in effects such as degradation, stabilization, destabilization, localization, or modulation of activity. The natural products cyclosporin, FK506 (tacrolimus), and rapamycin all work by a molecular glue mechanism and act as immunosuppressants. Related approved drugs temsirolimus and everolimus (Figure 8.15) are semisynthetic ester and ether analogs of rapamycin, with the ester or ether modification (shown in red) imparting improved physical properties, particularly water solubility. FK506 and rapamycin and its analogs bind to a protein called FKBP12, and that macromolecular complex (a neoprotein) inhibits one of the mTOR proteins. In the example of rapamycin, the structural determinants for binding to FKBP12 include the pipecolic acid and dicarbonyl moieties (shown in green). A distinct area of the molecule, the conjugated triene (shown in blue for temsirolimus), binds to mTOR and prevents it from phosphorylating downstream proteins and propagating the signaling cascade. Cyclosporin differs in that it binds to cyclophilin, and that neoprotein inhibits the protein calcineurin, resulting in a reduced function of T cells. It should be noted that calcineurin and cyclophilin do not typically interact without the cyclosporin glue. All these molecules are approved to prevent organ rejection, and their discovery paved the way for organ transplantation being a viable therapeutic strategy for many patients. This molecular glue mechanism was discovered well after the therapeutic benefit of the molecules was proven.

When molecular glues result in degradation, they are termed **molecular glue degraders**. Molecular glue degraders are distinctly different from PROTACS in that a bivalent ligand is not required, and for that reason, many of these were not designed, but their mechanism only delineated later. Thalidomide is one of the first examples of a drug that was found to work through a molecular glue degrader mechanism, and that understanding came well after its clinical use for leprosy and multiple myeloma was discovered. A group at Tokyo Institute of Technology found that thalidomide binds to CRBN, and later work found that once bound, the CRBN–thalidomide neoprotein recognizes various neosubstrates, including transcription factors called Ikaros (IKZF1) and Aiolos (IKZF3), resulting in their degradation, and providing anti-myeloma effects.

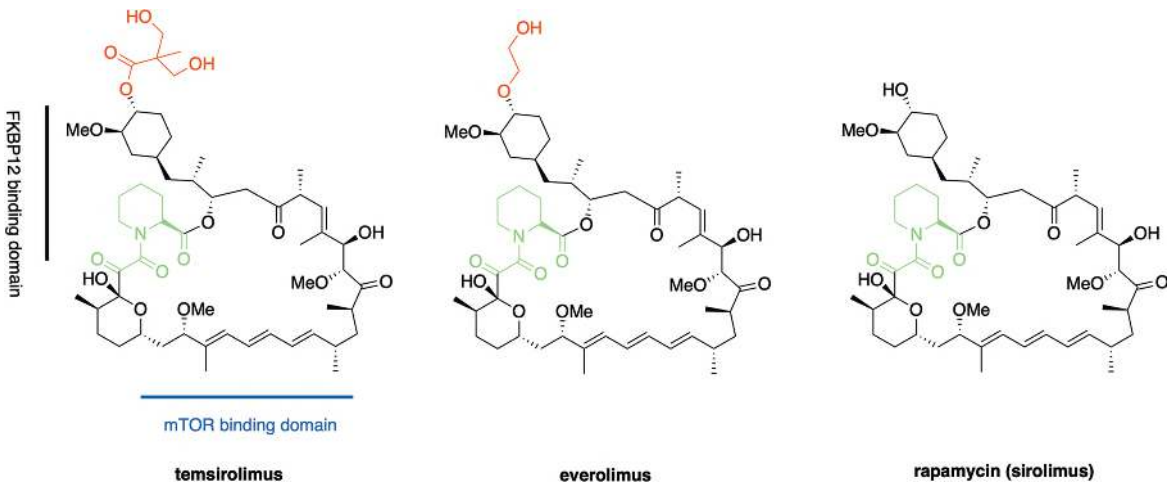


Figure 8.15 Structures of temsirolimus and everolimus, which are semisynthetic analogs, and natural product rapamycin. Ester or ether linkages, which impart improved physical properties, are shown in red, and the pipecolic acid-dicarbonyl region, which is essential for binding to FKBP12, is shown in green. A gray line highlights the FKBP12 binding domain, and a blue line shows the mTOR binding domain in the temsirolimus structure.

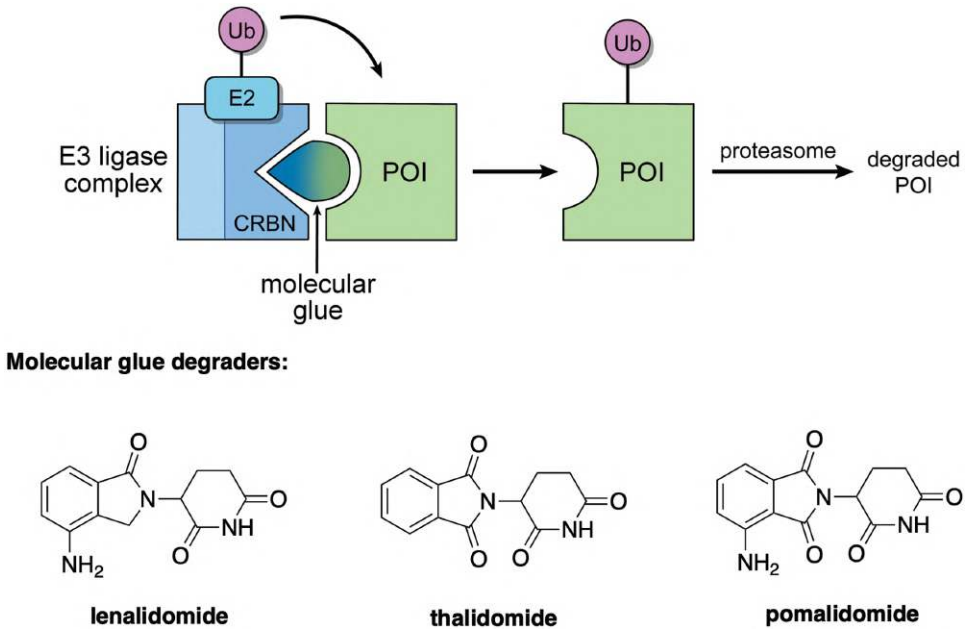


Figure 8.16 General description of molecular glue degraders lenalidomide, thalidomide, and pomalidomide showing binding to both the E3 ligase complex and POI.

Other analogs of thalidomide, such as pomalidomide and lenalidomide, work through a similar mechanism and are members of a family of drugs called IMiDs (immunomodulatory drugs). Figure 8.16 outlines this mechanism for lenalidomide. Like a designed PROTAC, the molecular glue degrader binds to both an E3 ligase and a POI and brings the E3 in proximity to the neosubstrate POI and promotes the ubiquitination of the POI, thereby eventually initiating its degradation by the UPS.

Molecular glues are likely to see continued growth. For instance, it is becoming evident that some known drugs work through such a mechanism, either via formation of neoproteins or via activation of degradation pathways. A few examples of these are highlighted in later chapters. However, intentional design can be challenging since there are considerable gaps in our understanding of how to predict which small molecules can have glue-like properties, as well as how to direct neoproteins to neosubstrates; however, this is a very active area of research.

8.3.3 Small Molecules That Bind to a Protein Can Stabilize or Destabilize It, Leading to Therapeutic Effects

When a protein binds transiently in order to modulate a pharmacological effect, small molecules that either stabilize or destabilize it can be used to prolong the effect (stabilize), reduce the effect (destabilize), or merely change the equilibrium. The most notable small-molecule examples in this category are those that modulate tubulin polymerization. Tubulins, which exist in two forms, α and β, are spherical proteins that interact to form a heterodimer that also contains two molecules of GTP. The heterodimers interact further in a head-to-tail orientation to form a long protein fiber called a protofilament that undergoes further elongation by polymerization, eventually leading to a pipelike structure called a microtubule. Microtubules are not static structures but are under active states of polymerization and depolymerization, and this process is highly regulated by accessory proteins (Figure 8.17a). In the depolymerization process, protofilaments curve and peel off from the straight microtubule, releasing heterodimers. Two general classes of small-molecule modulators are known: those that destabilize the polymer, leading to excessive depolymerization, and those that stabilize the polymer, thereby preventing depolymerization. These small molecules are thought to mimic the accessory proteins that regulate the polymerization/depolymerization process. Multiple binding sites for small molecules on tubulin have been characterized, and structural studies illustrate how these molecules disrupt tubulin oligomers.

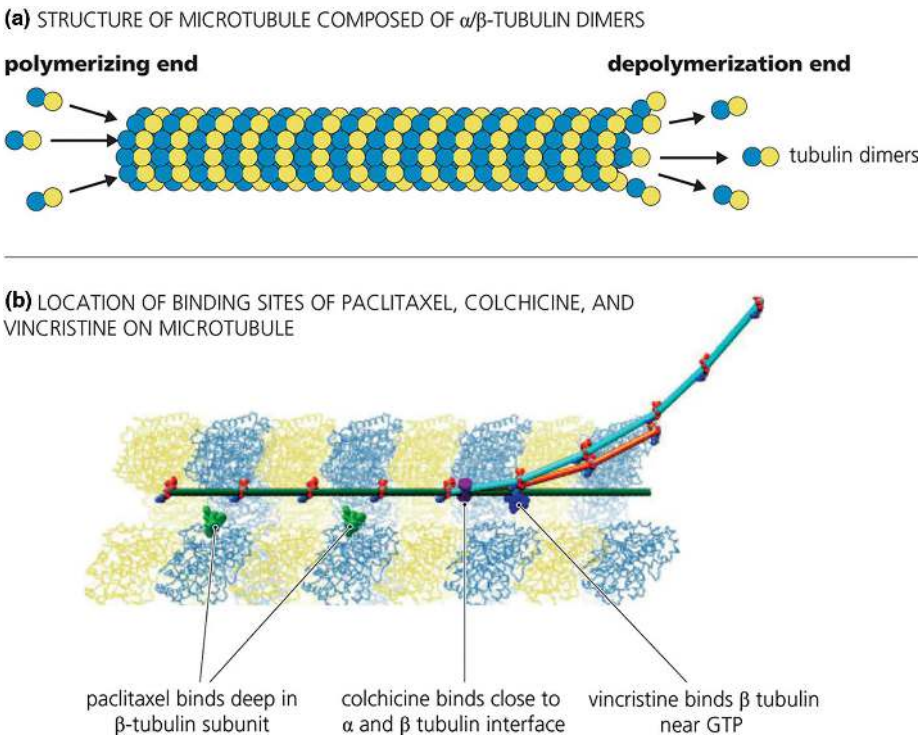


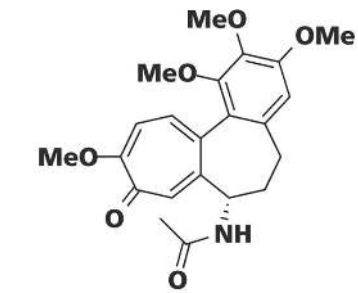
Figure 8.17 Structure of microtubule and small-molecule modulator binding sites. (a) Schematic of tubulin polymerization: α-tubulin (yellow) and β-tubulin (blue) form α/β-tubulin heterodimers. Further association of the dimers generates protofilaments (not shown) and, eventually, microtubules, which depolymerize by curving and then releasing heterodimers. (b) Binding sites on α-tubulin (yellow) and β-tubulin (blue) of small-molecule modulators paclitaxel (green; binds to a site on β-tubulin, on the inside surface of the microtubule, between adjacent protofilaments), colchicine (magenta; binds deep in the β-tubulin monomer, near its interface with the α-tubulin monomer), and vincristine (blue; binds β-tubulin near GTP). GTP molecules are shown in red/orange. The green line represents the axis of growth of protofilaments (stabilized by paclitaxel); cyan and orange lines represent axes of peeling protofilaments, stabilized by colchicine and vincristine, respectively.

Source: From Stanton, R. A.; Gernert, K. M.; Nettles, J. H.; and Aneja, R. (2011). *Med Res Rev* 31:443–481. With permission from John Wiley & Sons Inc.

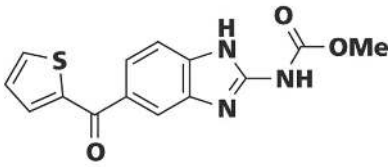
Colchicine (Figure 8.18), a natural product derived from the autumn crocus, is a drug used to treat gout. The plant's medicinal effects have been known for many centuries: the Ebers papyrus (ca. 1,500 BC) described its use to treat swelling and rheumatism. It binds to the tubulin heterodimer deep in the β-tubulin monomer, near its interface with the α-tubulin monomer (see Figure 8.17b). That binding induces kinks in the interface, leading to curved polymers and a preference for depolymerization. Other molecules, such as nocadazole, bind at or near the same site and are frequently used as tools by researchers studying tubulin and cancer biology.

The vinca alkaloids, such as vincristine and vinblastine (Figure 8.19), are natural products isolated from the Madagascar periwinkle plant that have been used as anti-cancer agents since the 1960s. Like colchicine, they destabilize tubulin polymers. However, they bind at a distinct interfacial binding site (called the vinca site; see Figure 8.17b) between the α- and β-subunits within a protofilament, and this leads to a curved structure that favors depolymerization. Compounds of diverse structure, some with high complexity, that bind at or near the vinca binding site are shown in Figure 8.19 and include monomethyl auristatin E, maytansine, and halichondrin. This diversity is in contrast to the colchicine binding site: while recognizing multiple scaffolds, that site does not bind compounds of such drastically different structures. A number of compounds that bind and inhibit polymerization or destabilize the polymer have either advanced into clinical trials or are approved drugs to treat certain cancers. Additional approved drugs in this class include a fragment of the natural product halichondrin B (eribulin; see Chapter 1) and antibody conjugates of maytansine and auristatin analogs, such as MMAE.

Taxol (paclitaxel), a natural product isolated from the Pacific yew tree, was identified as a highly potent anti-proliferative agent through a US government-sponsored natural product screening effort. Its mechanism of stabilizing microtubules rather than destabilizing them, when determined, was unique from others previously known. It binds to a site on β-tubulin (see Figure 8.17b), on the inside surface of the microtubule between adjacent protofilaments, through a large number of H-bonding and hydrophobic interactions. Upon binding, the resultant structures are stabilized and resistant to disassembly. Since the discovery of this mechanism, several other natural products that act in a similar fashion have been



colchicine



nocadazole

Figure 8.18 Structures of colchicine and nocadazole. Both compounds bind to tubulin at the same site and inhibit tubulin polymerization.

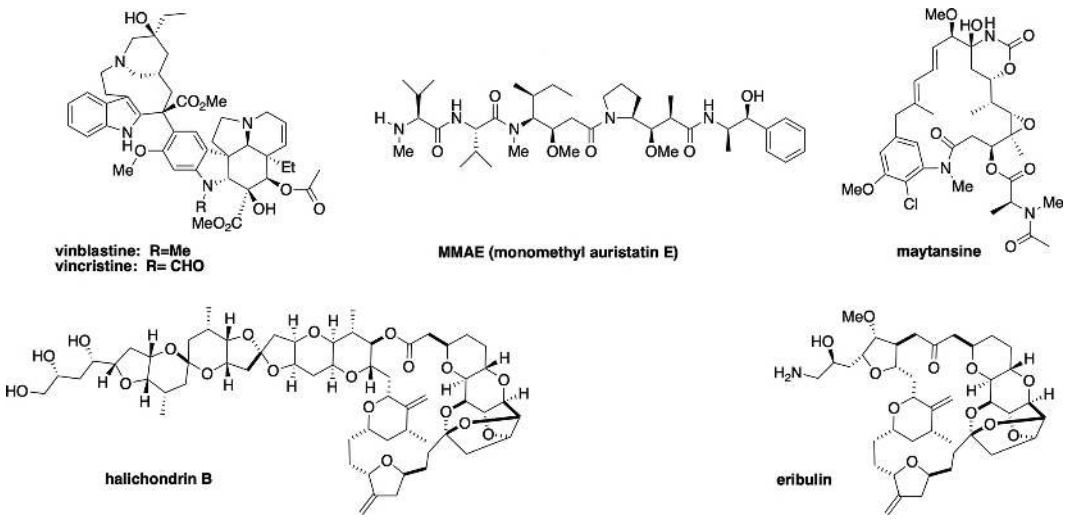


Figure 8.19 Structures of compounds that bind to the vinca binding site and destabilize microtubules. Vincristine, vinblastine, monomethyl auristatin E, maytansine, halichondrin B, and eribulin represent the vastly diverse structures that bind at or near the vinca binding site.

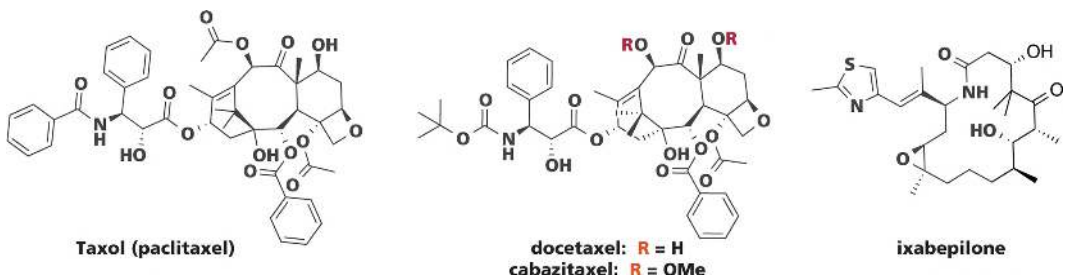
identified, such as the epothilones, a family of natural products isolated from a soil bacterium. The epothilones are competitive inhibitors of paclitaxel, suggesting that they bind at or near the same site on β-tubulin. Ixabepilone, a semisynthetic derivative of epothilone B, is approved for the treatment of certain cancers, as are Taxol (paclitaxel) and two taxoids, docetaxel and cabazitaxel, which differ primarily in the constitution of the ester side chain (**Figure 8.20**). Paclitaxel and other compounds with a similar mechanism of action have been characterized as molecular glues, in this case binding to two subunits of tubulin rather than two different proteins.

Although several high-resolution co-crystal structures of molecules bound to tubulin polymers are available, designing new or improved molecules has primarily relied on developing an understanding of structure–activity relationship studies and developing pharmacophore models. **Figure 8.21** shows one such model that superimposes the diverse structures of taxoids and epothilones, which are believed to bind to the same site.

In this example of a binding model, common structural features of paclitaxel and epothilone B include the *gem*-dimethyl and adjacent alcohol (red), cyclic ethers (oxetane in paclitaxel and epoxide in epothilone B; black), and the H-bond acceptor displayed by the carbonyl oxygen of paclitaxel and alcohol of epothilone B (red). The macrocyclic rings of the two compounds overlap, as do the thiazole side chain of epothilone and the phenyl group in the C-2 side chain of the taxane (blue) (**Figure 8.21b**).

However, this area is still controversial, as is the precise bioactive conformation of Taxol. Medicinal chemistry efforts to design improved tubulin stabilizers continues to garner significant interest. For instance, the approved anti-cancer drug ixabepilone is a lactam analog of the natural epothilone B that provides improved metabolic stability compared to the natural product, and many efforts have gone toward modifying taxoids to reduce their tendency to be effluxed. To date, however, the most effective compounds have been very close to the original natural product structure.

Figure 8.20 Structures of microtubule-stabilizing anti-cancer drugs. Taxol (paclitaxel), docetaxel, cabazitaxel, and ixabepilone are shown. Taxol was the first stabilizing compound identified, and docetaxel and cabazitaxel are close analogs. Ixabepilone is a semisynthetic analog of epothilone and is also approved to treat certain cancers.



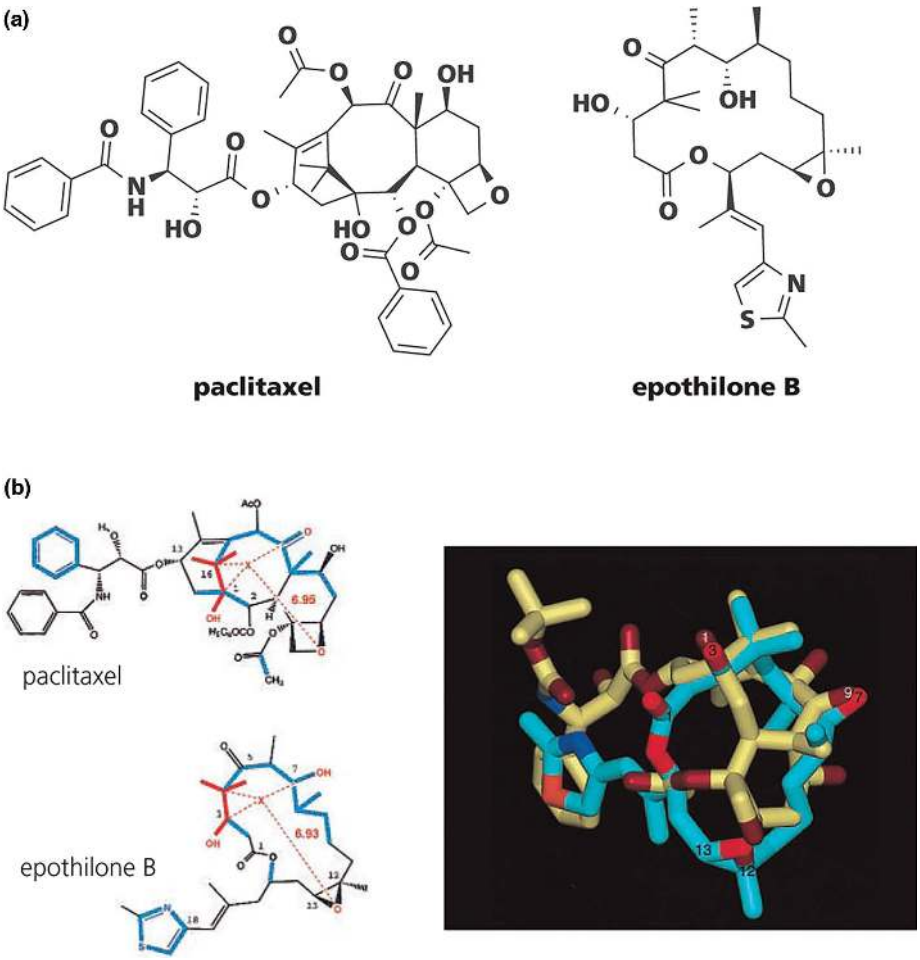


Figure 8.21 Structures of paclitaxel and epothilone B and common features. (a) Structures of paclitaxel and epothilone B. (b) *Left*, common structural features of the two molecules: gem-dimethyl with adjacent alcohol, cyclic ether, and H-bond acceptor (carbonyl and alcohol), shown in red; X defines the center of the model. Bonds/atoms shown in blue are those that overlap in the figure on the right. *Right*, molecular modeling–derived overlay (paclitaxel in yellow, epothilone B in cyan), showing similar footprint despite different scaffolds.

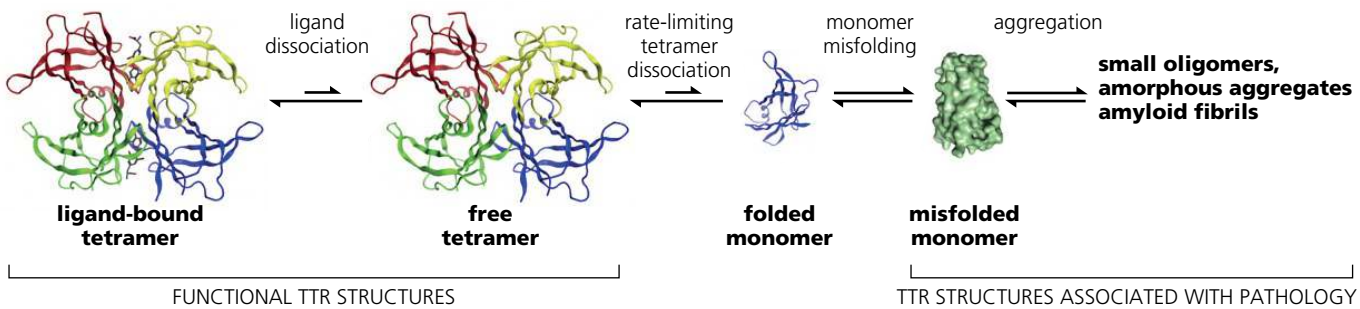
Source: Adapted from Giannakakou, P.; Gussio, R.; Nogales, E.; et al. (2000). *Proc Natl Acad Sci USA* 97:2,904–2,909. With permission from the National Academy of Sciences.

Typical drug targets are either overactive or underactive proteins. However, a third possibility is that a protein is misactive. For example, if a protein inappropriately aggregates, it may become pathogenic as the aggregate (gain of function) or incapable of performing its normal function (loss of function). The misfolding of the protein α-synuclein resulting in the formation of fibrils that are components of Lewy bodies found in patients with Parkinson's disease, and the misfolding of β-amyloid and tau proteins to form amyloid plaques and tangles, respectively, found in the brains of Alzheimer's patients is believed to contribute to the neurodegeneration characteristic of these diseases. In instances such as these, a drug discovery strategy is to use a small-molecule PPI modulator to revert the protein to a correctly folded (nonaggregated) conformation or sequester the misactive protein.

Transthyretin (TTR) is a tetrameric protein whose role is to transport certain proteins and hormones. Transthyretin amyloidosis results from the TTR monomer misfolding and aggregating. Initially considered a rare disease, it is now understood to

Figure 8.22 Equilibrium between various states of TTR tetramers, monomers, and pathogenic aggregates. The functional, ligand-bound tetramer is in equilibrium with the free tetramer as well as the folded monomer. When the folded monomer becomes denatured, it can misfold and is prone to oligomerization; oligomers form disease-causing aggregates and amyloid fibrils.

Source: Adapted from Johnson, S. M.; Connelly, S.; Fearns, C.; et al. (2012). *J Mol Biol* 421:185–203. With permission from Elsevier.



be much more common and found not only in those with specific genetic mutations but also among those without these mutations. Figure 8.22 illustrates the equilibrium between various states of TTR, from the thyroxine (hormone)–bound tetramer to the misfolded and aggregated monomer. The functional ligand-bound tetramer (*left*) is the most thermodynamically stable structure and is in equilibrium with the free tetramer, as well as with the correctly folded monomer. However, when the monomer is misfolded, it becomes more prone to aggregation and generates oligomers, aggregates, and fibrils, which are pathogenic. In some patients, transthyretin amyloidosis, a serious disease that results from TTR aggregates being deposited in various tissues and causing damage, is caused by mutations in the TTR protein that destabilize the free tetramer, shift the equilibrium to the right, and allow more of the free monomer to be available to misfold and eventually aggregate.

Jeffery Kelly's lab at the Scripps Research Institute developed a strategy to stabilize the properly folded tetramer via binding of a small molecule, thereby shifting the equilibrium away from the misfolded monomer. The TTR tetramer binds the small molecule thyroxine; however, at any given time, only a small percentage of the two thyroxine binding sites on TTR are occupied. An X-ray crystal structure of thyroxine bound to TTR showed thyroxine bound to the inner and outer subsite through multiple hydrophobic and H-bonding interactions, and this structure was used to design molecules that make similar interactions (Figure 8.23). In addition, screening approaches were also employed.

A general feature of all active molecules was the presence of substituted aryl rings connected by a linker. The substituents on the ring occupying the inner subsite make H-bonds with nearby serine residues and/or hydrophobic interactions. Aryl ring substituents in the outer subsite make salt bridges with an adjacent lysine residue as well as hydrophobic interactions that are similar to those made by the iodine atoms in thyroxine (Figure 8.23b). Tafamidis is approved to treat transthyretin-related hereditary amyloidosis. Its dichlorophenyl group sits in the inner subsite, occupying the same site as the terminal diiodophenyl group in thyroxine. The carboxylic acid makes a water-mediated interaction with both the glutamic acid and the lysine residues, also seen in the thyroxine–TTR co-structure. Acoramidis, also approved for ATTR cardiomyopathy, binds to the same site and makes similar biomolecular interactions to stabilize the tetramer.

As described earlier, drugs that work via PPIs are much less common than those that target enzymes or receptors. Despite the challenges, understanding how to design and optimize small molecules that modulate PPIs is an emerging area with

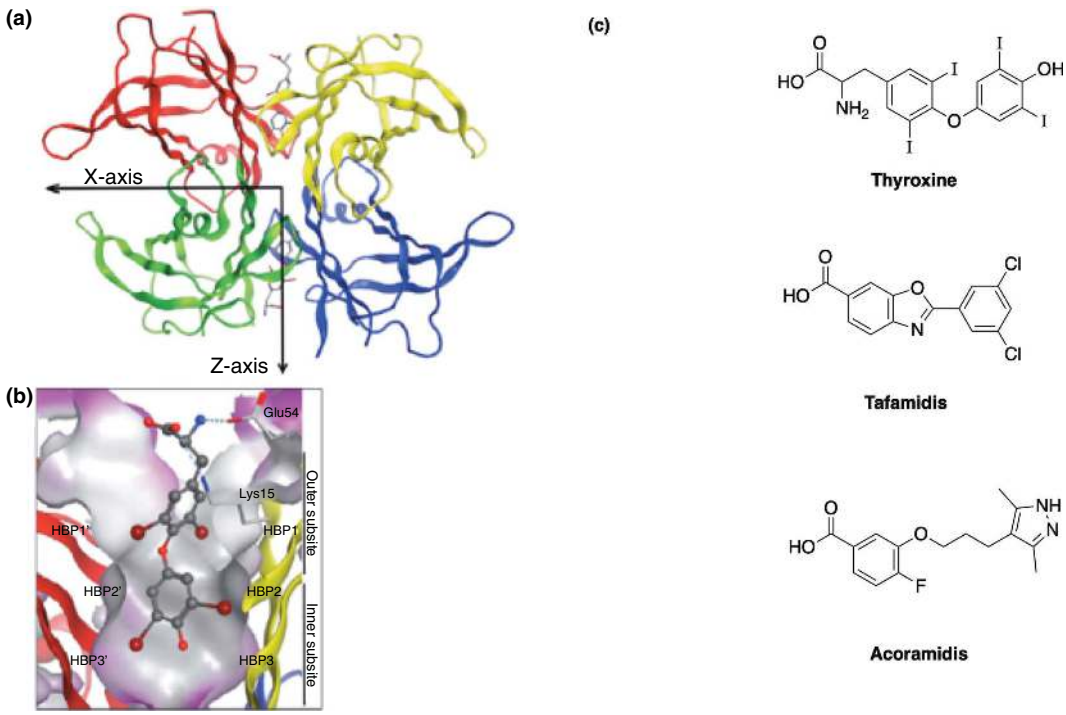


Figure 8.23 Design of molecules to stabilize TTR tetramer. (a) X-ray structure of TTR tetramer (each monomer is shown in a different color; two bound molecules of thyroxine are seen at yellow–red and blue–green interfaces. (b) Binding site of thyroxine with molecular interactions (HBP = hydrophobic pocket). The inner subsite makes hydrophobic interactions with one aryl ring of thyroxine; the second aryl ring sits in the outer subsite and makes interactions with hydrophobic groups as well as with the adjacent lysine (Lys15), and the terminal amine interacts with a glutamic acid (Glu54). (c) Structures of natural ligand thyroxine and drugs tafamidis and acoramidis.

Source: Adapted from Johnson, S. M.; Connelly, S.; Fearns, C.; et al. (2012). *J Mol Biol* 421:185–203. With permission from Elsevier.

considerable success since the 2000s. There are a number of targets and approaches that are under active investigation and from which drug candidates are likely to soon emerge. One such target class are heat shock proteins (HSPs), which are up-regulated in response to stress and have roles in cancer and diseases involving protein misfolding, with several drugs in clinical trials. Another target class, bromodomains, is also an active area with a number of compounds in clinical trials. These proteins are involved in binding to acetylated lysines on histones and have a role in the proliferation of cancer cells. New areas, such as the design of small molecules that will correct deficient or mutated proteins, is another exciting area of opportunity.

8.4 LIPIDS OF CELL MEMBRANES AS DRUG TARGETS

The hydrophobic interactions necessary for membrane integrity, and even for properly maintaining protein structure, can be affected by drugs that target another atypical drug target—lipids. Until recently, lipids and the membrane structures they form were considered inert structural molecules that were devoid of physiological properties. However, today there is a greater appreciation that they can have a significant impact on physiological and disease processes. For example, certain lipid structures can sequester proteins and thereby indirectly play a role in cell signaling and protein transport. In addition, lipids can directly bind enzymes and either activate or inactivate them. Therefore, by modulating lipids with small molecules, one has the potential to affect a large number of systems.

Unfortunately, understanding how lipid composition affects membranes and proteins is a very immature field. Consequently, there is no universal agreement on the mechanism of drugs that target lipids, let alone how to design drugs to do so. One reason for this lack of understanding is that few methods are available to study lipid bilayers in their natural setting, primarily because the lipid bilayer is a dynamic supramolecular structure rather than a single molecule with a defined action, such as an enzyme. In most cases, models must be used based on membranes that are reconstituted. Methods such as NMR and specialized microscopy can visualize the changes in membrane structure that result from drug interaction; however, drugs that interfere with the lipid structure and membranes are challenging to directly evaluate, since the membrane is not a single entity but a fluid structure. One example of a useful detection method is atomic force microscopy (AFM) that visualizes changes in the cell surface. Phenotypic assays can also be included in the compounds' characterization to determine whether interference in lipid structure has a physiological effect.

8.4.1 Drugs Acting at Lipids in Membranes Act as Anti-Infective Agents

Only a handful of drugs are known to work by modifying the lipid membrane, and they are primarily anti-infective agents. Some of these agents work by disrupting the cell membrane and allowing ions and nutrients to escape, thereby causing cell death. Others bind to lipid components of the cell membrane, such as lipid II of bacterial cell membranes or ergosterol of fungal cell membranes. However, the specific mechanism by which the membrane is disrupted varies, and designing molecules to do so selectively is quite challenging. One example is shown in what follows; other examples can be found in Chapter 12.

Daptomycin (**Figure 8.24**), isolated from cultures of *Streptomyces roseosporus*, is a branched cyclic lipopeptide. The macrocycle is made up of 10 amino acids, and the branch consists of three linked amino acids that terminate with a decanoic amide moiety. Of the 13 amino acids that comprise daptomycin, five are unnatural (shown in blue), either in stereochemistry (D rather than L) or in structure (3-methylglutamate and kynurenine). In addition, a 4–amino acid sequence in the ring (DADG, shown in red) is a calcium-binding motif observed in many calcium-binding proteins. Daptomycin is approved to treat Gram-positive skin infections, including those caused by methicillin-resistant *Staphylococcus aureus* (MRSA), and bacteremia. It represents one of only three new classes of antibacterial drugs approved since the 1970s.

Daptomycin targets the Gram-positive bacterial membrane, and its activity requires the presence of Ca^{2+} at physiological concentrations. Its mode of action is still controversial. One hypothesis is that a micellar structure containing multiple monomers is formed

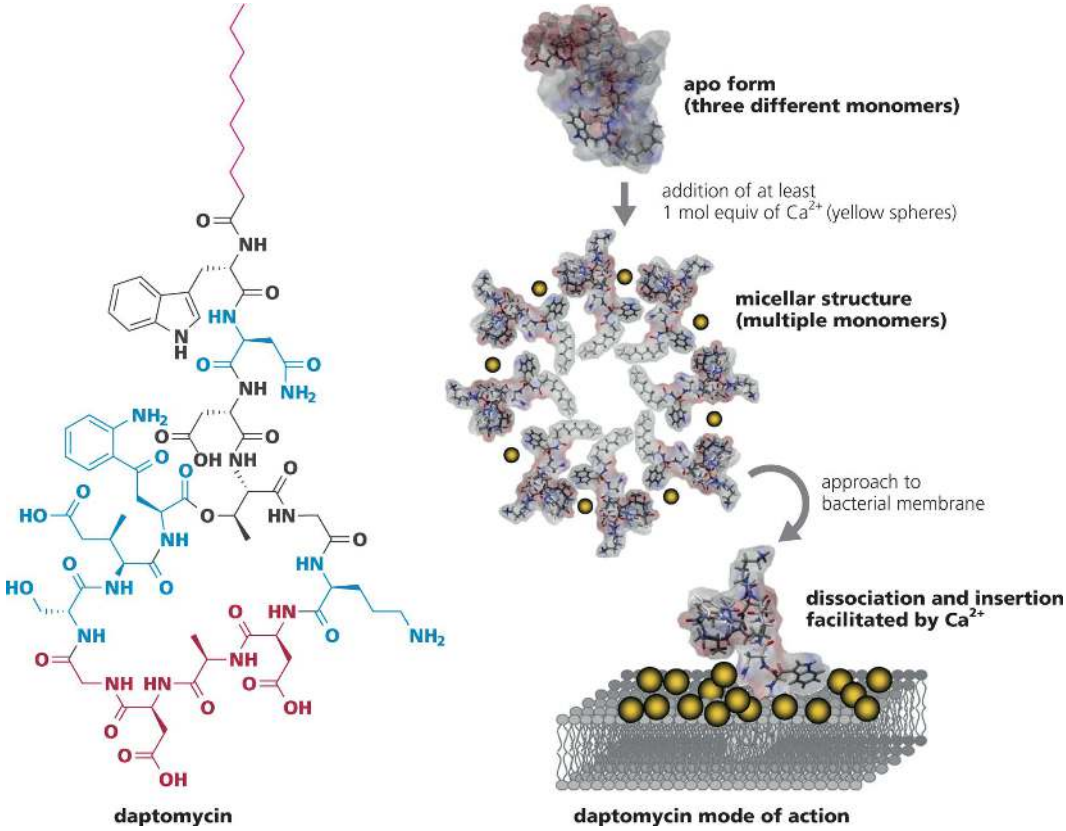


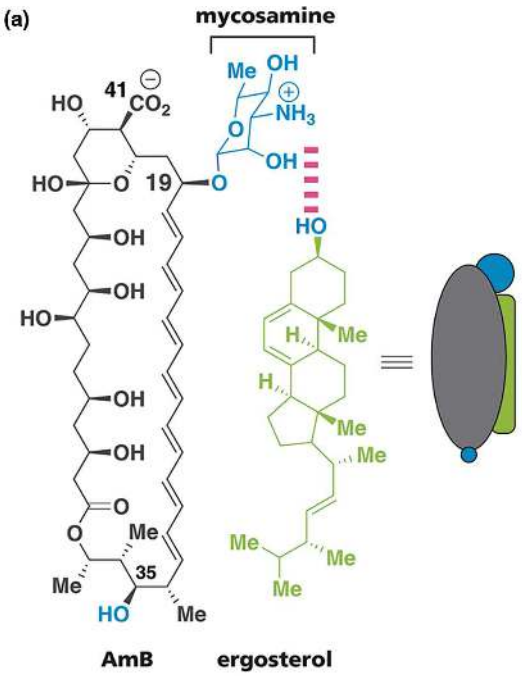
Figure 8.24 Structure of daptomycin and proposed mode of action. *Left*, daptomycin consists of a 10–amino acid macrocycle and pendant-acylated tripeptide. Unnatural amino acids are shown in blue, calcium-binding sequence is depicted in red at the bottom, and lipophilic tail is shown in magenta at the top. *Right*, micellar structure forms in the presence of calcium ions (depicted as yellow spheres) without change in daptomycin's conformation, which favors approach to the cell membrane. Once at the membrane, the supramolecular structure dissociates; daptomycin embeds into the membrane and disrupts it.

Source: (b) Adapted from Scott, W. R. P.; Baek, S.-B.; Jung, D.; et al. (2007). *Biochim Biophys Acta, Biomembr* 1768:3,116–3,126. With permission from Elsevier.

in the presence of calcium ions, allowing the daptomycyin micelle to approach the bacterial cell wall (see Figure 8.24). Once there, the supramolecular structure dissociates, and the monomers insert into the bilayer and oligomerize, thereby disrupting the membrane structure and causing efflux of potassium ions, loss of membrane potential, and eventual cell death. However, other models are also favored.

Another example of an anti-infective agent targeting lipids is the antifungal drug amphotericin B (AmB). This polyene macrolide consists of a polar face (displaying polyol), a hydrophobic face (displaying polyene), and a mycosamine unit. Fungi are eukaryotes that use ergosterol to form their cell membranes, as opposed to cholesterol that is used by animals. While it has been proposed for many years that amphotericin B forms pores in fungal cell membranes by making hydrophobic interactions with ergosterol and forming a macromolecular structure that makes the membrane permeable (Figure 8.25a), this model has been supplanted by one in which amphotericin B acts by sequestering ergosterol and removing it from the lipid bilayer, as seen in Figure 8.25b, as its primary mechanism.

Drugs that work by interacting with lipids constitute an evolving area. The few that are known have been identified by phenotypic screening, and their mechanisms were elucidated well after they were discovered. In many cases, their specific mode of action is still being studied and elucidated. However, as more is understood, strategies to design drugs *a priori* to interfere with lipids may be forthcoming. By modulating lipid structure, small-molecule modulators have the potential to affect membrane-bound protein structure, lipid signaling, and cell membrane integrity. While drugs in this class are primarily anti-infective agents, future agents may target central nervous system disorders, cancer, and a wider variety of therapeutic areas.



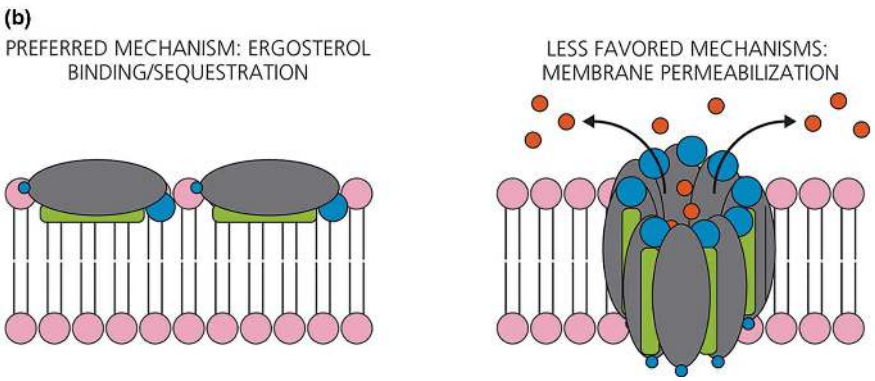


Figure 8.25 Structures of amphotericin B and ergosterol and proposed mechanisms. (a) Amphotericin B (AmB) consists of a macrocyclic polyol that displays a polar and an apolar face (black), with a polar mycosamine unit (blue) appended. Ergosterol (green) is the main constituent of fungal membranes. The polar carbohydrate is proposed to hydrogen-bond (red dashed line) to the alcohol of ergosterol; the polyene has hydrophobic interactions with ergosterol, and the C35-OH (blue) can hydrogen-bond to the polar end of the membrane lipid. (b) Two proposed mechanisms based on lipid interaction. The preferred mechanism involves amphotericin B sequestering ergosterol via 1:1 binding. The less-favored mechanism requires the ergosterol–amphotericin B complex to form a macromolecular structure that embeds in the membrane, making pores.

Source: Adapted from Gray, K. C.; Palacios, D. S.; Dailey, I.; et al. (2012). *Proc Natl Acad Sci USA* 109:2,234–2,239. With permission from the National Academy of Sciences.

8.5 SUMMARY

While most drug discovery efforts have focused on receptors and enzymes as drug targets, there are many other opportunities. Although there are inherent difficulties in targeting protein–protein interactions and interfering with lipid structures, many natural products have been found to do just that. Drugs such as Taxol and daptomycin are large complex molecules whose therapeutic effects were identified well before their mechanism of action. While the area of designing small modulators of lipid structures is in its infancy, designing modulators of PPIs is much more advanced. Recent approaches rely on understanding the binding interface and designing molecules that either mimic one of the binding partners or can interact with specific residues at the interface. In addition, new modalities, such as PROTACs and molecular glues, take advantage of the ability of small molecules to bind to proteins and direct or change the protein's function.

CASE STUDY 1 DISCOVERY OF VEPDEGESTRANT

One of the first PROTACs to reach late-stage development is vepdegestrant (ARV-471), invented at Arvinas. This drug targets the estrogen receptor (ER) and, as of early 2025, is in Phase III clinical trials for those with advanced or metastatic ER-positive/HER2-negative (ER+/HER2–) breast cancer. Metastatic breast cancer is the second leading cause of death in US women, and 70% of breast cancers are ER-positive. There are two isoforms of ER, alpha and beta, and they can homodimerize or heterodimerize and then translocate to the nucleus. In the nucleus, they bind co-activators to form a transcriptionally active complex that promotes proliferation and survival of breast cancer cells. Currently, most ER-positive breast cancers are treated with endocrine therapy, using SERMs (selective estrogen receptor modulators), such as tamoxifen, and also aromatase inhibitors (AIs), like anastrazole or letrazole. Resistant tumors are treated differently and respond to a combination of CDK4/6 kinase inhibitors like palbociclib and fulvestrant, a selective estrogen receptor degrader (SERC). Fulvestrant works by a different mechanism than PROTACs. It is a full ERα competitive antagonist that prevents ER dimerization and translocation to nucleus, and therefore, the degradation of ER is accelerated, but one shortcoming is its poor solubility and a lack of oral bioavailability.

Drug Discovery Efforts

The design of PROTACS that would degrade the estrogen receptor focused on multiple ER binders as well as multiple E3 recruiting ligands and linkers during the early drug discovery efforts. Some promising results were found with analogs based on raloxifene, a SERM, and lenalidomide, an E3 ligase ligand. As seen in **Figure 8.26**, compound **1**, with a piperazine-based linkage, had *in vivo* exposure as well as promising DC_{50} results (concentration to degrade 50% of ER protein). A similar analog based on raloxifene and lenalidomide, compound **2**, containing a polyethylene linkage, showed off-target activity, acting as a GSPT1 degrader. Although off-target, GSPT1 (G1 to S phase transition 1) is a translation termination factor overexpressed in some cancer cells and is also a cancer target. More promising for selective ER degradation were the PROTACS, such as compound **3**, based on the third-generation SERM lasofoxifene combined with lenalidomide E3 ligase ligand. This compound exhibited *in vivo* efficacy in an MCF xenograft (a breast cancer cell line), and when modified with a trifluoromethyl substituent on the pendant aryl ring, oral efficacy was seen in dogs. With an optimized piperazine-based linker, compound

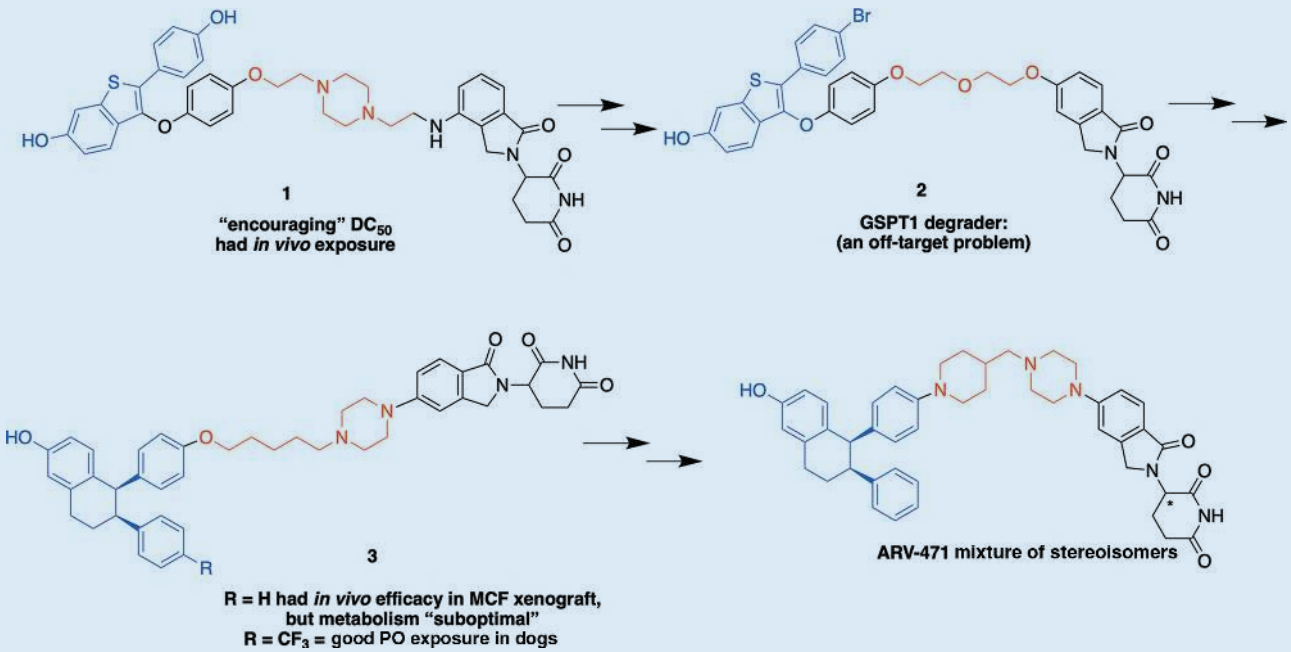


Figure 8.26 Early discovery leads in the development of vepdegestrant (ER ligand in blue, linker in red). Compounds **1** and **2** were based on the ER ligand raloxifene, with 3 and ARV-471 based on lasofoxifene.

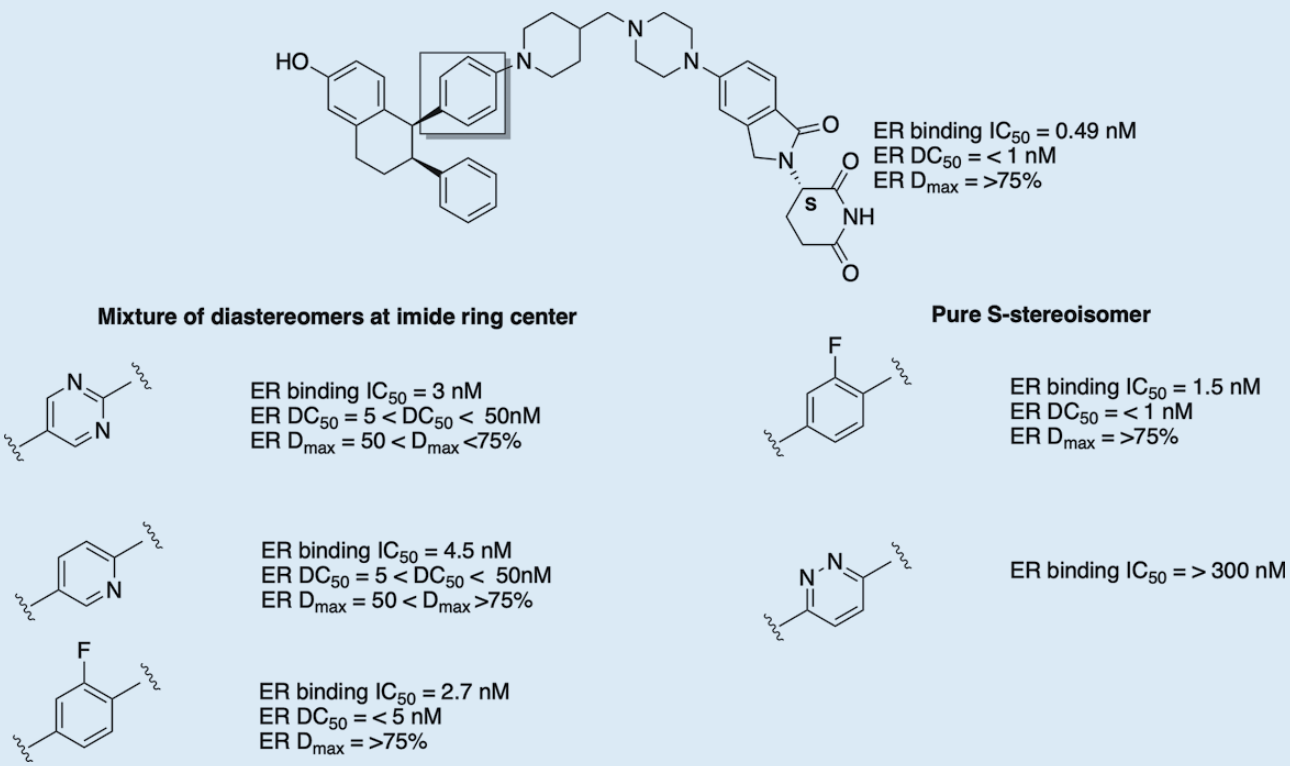


Figure 8.27 Variations in the aryl ring of the lasofoxifene portion of vepdegestrant analogs.

ARV-471 (as a mixture of diastereomers) had *in vivo* efficacy suitable for further preclinical development. Further investigation showed that the S enantiomer of the lenalidomide portion was more effective than the R, leading to vepdegestrant.

In addition to optimization of the linker through conformational constraint, bioisosteres for the pendant aryl ring of lasofoxifene were investigated, as seen in Figure 8.27, although none provided an improvement over vepdegestrant.

Analogs were tested for binding to the estrogen receptor (ER binding IC_{50}) and ability to degrade the ER, with DC_{50} measuring concentration to result in 50% degradation, and ER D_{max} measuring the maximum percent degraded. As seen in Figure 8.27, of the analogs tested, vepdegestrant had most potent binding, as well as most efficient degradation of the receptor. ER degradation by vepdegestrant was evaluated in multiple ER positive cell lines, with DC_{50} of ~1 nM, including clinically relevant drug-resistant *ESR1* variants Y537S and D538G. A low-resolution cryo-EM structure of the ternary complex between vepdegestrant, the ER ligand binding domain, and a portion of the cereblon E3 ligase complex was solved and is shown in Figure 8.28.

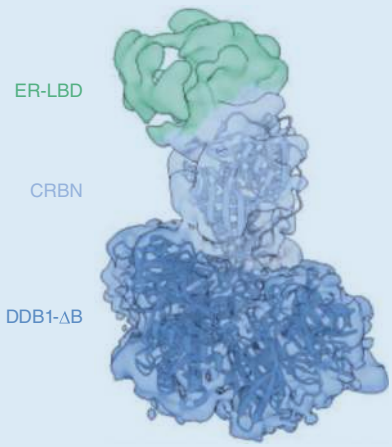


Figure 8.28 Low-resolution cryo-EM structure of ternary complex between the vepegestrant, the ER ligand binding domain (ER-LBD), and a cereblon (CRBN) containing portion of the E3 ligase (DDB1-ΔB).

Source: Adapted from Gough, S. M.; Flanagan, J. J.; Tey, J.; et al. (2024). *Clin Cancer Res* 30 (16):3,549–3,563. With permission from AACR Journals.

Synthesis

The process synthesis developed for vepdegestrant, outlined in Figure 8.29, followed the same bond disconnection strategy as the discovery phase synthesis. Improvements included the avoidance of chiral chromatography and streamlining by crystallizing intermediates. The convergent synthesis involved the preparation of key aldehyde and piperazine components **4** and **5**, containing the lasofoxifene and lenalidomide binding portions, respectively. Reductive amination was optimal using triacetoxyborohydride to form the key bond in the final step. Synthesis of the aldehyde component began with the triflate enol ether **6**, and the two aryl rings were introduced in successive Suzuki couplings to afford compound **7**. In the original synthesis, this intermediate was hydrogenated, and the enantiomers of the resulting *cis*-product were separated by chiral chromatography. In the process synthesis, compound **7** was coupled with **8** in a Buchwald–Hartwig reaction to afford the penultimate aldehyde precursor **9**. Numerous chiral hydrogenation catalysts for this hindered alkene were tried, but generating optically pure product was unsuccessful; ultimately, a reduction of the olefin using palladium on carbon, followed by

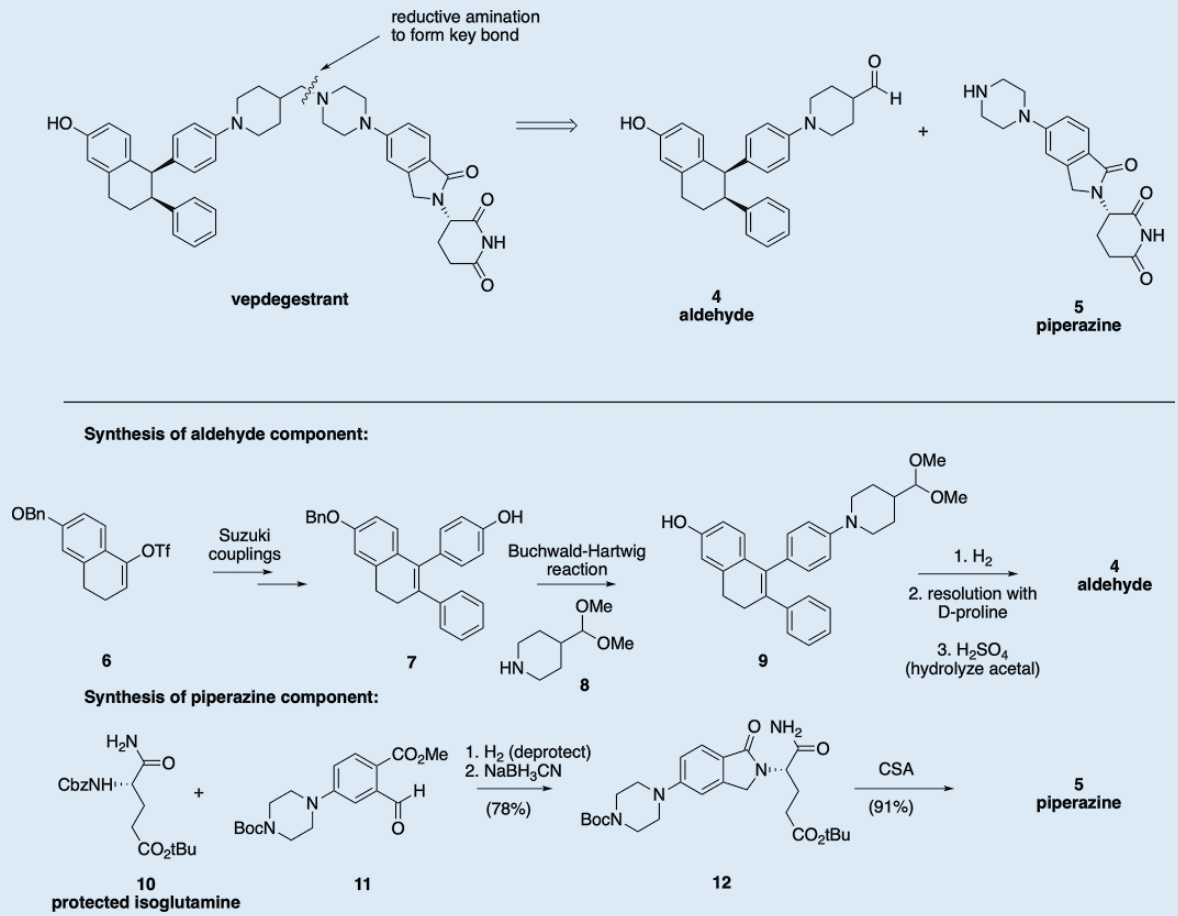


Figure 8.29 Outline of process synthesis of vepdegestrant.

resolution with D-proline and treatment with sulfuric acid to hydrolyze the acetal, afforded aldehyde **4**.

Synthesis of the piperazine **5** started with Cbz-protected isoglutamine **10**. Deprotection and then reductive amination with the aldehyde **11**, followed by cyclization onto the methyl ester, afforded the cyclized product **12**. Cyclization with camphor sulfonic acid gave the piperazine coupling component **5** in high yield.

Preclinical and Clinical Results

Vepdegestrant exhibited good oral bioavailability and *in vitro* efficacy in multiple ER-dependent cancer cell lines, where degradation of the ER was detected. Investigation for tumor growth inhibition was carried out in xenograft models with both an ER-dependent breast cancer cell line (MCF-7) and a patient-derived xenograft model (PDX) with an **es**trogen **r**eceptor gene mutation *ESR1* (Y537S). Tumor growth inhibition (TGI) was high in both models, with doses of 30mg/kg resulting in >100% TGI (tumor regression), and at 30 mg/kg TGI, results were superior to the standard of care, fulvestrant, at 200 mg/kg. Based on Western blot studies, very high percentages of ER reduction were seen.

Phase I clinical trials showed that vepdegestrant was well tolerated and set the dose at 200 mg/kg. Further trials in combination with palbociclib were done in patients who had previously received other treatments. Side effects included neutropenia and decreased blood cell count, which are common with palbociclib. A high clinical benefit rate of 63% was found, and in patients with mutant ESR1, the rate was 72%. The median progression-free survival (MPS) was 11 months, with a very high level of reduction of circulating tumor DNA. Based on these promising results, vepdegestrant has advanced to Phase III clinical trials and, in February of 2024, received FDA Fast Track designation.

CASE STUDY 2 DISCOVERY OF MARAVIROC

The chemokine receptors CXCR4 and CCR5 are members of a large family of GPCRs that bind to protein ligands (chemokines) and play a role in the immune response. The endogenous ligands for the chemokine receptors are the peptides RANTES, MIP-1α, MIP-1β, and TNF. In addition to their physiological role, in 1999 CCR5 and CXCR4 were identified as essential co-factors for entry of the HIV into T cells. The glycoprotein gp120, found on the surface of the viral particle, binds to a CD4 molecule on the surface of the host immune cells and then uses either the CCR5 or CXCR4 receptor as a second (co-)receptor. Once the viral particle is bound, it fuses to the cell, releases its genetic materials, and infects the cell (Figure 8.30).

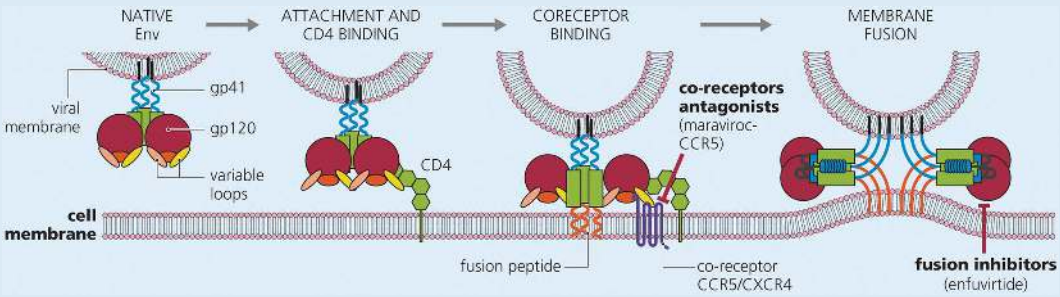


Figure 8.30 Fusion of HIV viral particle. Glycoprotein gp120 (shown in red) binds to a T cell CD4 receptor (depicted as green hexagons) on the surface of the host T cell and then associates with the CCR5 or CXCR4 co-receptor (shown in purple). The fusion of the viral particle, mediated by fusion peptides, and the release of viral nucleic acid material into the cell occur. Maraviroc blocks the protein–protein interaction between gp120 and CCR5 required for fusion. Fusion inhibitors block a later step and prevent the virus from releasing its genetic material into the T cell.

Source: Adapted from Didigu, C. A. and Doms, R. W. (2012). *Viruses* 4:309–324. With permission from MDPI.

The validation of the role of CCR5 in HIV infection was based on reports describing individuals who were both healthy and resistant to HIV infection and had a genetic mutation resulting in a lack of the CCR5 receptor. This observation suggested that compounds which blocked the binding of HIV particles to CXCR4 and CCR5 could be effective anti-HIV agents. Although CXCR4 and CCR5 are GPCRs, the goal was not to prevent the signal transduction they initiate but, rather, to block binding between gp120 and the chemokine receptor. As such, the target is treated as a PPI rather than a classical receptor.

High-throughput screening in an assay for the inhibition of binding of radiolabeled ligand (MIP-1β) to CCR5 led to a series of hits that were filtered on the basis of Lipinski's rules, as well as target affinity and ligand efficiency. This process led to two compounds that were pursued, UK-107543 (**1**) and UK-179645 (**2**), shown in Figure 8.31. Both inhibited binding of the natural ligand to CCR5 but had no antiviral activity. Additionally, UK-107543 inhibited CYP450 2D6, which predicted future issues with drug–drug interactions.

Molecular modeling suggested that the pyridine nitrogen in the initial hit **1** (shown in red) may contribute to CYP inhibition, so it was replaced with the isosteric carbon to give **3**; this modification reduced the CYP inhibition but increased lipophilicity (Figure 8.32). The incorporation of an amide, inspired by the structure of **2**, afforded **4**, which had reduced lipophilicity and now exhibited antiviral activity. Synthesis of both enantiomers of **4** proved that the more potent isomer was the *S*-enantiomer, with 44 times higher inhibition of MIP-1β binding; the *R*-enantiomer showed no measurable antiviral activity, suggesting only the eutomer contributed to the efficacy observed. Further modifications focused on maintaining potency, avoiding hERG liabilities, and optimizing permeability. The replacement of the phenyl amide with a tetrahydropyran amide and conformational constraint of the piperidine into a tropane afforded a highly potent compound, with no hERG liabilities but still inadequate PK characteristics. This constrained piperidine modification was predicted to reduce Cyp450 2D6 inhibition due to increased steric hindrance that would remove a key interaction with an aspartic acid residue found in the Cyp enzyme (Figure 8.32). The tetrahydropyran group was incorporated as a way to limit hERG liability, by adding polar functional groups without limiting permeability.

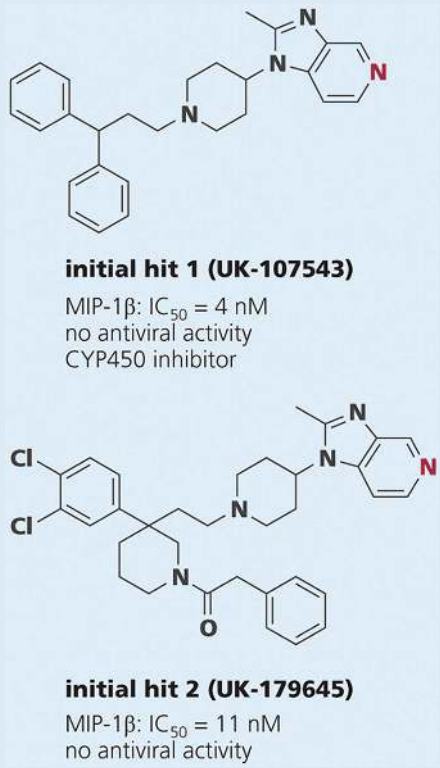


Figure 8.31 Structures of early hits in maraviroc drug discovery. Compounds **1** and **2** (with the pyridine nitrogen shown in red) both inhibited MIP-1β interaction with CCR5, indicative of binding to CCR5, but they exhibited no anti-HIV activity.

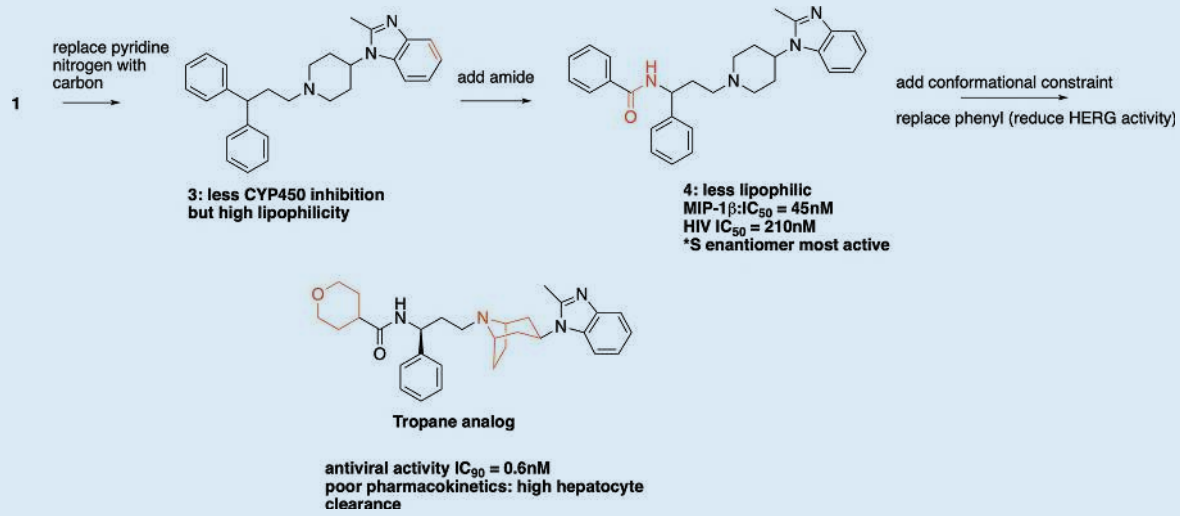
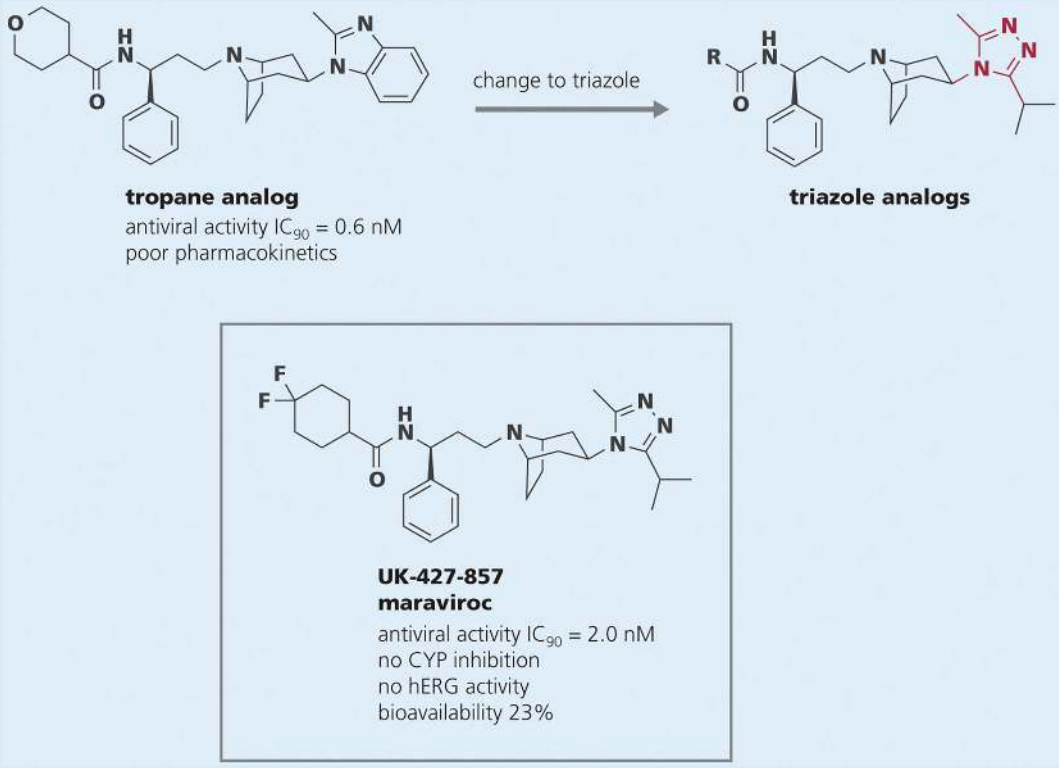


Figure 8.32 Structures illustrating lead development to give rigid tropane analog. Isosteric replacement of pyridyl nitrogen of **1** with carbon (**3**) reduced CYP450 inhibition but made the analog too lipophilic. Inserting an amide linkage improved lipophilicity and cell-based anti-HIV activity (**4**), and there was a strong preference for the (S)-enantiomer. Conformational constraint of the piperidine and replacement of phenyl with tetrahydropyran led to a tropane analog, which had improved antiviral activity but poor pharmacokinetics (changes in each series shown in red).

The tropane analog in Figure 8.32 was shown to be a selective antagonist of CCR5 binding to its chemokine ligands. It was also a potent inhibitor of HIV infection in blood cells and was noncytotoxic. The anti-HIV activity is proposed to result from the conformational changes to CCR5 that disrupt the interaction with gp120. However, the tropane analog exhibited poor pharmacokinetics, with rapid metabolism and <10% oral availability in dogs. Replacement of the benzimidazole with a triazole reduced both molecular weight and lipophilicity. Further modification of the amide eventually led to UK-427857 (maraviroc), which exhibited potent anti-HIV activity and no CYP or hERG activity (Figure 8.33). Although its bioavailability is still somewhat low (23%), it represents an improvement over the benzimidazole series.

Figure 8.33 Development of maraviroc from tropane analog. Replacement of benzimidazole with a triazole (in red) and varying the amide group R reduced lipophilicity and improved druglike features and resulted in maraviroc.



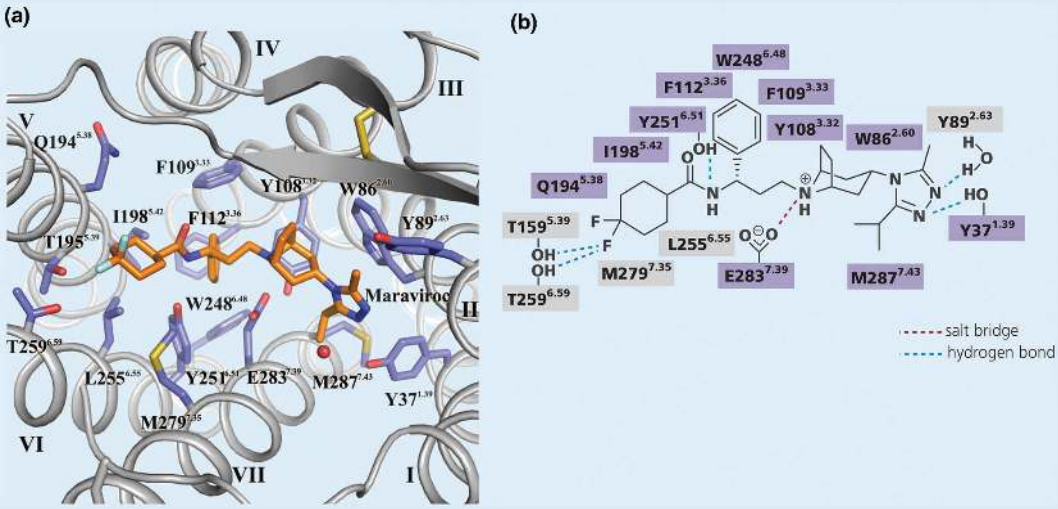


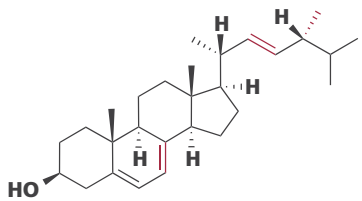
Figure 8.34 X-ray structure of maraviroc bound to CCR5. (a) Maraviroc (depicted in orange) is shown in its binding pocket, away from its originally presumed site of action. (b) Structure of maraviroc showing key interactions, including an ion–ion interaction between the tropane nitrogen and aspartate 283 (E283), hydrogen bonds between the fluorine and threonine hydroxyls (T195 and T259), and hydrogen bonding between the triazole and tyrosine 37 and between tyrosine 251 and the amide nitrogen. Residues previously identified by mutagenesis studies as important for binding are in purple squares.

Source: From Tan, Q.; Zhu, Y.; Li, J.; et al. (2013). *Science* 341:1,387–1,390. With permission from AAAS.

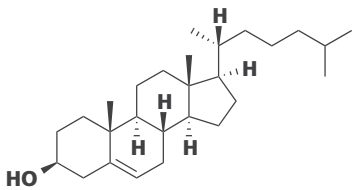
Maraviroc was approved by the FDA in 2007 for use in patients infected with HIV. It is interesting to note that an X-ray co-crystal structure of maraviroc and CCR5 was published in 2013, showing that maraviroc binds at an allosteric site on CCR5 rather than the presumptive gp120 binding site (**Figure 8.34**). A number of key interactions are evident and include an ion–ion interaction between the tropane nitrogen and aspartate 283 (E283), hydrogen bonds between the fluorine and threonine hydroxyls (T195 and T259), and hydrogen bonding between nitrogens in the triazole and tyrosine 37 (Y37) and between the amide nitrogen and tyrosine 251 (Y251). Binding of the drug prevents the helix movements needed for gp120 to bind to the receptor.

8.6 REVIEW QUESTIONS

1. Explain why some drug targets (such as enzymes and receptors) are more druggable than others.
2. Describe three biophysical methods that can be used to detect binding between a small molecule and a protein.
3. How do molecular glues differ from bifunctional degraders; give an example of each.
4. Name two hit identification strategies that have been particularly successful in identifying modulators of PPIs.
5. Define hot spots and their importance in drug discovery.
6. Explain why drug design to target lipids is particularly challenging.
7. Explain the key components of a PROTAC.
8. On the basis of the examples in this chapter, identify at least two compounds that incorporate the medicinal chemistry concepts of conformational constraint and at least two that contain a privileged structure.
9. Explain the concept of stapling a peptide.
10. Compare the structures of ergosterol and cholesterol. What are the differences?



ergosterol



cholesterol

8.7 APPLICATION QUESTIONS

1. Research an example of a β-turn mimetic reported in the literature. Draw the structure and compare it to the peptide secondary structure.
2. Research one of the emerging PPI targets in Section 8.3.3. Describe how small-molecule modulators of it were identified, and draw their structures.
3. Research the magainins. Describe their structure and mechanism of action.
4. Design three different mimics of RGD (see **Figure 8.9**) that incorporate isosteres for the acid group.
5. Research one PROTAC (not covered in this chapter) that is in clinical trials. Identify the key components and challenges in its development.
6. Tacrolimus (or FK506) and cyclosporin are important drugs used to prevent transplant rejection. They have very complex mechanisms of action that involve protein–protein interactions. Choose one, and research its structure and general mechanism of action.

7. Review the structure of amphotericin B. Does it obey Lipinski's rules? For logP, use a program such as ChemDraw to generate a calculated value.
8. Review the MDM2 inhibitor structures in **Figure 8.7**. Orient them to propose how they share a common pharmacophore. Propose a pharmacophore that is consistent with this orientation.
9. Give one example of a small-molecule drug that is an interfacial binder.

FURTHER READING

Protein–Protein Interactions

Arkin M, Tang Y & Wells JA (2014) Small-molecule inhibitors of protein–protein interactions: progressing toward the reality. *Chem Biol* 21:1102–1114.

Azzarito V, Long K, Murphy NS & Wilson AJ (2013) Inhibition of α-helix-mediated protein–protein interactions using designed molecules. *Nat Chem* 5:161–173.

Bonetta L (2010) Interactome under construction. *Nature* 468:851–854.

Fry DC (2012) Small-molecule inhibitors of protein–protein interactions: how to mimic a protein partner. *Curr Pharm Des* 18:4679–4684.

Jin L, Wang W & Fang G (2014) Targeting protein–protein interaction by small molecules. *Annu Rev Pharmacol Toxicol* 54:435–456.

Petros AM, Dinges J, Augeri DJ et al. (2006) Discovery of a potent inhibitor of the antiapoptotic protein Bcl-xL from NMR and parallel synthesis. *J Med Chem* 49:656–663.

Pobbati AV, Chan SW, Lee I et al. (2012) Structural and functional similarity between the VgII1-TEAD and the YAP-TEAD complexes. *Structure* 20:1135–1140.

Scott DE, Bayly AR, Abell C et al. (2016) Small molecules, big targets: drug discovery faces the protein-protein interaction challenge. *Nat Rev Drug Disc* 15:533–550.

Walensky LD & Bird GH (2014) Hydrocarbon-stapled peptides: principles, practice, and progress. *J Med Chem* 57:6275–6288.

Chemically Induced Proximity and PROTACs

Bekes M, Langley DR & Crews CM (2022) PROTAC targeted protein degraders: the past is prologue. *Nat Rev Drug Discov* 21:181–200.

Hornberger KR & Arauho EMV (2023) Physicochemical property determinants of oral absorption for PROTAC protein degraders. *J Med Chem* 66(12):8281–8287.

Liu X & Ciulli A (2023) Proximity-based modalities for biology and medicine. *ACS Cent Sci* 9(7):1269–1284.

Schreiber SL (2024) Molecular glues and bifunctional compounds: therapeutic modalities based on induced proximity. *Cell Chem Biol* 31:1050–1063.

Tubulin Modulators

Dumontet C & Jordan MA (2010) Microtubule-binding agents: a dynamic field of cancer therapeutics. *Nat Rev Drug Discov* 9:790–803.

He L, Jagtap PG, Kingston DGI et al. (2000) A common pharmacophore for taxol and the epothilones based on the biological activity of a taxane molecule lacking a C-13 side chain. *Biochemistry* 39:3972–3978.

Stanton RA, Gernert KM, Nettles JH & Aneja R (2011) Drugs that target dynamic microtubules: a new molecular perspective. *Med Res Rev* 31:443–481.

RGD Mimetics

Andronati SA, Karaseva TL & Krysko AA (2004) Peptidomimetics-antagonists of the fibrinogen receptors: molecular design, structures, properties and therapeutic applications. *Curr Med Chem* 11:1183–1211.

Transthyretin

Johnson SM, Connelly S, Fearns C et al. (2012) The transthyretin amyloidoses: from delineating the molecular mechanism of aggregation linked to anthology to a regulatory-agency-approved drug. *J Mol Biol* 421:185–203.

Johnson SM, Wiseman RL, Sekimina Y et al. (2005) Native state kinetic stabilization as a strategy to ameliorate protein misfolding diseases: a focus on the transthyretin amyloidosis. *Acc Chem Res* 38:911–921.

Penchala SC, Connelly S, Wang Y et al. (2013) AG10 inhibits amyloidogenesis and cellular toxicity of the familial amyloid cardiomyopathy-associated V122I transthyretin. *Proc Natl Acad Sci USA* 110:9992–9997.

Vepdegestrant

Avery S, Buske JM, Chen D et al. (2024) Development of a commercial manufacturing process for vepdegestrant, an orally bioavailable PROTAC estrogen receptor degrader for the treatment of breast cancer. *Org Process Res Dev* 28:4079–4090.

Chen CH & Dong H (2023) Methods of manufacturing a bifunctional compound. WO2023009521A1.

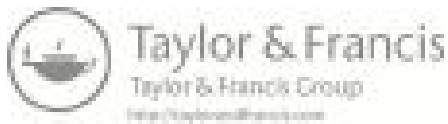
Rej RK, Thomas II JE, Acharyya RK, Rae JM & Wang S (2023) Targeting the estrogen receptor for the treatment of breast cancer: recent advances and challenges. *J Med Chem* 66:8339–8381.

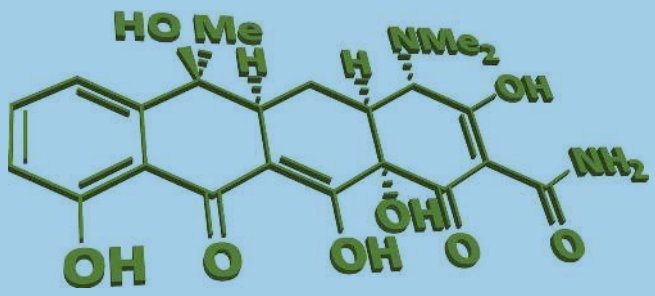
Snyder LB et al. (2021) The discovery of ARV-471, an orally bioavailable estrogen receptor degrading PROTAC for the treatment of patients with breast cancer. *Cancer Res* 81: Abstract 44.

Maraviroc

Palani A & Tagat JR (2006) Discovery and development of small-molecule chemokine coreceptor CCR5 antagonists. *J Med Chem* 49:2851–2857

Wood A & Armour D (2005) The discovery of the CCR5 receptor antagonist, UK-427,857, a new agent for the treatment of HIV infection and AIDS. *Prog Med Chem* 43:239–271.


Taylor & Francis
Taylor & Francis Group



DNA and RNA as Drug Targets

9

LEARNING OBJECTIVES

- Explain the rationale for targeting DNA and RNA.
- Name and describe the various ways small molecules can make noncovalent interactions with DNA.
- Identify several different mechanisms by which small molecules can cause modifications to nucleic acids.
- Describe how antisense oligonucleotides work.
- Learn examples of drugs that work by targeting DNA and RNA.

9.1 INTRODUCTION

Deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) are macromolecular polymers, each made up of four different nucleotides linked through 3′–5′ phosphodiester linkages (**Figure 9.1a**). The individual monomers that make up DNA are 2-deoxynucleotides that contain the nucleobases adenine (A), guanine (G), thymine (T), and cytidine (C). RNA monomers contain the bases adenine, guanine, uracil (U), and cytidine.

A DNA duplex consists of two polymers linked at the 3′ and 5′ positions by a phosphodiester, held together by noncovalent interactions between specific nucleobases in an anti-parallel fashion, forming a double helix (A–T and G–C; Watson–Crick pairing, shown in **Figure 9.1b**). The base pairs are stacked in the interior of the double helix, and the sugar–phosphate backbone sits on the exterior. B-DNA is considered the most relevant form of DNA; it displays two distinct grooves, a deep, narrow minor groove and a wider, shallow major groove (**Figure 9.2a**). Rather than double-stranded helices, RNA is often found as a single strand. It can self-associate with self-complementary regions after folding back on itself and form secondary structures with hairpin loops and bulges and can also form a double helix, usually in the A-form (**Figure 9.2b**). This is a shorter, wider form of a double helix compared to the DNA helix with a deep, narrow major groove and a wide, shallow minor groove.

Although the structural difference at the molecular level between RNA and DNA is only in the substitution of uracil for thymine and an additional hydroxyl group in the furanose ring, as can be seen in **Figure 9.1a**, this translates to significant differences in secondary and tertiary structure. There are also significant differences in the function of the two oligonucleotides. The function of DNA is to maintain and transmit genetic information, either to a daughter cell or to RNA (transcription). It is located in the nucleus, where it is typically bound to proteins called histones. A primary role of RNA is to direct the synthesis of proteins on the ribosome (translation). There are several forms involved in this process: messenger RNA (mRNA) is transcribed from DNA and carries genetic

DOI: 10.1201/ 9781003465607-11

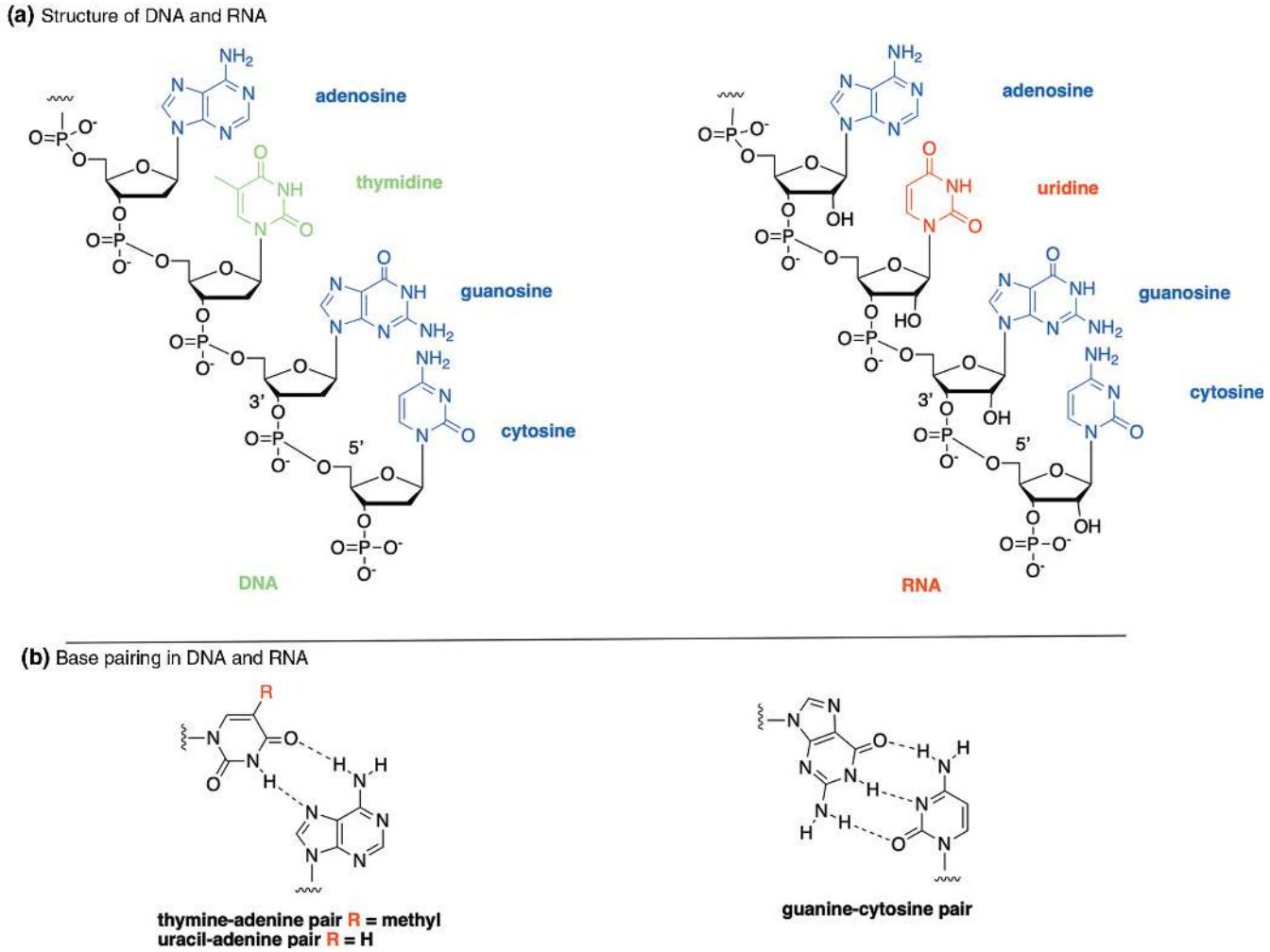


Figure 9.1 Backbone structure and base pairing of DNA and RNA. (a) DNA and RNA segments containing the four individual nucleobases adenine, thymine (green)/uracil (red), guanine, and cytosine. (b) Watson–Crick base pairing between adenine and thymine (R = Me)/uracil (R = H) and between cytosine and guanine.

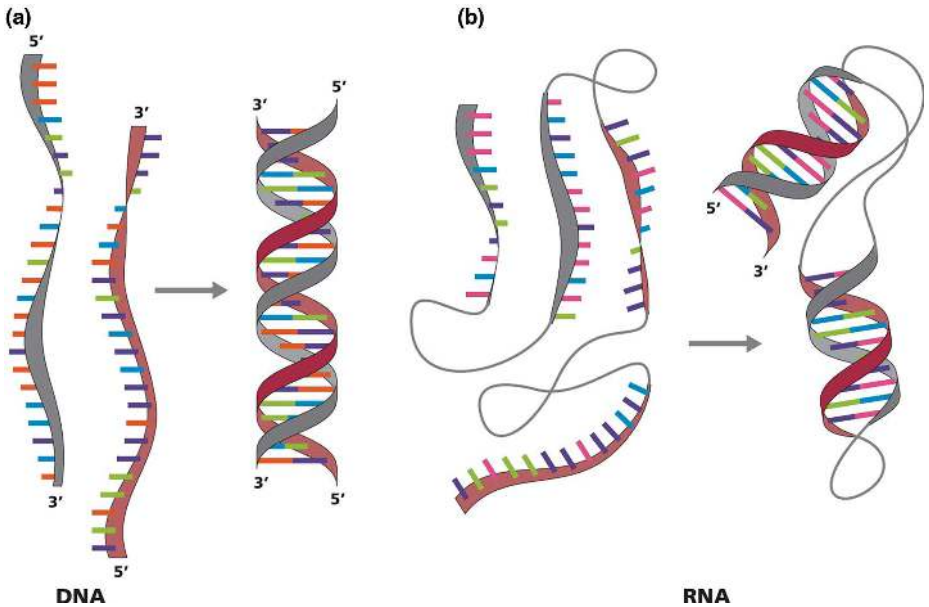


Figure 9.2 Macromolecular structures of DNA and RNA. (a) DNA molecules exist in pairs (red and gray strands) with complementary sequences that enable the two molecules to form a continuous anti-parallel double helix. (b) RNA molecules frequently exist as single strands that contain self-complementary regions. In this diagram, the segments colored red contain sequences that are complementary to sequences colored gray. In this example, the RNA molecule self-assembles and adopts a conformation that forms two double helices. For both DNA and RNA, each colored projection on each strand represents a specific nucleobase.

Source: From Kuriyan, J.; Konforti, B.; and Wemmer, D. (2012). *The Molecules of Life: Physical and Chemical Principles*. Garland Science.

information into the cytoplasm; transfer RNA (tRNA) moves to the ribosome and carries the three-base code for each amino acid needed for protein synthesis; and ribosomal RNA (rRNA) is part of the ribosome, a protein–RNA complex where protein synthesis is carried out. More recently, even more types and functions of RNA have been identified: Ribozymes are RNAs that possess enzymatic activity, whereas microRNA (miRNA) and small interfering RNAs (siRNA) are naturally occurring 20- to 25-nucleotide segments of double-stranded RNA whose typical function is to block gene expression. As such, RNA is essential for protein synthesis and, ultimately, cell division.

Historically, many compounds that bind to nucleic acids were isolated or derived from natural products. The original compounds, often detected on the basis of their antibacterial and/or anti-proliferative effects, were studied, and the mechanism of action was determined later. Since DNA and RNA are conserved across all living species, it is not surprising that various organisms have evolved small molecules that bind to nucleic acids as a defense mechanism. Other agents that bind to DNA have more nefarious origins (*vide infra*), and their beneficial effects were based on clinical observations. More recently, scientists have been developing methods to design compounds that bind with high affinity and high selectivity to specific sequences of DNA and RNA.

Methods to measure the effects of molecules that bind to DNA rely on biophysical techniques, such as surface plasmon resonance (SPR), electrophoretic mobility shift assays (EMSA), and melting-temperature assays. These assays, in the broadest terms, detect differences in size or stability between the duplex DNA and a small molecule bound to duplex DNA. Other techniques use spectroscopic changes that occur upon binding, such as fluorescence competition assays. These may rely on changes in fluorescence of an external probe that binds to DNA. Identical or similar assays are also used to detect small-molecule binding to RNA. Secondary assays used to measure downstream events that reflect changes to DNA or RNA, such as protein synthesis and cell proliferation, are similar to those previously described.

Small molecules that interact with DNA are primarily used to treat cancer, a disorder characterized by uncontrolled growth. Some of these drugs are also approved to treat autoimmune diseases, which are characterized by the proliferation of immune cells. This focus is due to the fundamental role DNA plays in replication. However, the challenges in developing selective DNA-binding agents usually make them second- or third-line therapies in autoimmune diseases, where other therapies are available. The current state of medicinal chemistry has not yet evolved to allow design of highly sequence-selective DNA-targeted drugs, and therefore, toxicity due to effects on non-target cells is a significant liability. Drugs that target RNA primarily affect protein synthesis, either in bacteria or in humans. Both DNA- and RNA-binding drugs act by a variety of binding mechanisms, and a summary of those mechanisms is given in **Table 9.1**. However, drugs that exclusively target DNA or RNA as their mechanism of action are uncommon, and most work via multiple effects. This area is rapidly evolving, with the first examples of drugs that bind to specific RNA sequences being approved.

9.2 DNA AS A DRUG TARGET

The role of DNA during cell division and proliferation has made directly targeting it a fruitful avenue to identify drugs, particularly anti-cancer agents. Binding to and/or damaging DNA prevents its accurate replication, often causing the processes of cell division and, eventually, cell proliferation to stop. Drugs that work through this mechanism are often highly potent and effective anti-cancer agents. However, these mechanisms are not selective, and significant toxicities due to effects on noncancerous dividing cells, such as those in the gastrointestinal tract or part of the immune system, limit the usefulness of these agents. Furthermore, certain DNA-damaging agents can cause secondary cancers independently, often years after the original treatment. The mechanisms by which drugs damage DNA involve two primary modes of interaction, reversible and irreversible. Reversible, noncovalent binders may either insert between the base pairs (intercalation) or lie in the minor groove or act via a combination of both. Some compounds that work by an irreversible mechanism contain a reactive electrophilic group that alkylates nucleophilic sites on DNA, while others generate a reactive radical that abstracts hydrogen atoms from the furanose ring in DNA, ultimately leading to chemical modification of the DNA strand.

Table 9.1 Examples of Drugs Directly Targeting DNA and RNA, Their Mechanisms of Action, and Therapeutic Use

Nucleic Acid Target	Drug Mechanism	Drug	Indication
DNA	Intercalation and minor groove binding	Doxorubicin, daunomycin, epirubicin	Cancer
		Mitoxantrone	Cancer, multiple sclerosis
DNA	Minor groove binding	Netropsin*	Probe
		Trabectedin	Cancer
DNA	Alkylation or metalation (cisplatin)	Cyclophosphamide, chlorambucil, cisplatin, thiotepa	Cancer, autoimmune diseases (cyclophosphamide)
DNA	Strand cleavage	Bleomycin, calicheamicin (as antibody conjugate)	Cancer
RNA	rRNA binding	Tetracycline	Acne, bacterial infections
		Erythromycin	Bacterial infections
RNA	Antisense	Mipomersen	Familial hypercholesterolemia
		Nusinersen	Spinal muscular atrophy
		Etiplersen	Duchenne muscular dystrophy
RNA	Splicing modifier	Risdiplam	Spinal muscular atrophy
		Branaplam*	Huntington's disease

*Indicates a research compound/not yet clinically approved compound.

9.2.1 Intercalation Is One of the Main Mechanisms by Which Small Molecules Interact with DNA

Drugs that bind to DNA reversibly act directly on DNA through two main mechanisms: intercalation and groove binding. **Intercalators**, typically planar aromatic molecules, insert between the DNA base pairs via π stacking and additional hydrophobic and electrostatic interactions (Figure 9.3a). The antibiotic proflavine, with its planar tricyclic aromatic system, is a prototypical intercalator. This type of binding is nonspecific with respect to the DNA sequence but prefers GC-rich regions. Larger molecules with appropriate polar appendages may make additional interactions with the grooves of the DNA duplex. When they make interactions in both the major and the minor grooves, like the anthracycline nogalamycin does, they are termed **threading intercalators** (Figure 9.3b, c). The aglycon of nogalamycin intercalates between the base pairs, the carbohydrate moiety (nogalose) sits in the minor groove, and the bicyclic acetal lies in the major groove.

Once bound, intercalators produce significant distortion of the DNA structure, resulting in the unwinding and elongation of the helix. This change in structure affects the stability of the duplex, prevents DNA's template function, and disrupts binding and recognition by certain DNA processing enzymes, such as topoisomerase, all leading to an inability to replicate the genetic material or causing extensive DNA damage. Tuning the activity of intercalators so they are only effective in cancerous or proliferating cells has been challenging, and this difficulty limits the usefulness of agents that act solely though this mechanism.

Several natural products that are approved for use in certain cancers work through an intercalation mechanism. The planar aromatic phenoxazone ring system of dactinomycin (also known as actinomycin D) intercalates between G–C base pairs, and the peptide side chains make H-bonds on the outside of the helix (Figure 9.4a). As seen in Figure 9.4b, an X-ray structure of the complex between dactinomycin and a segment

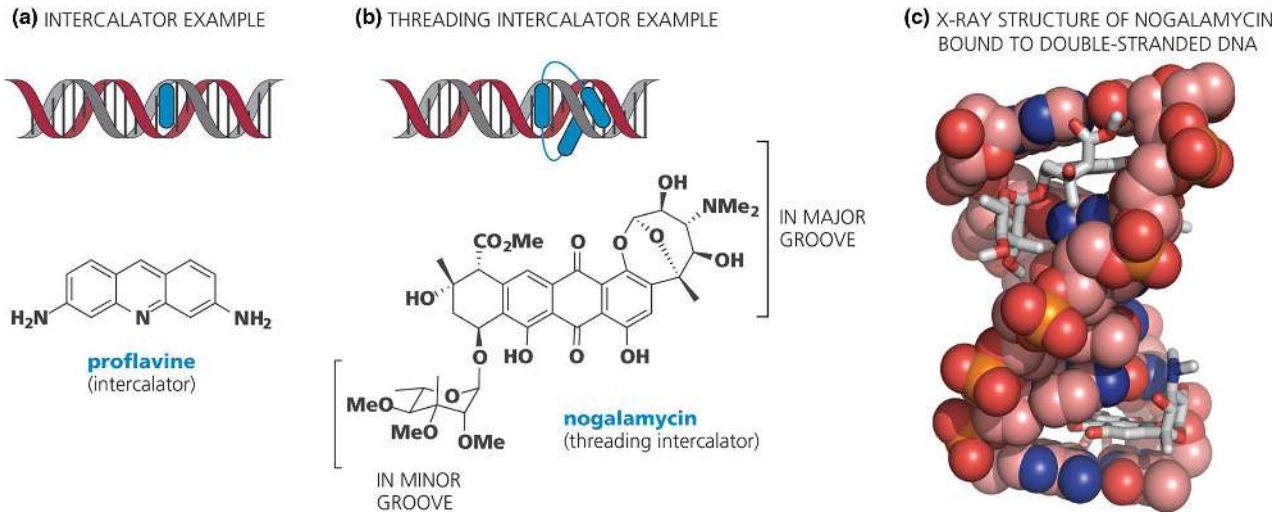
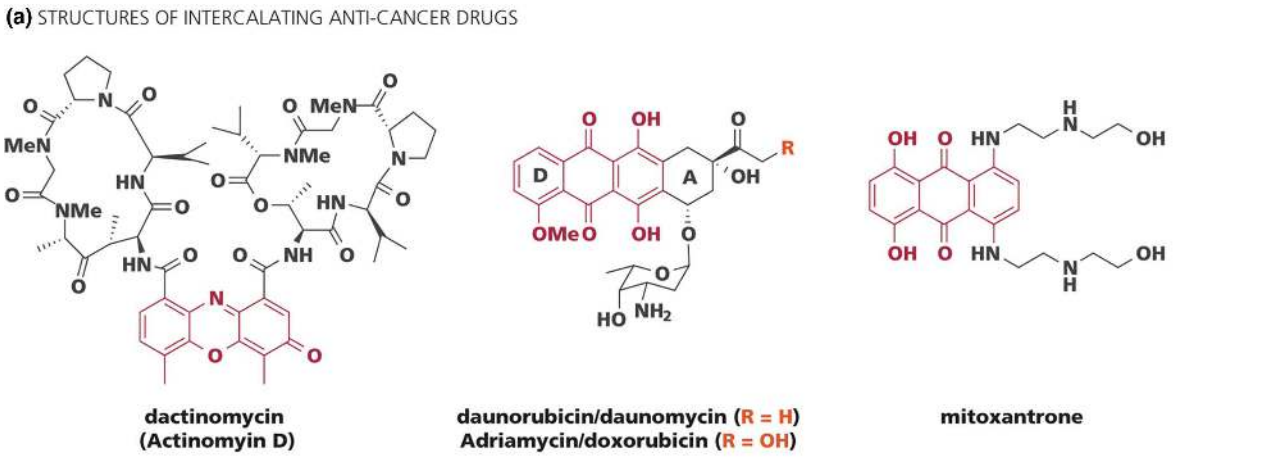


Figure 9.3 Schematic of DNA intercalators and threading intercalators, with examples of each. Red and gray represent two strands of the DNA helix; the blue bar represents a small-molecule binder. (a) Intercalators (blue) bind between the base pairs (black) of the DNA helix. Proflavine, with a planar aromatic structure, is a prototypical intercalator. (b) Threading intercalators (blue) bind between base pairs (black) but also bind to the minor and major grooves of the helix. Nogalamycin is an example of a threading intercalator, with the carbohydrate binding to the minor groove and the bicyclic acetal binding to the major groove. (c) X-ray structure of two molecules of nogalamycin (shown in stick format) bound to a short sequence of DNA helix. The planar moiety can be seen intercalated between base pairs at the bottom of the figure, with other portions of the molecule bound to helix grooves.

Source: (c) from Van Vranken, D, and Weiss, G. A. (2012). *Introduction to Bioorganic Chemistry and Chemical Biology*. Garland Science.



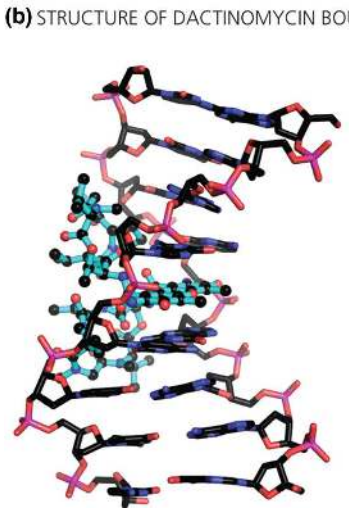


Figure 9.4 Examples of compounds that intercalate into double-stranded DNA. (a) Structures of dactinomycin/actinomycin D and of anthracyclines (daunorubicin/daunomycin, Adriamycin/doxorubicin, and mitoxantrone). Planar aromatic regions that intercalate are shown in red. (b) X-ray structure of dactinomycin (shown in cyan) bound in DNA double-strand helix. The aromatic core intercalates between base pairs, and peptide side chains bind to the minor groove, resulting in the distortion of the helix and unwinding.

Source: (b) from Zlatanova and van Holde, K. (2015). Molecular Biology: Structure and Dynamics of Genomes and Proteomes, Garland Science.

of a double-stranded DNA helix shows the drug intercalating into the middle of the helix, with the two peptide side chains lying in the minor groove. The helix is distorted and unwound at the lower end. The anthracyclines are a class of archetypical intercalators: planar polycyclic aromatic molecules with positive ionizable groups appended. Two members of this class, daunorubicin (daunomycin) and Adriamycin (doxorubicin), both natural products isolated from soil bacteria in the 1960s, have found wide use to treat both solid tumors and leukemias. Their planar aromatic tricyclic portion (shown in red; see **Figure 9.4a**) intercalates between base pairs, the amino sugar lies in the minor groove of the helix, and the D ring protrudes into the major groove.

These drugs have multiple mechanisms of action, including intercalation, which prevents further processing by enzymes such as topoisomerase. While the anthracyclines are widely used, they exhibit serious cardiotoxicity. The synthetic analog mitoxantrone (see **Figure 9.4a**), used to treat certain cancers and multiple sclerosis, has similar activity but reduced cardiotoxicity.

One current avenue of research is the identification of compounds that recognize and bind to specific uncommon conformations of DNA, called non-canonical structures, such as hairpins, cruciform, and multi-strand structures, such as triplex and G-quadruplex, among others. There is evidence that these unusual structures contribute to genomic stability and regulation of transcription and translation, and therefore, associations have been established between the presence of these structures and various disease states, such as cancer and neurodegeneration. As an example, G-quadruplex structures form in regions that contain a high percentage of guanine residues and are composed of multiple sets of four G-residues (tetrads) stabilized by the presence of cations that become stacked upon each other and are found at regions of DNA near telomeres and oncogene promoters, suggesting they play a regulatory role in cell proliferation and cancer. Targeting these unusual structures with small molecules is an active area of research, as any drugs developed have the potential to be less toxic than those that bind to the more common B-DNA motifs.

9.2.2 Drugs May Bind in the Minor Groove of DNA through Reversible Noncovalent Interactions

As noted earlier, many drugs that intercalate may also contain moieties that make additional contacts with the minor groove. Compounds that act primarily through such minor groove interactions are known as **DNA groove binders**. As a class, DNA groove binders sit in one of the two grooves of the DNA duplex. While many proteins interact with DNA through binding in the major groove, small molecules usually bind in the narrower yet deeper minor groove. Minor groove binders typically adopt a crescent shape that is complementary to the minor groove and contain flat, heterocyclic rings between flexible linkers. Specific interactions between the groove binder and DNA often occur at A-T-rich regions and include H-bonding with accessible lone pairs of atoms on base pairs in the floor of the minor groove, as well as hydrophobic interactions with thymidine methyl groups, and nonpolar H-atoms as seen in **Figure 9.5a**. Other points of contact include electrostatic and hydrophobic interactions with the sugar–phosphate backbone. Natural products, such as netropsin, distamycin, and trabectedin, are examples of minor groove binders (**Figure 9.5b**). The polyheterocyclic amides netropsin and distamycin adopt crescent-shaped conformations that fit into the minor groove of DNA, as seen in **Figure 9.5c**. When a second molecule of distamycin binds, the groove is significantly distorted.

Compared to intercalators, minor groove binders generally exhibit greater affinity and can exhibit some sequence specificity. The natural products netropsin and distamycin prefer AT-rich regions of DNA. Many analogs of these have been studied in attempts to develop sequence-specific drugs. In spite of these efforts, no approved drugs work solely through minor groove binding. However, binding to the minor groove of DNA is one aspect of the mechanism of a number of anti-cancer drugs. Trabectedin is one approved drug that binds in the minor groove with a preference for CGG sequences, but it also alkylates the exocyclic amine (N_2) of guanine. The potential for designing sequence-specific DNA binders offers opportunities to design drugs that treat genetic disorders that are defined by certain DNA sequences, as opposed to targeting all DNA.

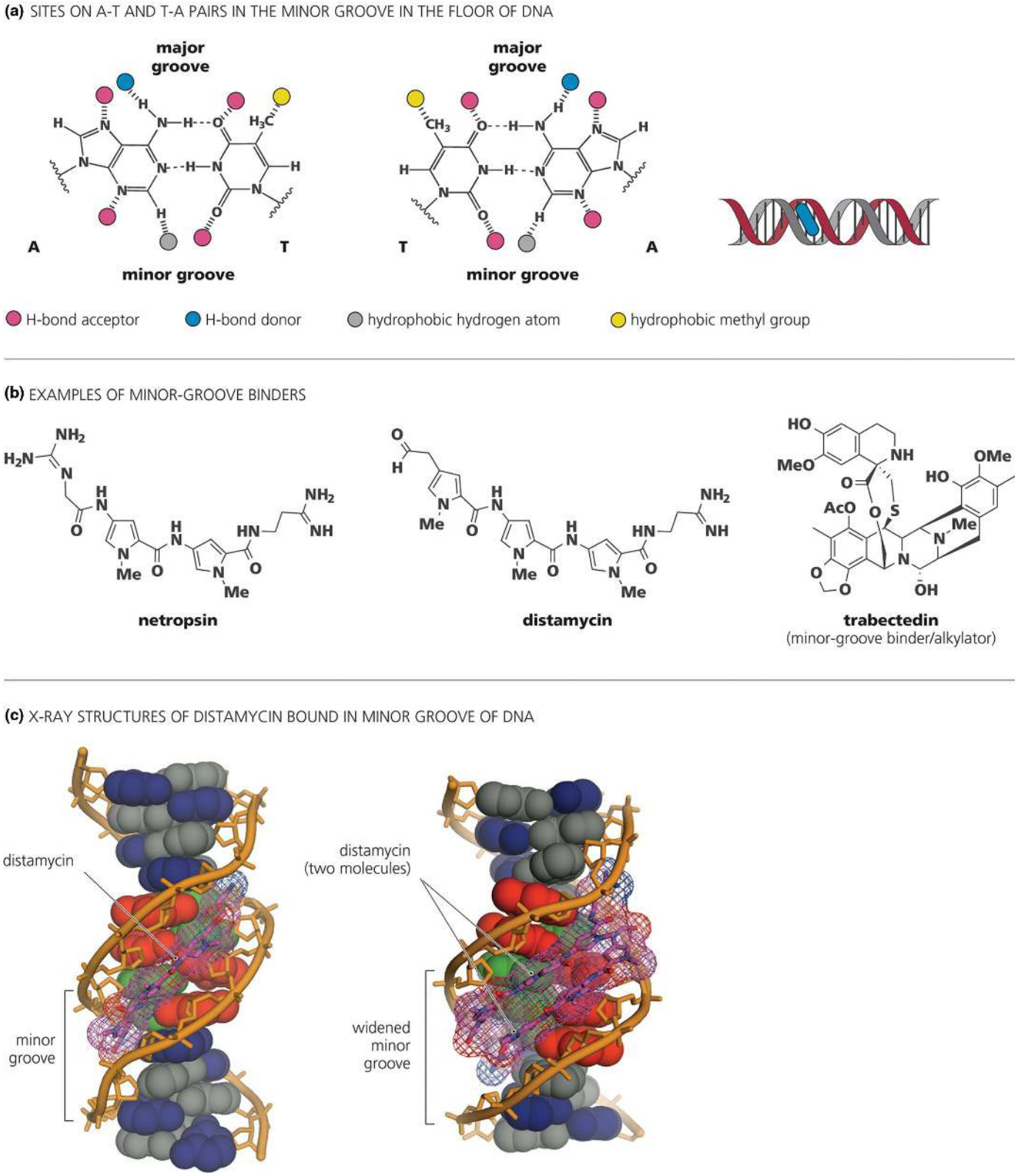


Figure 9.5 **Structures of minor groove binders and points of interaction.** (a) Sites on A–T and T–A base pairs available for interactions with minor groove binders. Colors indicate specific types of interactions: H-bond acceptor (pink), H-bond donor (blue), hydrophobic methyl group (yellow), and nonpolar H-atom (gray). The diagram on the *right* represents a compound (blue) seated in the minor groove of DNA. (b) Structures of minor groove binders netropsin, distamycin, and trabectedin. (c) X-ray structures showing binding of one or two molecules of distamycin (hashed space filling purple) in AT-rich regions (red and green spheres). The binding of a second molecule of distamycin results in widening of the minor groove.

Source: (a, c) From Kuriyan, J.; Konforti, B.; and Wemmer, D. (2012). *The Molecules of Life: Physical and Chemical Principles*. Garland Science.

9.2.3 Drugs That Interact with DNA by Irreversible Mechanisms Are Anti-Cancer Agents

In addition to molecules that bind to DNA through reversible noncovalent interactions, a number of drugs work by making covalent bonds to DNA. These are known as **DNA alkylating agents**. Despite their significant toxicities, compounds such as these continue to be used for anti-cancer therapy. In general, they contain, or generate upon bioactivation, a highly reactive electrophile that reacts with nucleophilic sites on DNA. Modification of N7 of guanine is common, as is alkylation at N3 and N7 of adenine and O6 of guanine (Figure 9.6).

With **bis-alkylating agents**, a second reaction on a neighboring base also occurs to cross-link the bases. The linkages can occur between two bases on the same strand (intrastrand) or between two bases on opposite strands (interstrand). Once chemically modified, either by mono- or bis-alkylation, the DNA is incompetent as a template. At that point, it is often recognized as damaged and degraded, or it may not be accepted as a substrate by the enzymes required for replication. Regardless of the mechanism, cell division and proliferation are arrested. Alkylating agents are known to alkylate other biomolecules, such as proteins, and therefore, they are usually quite toxic. Classes of DNA alkylators can be characterized by their electrophiles, and representative examples are shown in Table 9.2.

Nitrogen mustards, triazines, and nitrosoureas, through a complex series of biotransformations, generate highly reactive intermediates, such as aziridinium ions from nitrogen mustards and diazonium ions from triazines and nitrosoureas. The nitrogen mustards, of which chlorambucil is an example, were derived from mustard gas, first used in World War I as a chemical warfare agent (Figure 9.7). One result of gas (bis[2-chloroethyl] sulfide) exposure was a lowering of white blood cell count. This observation led scientists to pursue analogs that might be useful for treating leukemia,

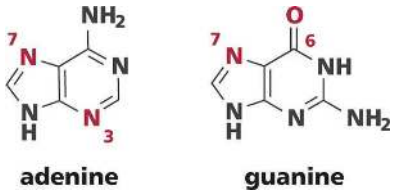
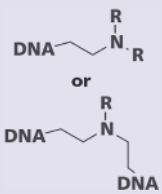
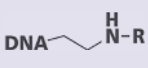
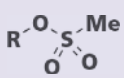


Figure 9.6 Nucleophilic sites on adenine and guanine. Nucleophilic sites (N3 and N7 in adenine; N7 and O6 in guanine) are marked in red.

Table 9.2 Classes of DNA Alkylators, the Reactive Electrophile, Reaction Product, and Select Examples of Drugs

Class	Electrophile	Drug Example	Reaction Product
Nitrogen mustards	Aziridinium ion 	Mechlorethamine, chlorambucil, cyclophosphamide	
Triazines	Diazonium ion 	Dacarbazine	**DNA-R**
Nitrosoureas	Diazonium ion 	Carmustine	**DNA-R**
Aziridines	Aziridine 	Thiotepa	
Sulfonates	Methylsulfonate 	Busulfan	**DNA-R**
Platins	Platinum	Cisplatin	**DNA-Pt-DNA**

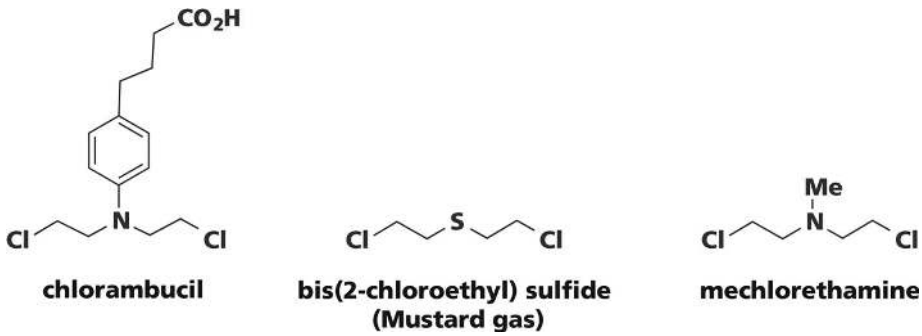


Figure 9.7 Structures of mustards. Chlorambucil is an example of a modern mustard used to treat certain cancers, including leukemias. Bis(2-chloroethyl) sulfide, also known as mustard gas, was used during World War I as nerve gas. In the 1940s, the replacement of the sulfur atom with an amine led to a less-toxic analog, mechlorethamine, used to treat certain cancers.

based on the rationale that less-toxic analogs might be effective in lowering abnormal white blood cell counts. Replacing the sulfur atom with nitrogen led to less-toxic analogs, with mechlorethamine being used clinically to treat Hodgkin's lymphoma in the early 1940s.

The mechanism of action of cyclophosphamide, a nitrogen mustard prodrug that bis-alkylates DNA, is illustrated in Figure 9.8. Bioactivation by cytochrome P450 oxidation and subsequent ring opening affords acrolein and phosphoramide mustard (Figure 9.8a), which then generates the highly reactive aziridinium ion (Figure 9.8b). The aziridinium ion reacts with N7 of a guanine residue in DNA to yield an alkylated derivative. This aziridinium formation process is repeated, and alkylation of a nearby guanine residue generates a cross-linked DNA adduct (Figure 9.8c).

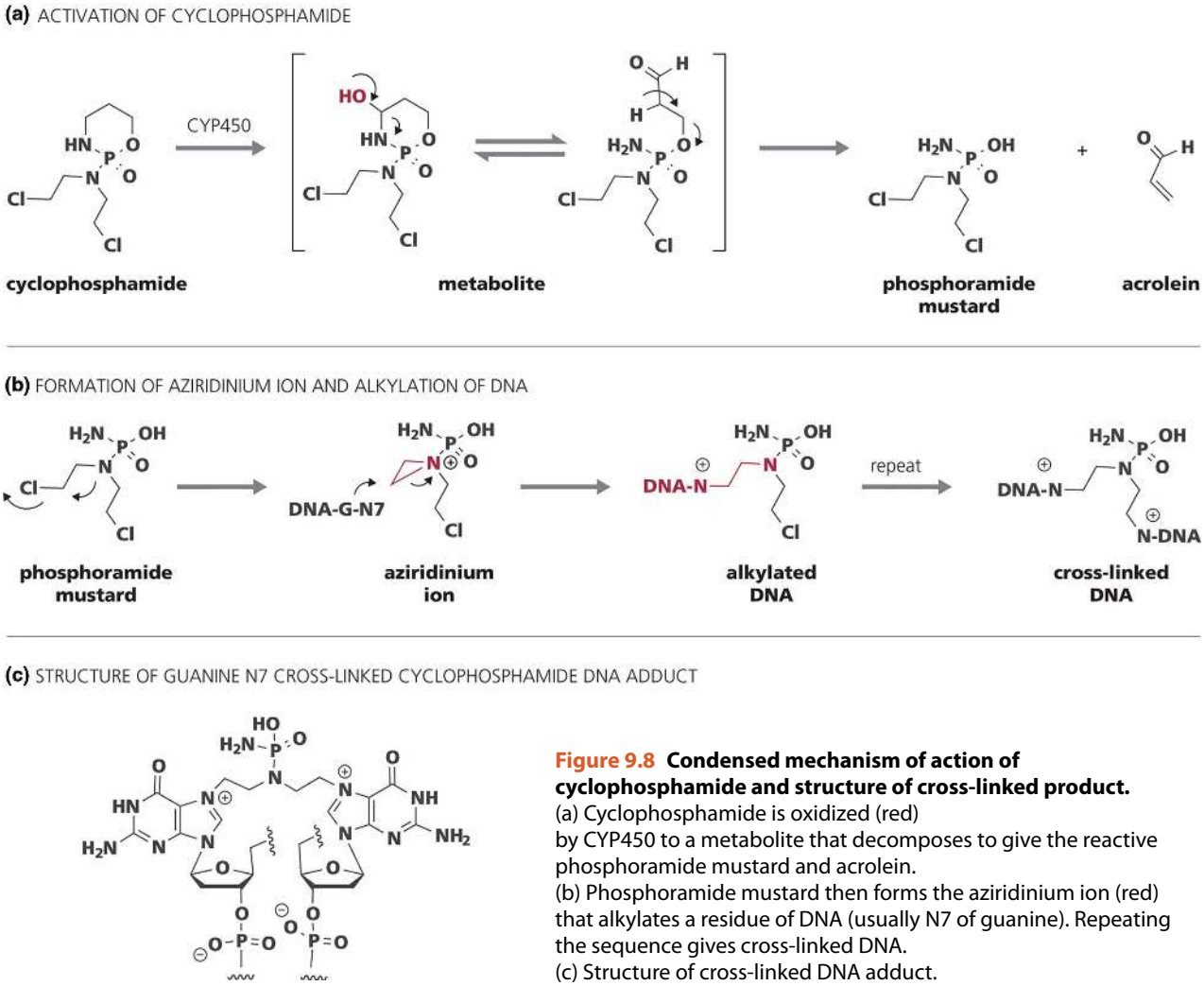


Figure 9.8 Condensed mechanism of action of cyclophosphamide and structure of cross-linked product.
(a) Cyclophosphamide is oxidized (red) by CYP450 to a metabolite that decomposes to give the reactive phosphoramide mustard and acrolein.
(b) Phosphoramide mustard then forms the aziridinium ion (red) that alkylates a residue of DNA (usually N7 of guanine). Repeating the sequence gives cross-linked DNA.
(c) Structure of cross-linked DNA adduct.

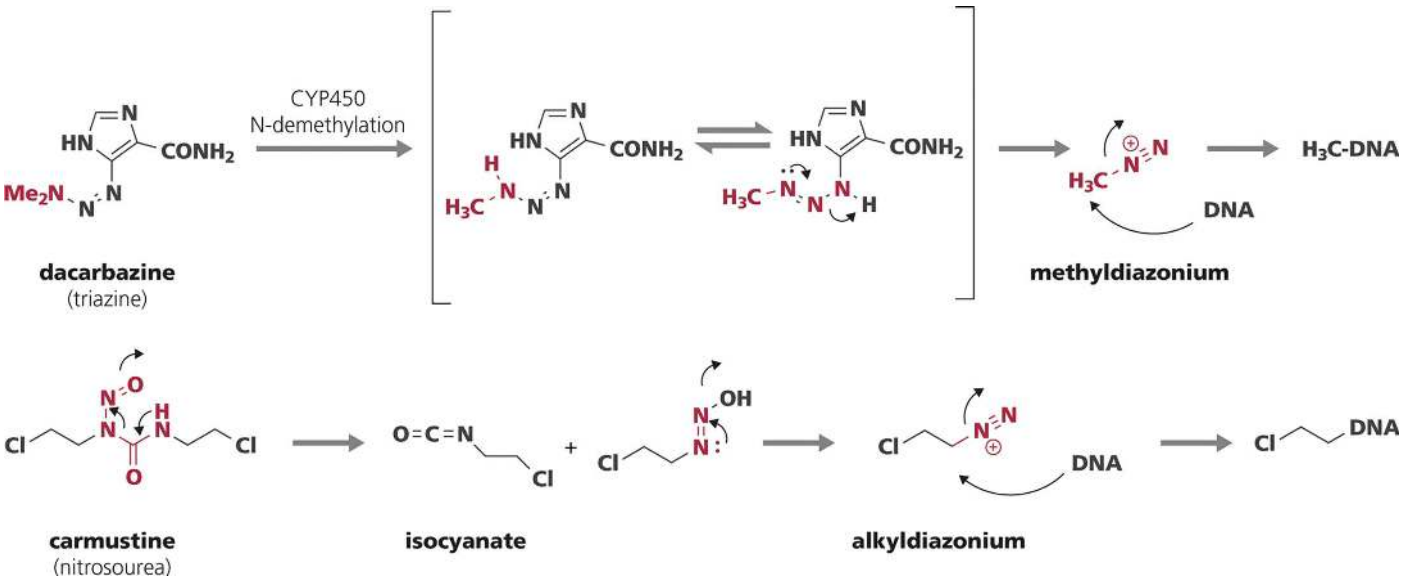
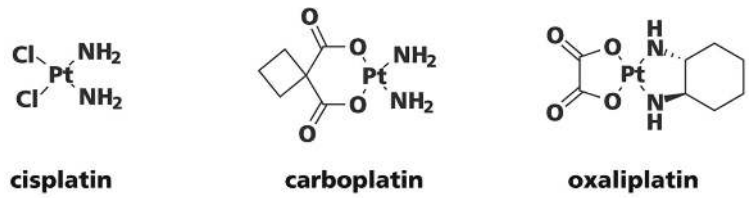


Figure 9.9 Alkylation of DNA by dacarbazine and carmustine (sites involved in reaction in red). *Top*, dacarbazine is N-demethylated *in vivo* to give a product that decomposes to a reactive methyldiazonium ion, which methylates DNA. *Bottom*, carmustine, a nitrosourea (red), decomposes to an isocyanate and an alkyldiazonium ion; the latter alkylates DNA.

Similarly, dacarbazine (a triazine) and carmustine (a chloroethyl-containing nitrosourea) generate reactive intermediates that alkylate DNA, as shown in Figure 9.9. Dacarbazine is first activated by metabolic N-demethylation. The resulting intermediate generates one equivalent of highly reactive methyldiazonium, which reacts with DNA to form a methylated product. The nitrosoureas, exemplified by carmustine, are activated by decomposition to an isocyanate and an alkyldiazonium ion, the latter of which is the alkylating agent.

In contrast to the previous examples, aziridines and sulfonates require no activation for reaction. The requirement for bioactivation by some agents is likely a mechanism by which the drug is protected from reacting with other nucleophiles in the body (e.g., proteins, other biomolecules) and eventually unmasked at the site of action.

While they are technically metalating agents and not alkylating agents, platinum-containing drugs (known as platins) work through a related mechanism. The effects of cisplatin (Figure 9.10) were discovered serendipitously by Barnett Rosenberg





X-RAY STRUCTURE OF CISPLATIN BOUND TO DNA FRAGMENT

Figure 9.10 Platinum-containing drugs. *Top*, structures of cisplatin, carboplatin, and oxaliplatin. *Bottom*, X-ray structure of cisplatin bound to DNA fragment, showing cross-linking. Pt atom = large while ball.

Source: Bottom, from Fuertes, M. A.; Alonso, C.; and Perez, J. M. (2003). *Chem Rev* 103:645–662. With permission from the American Chemical Society.

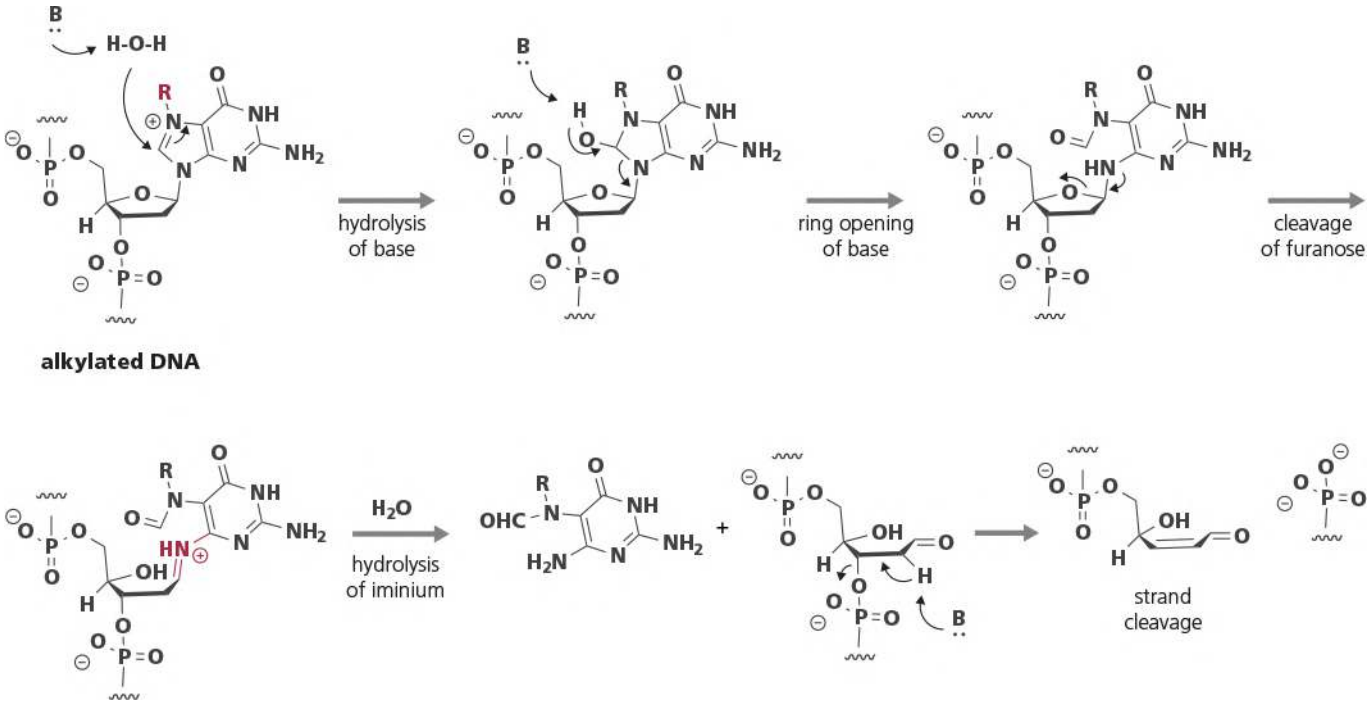


Figure 9.11 **Abbreviated depurination mechanism.** The hydrolysis of the iminium ion of alkylated DNA (alkyl group R is shown in red) leads to an unstable intermediate and ring opening of the purine base. This is followed by the cleavage of the C–O bond of the furanose and the generation of a reactive iminium ion (shown in red). Hydrolysis of the iminium releases the remnants of the guanine base and affords an aldehyde. Base-catalyzed elimination of the 3′ phosphate results in strand cleavage.

and Loretta Van Camp in the 1960s as a result of experiments designed to observe the effect of an electrical current on bacterial cell division. The bacteria failed to grow, and it was eventually determined this effect was due not to electricity but to the supposedly inert platinum electrode which had reacted with an ammonium chloride solution to form $Pt(NH_2)_2Cl_2$. Several years later, the compound, now called cisplatin, was shown to shrink tumors in mice; the drug was approved for human use in 1978. Its mechanism of action involves hydrolysis of the chlorine atoms and then binding to DNA and the displacement of two Pt ligands by N7 atoms of adjacent guanines. This results in the formation of intrastrand cross-links, helix distortion, and ultimately, cell death. Carboplatin (1989) and oxaliplatin (2002) are considered to be second-generation agents, as they exhibit reduced kidney toxicity compared to cisplatin. These agents form both intrastrand and interstrand complexes with purines on DNA strands, with the majority (about 65%) being intrastrand linkages between two guanines (see **Figure 9.10**).

Common among many of these DNA-modifying agents is that, eventually, strand cleavage results. Once a DNA nucleobase has been alkylated or modified, the resulting adduct is labile and can undergo depurination either by simple hydrolysis or by a mechanism such as the one outlined in **Figure 9.11**. The addition of water to the reactive iminium ion is followed by ring opening of the heterocyclic base. Cleavage of the furanose ring, followed by hydrolysis of the resulting iminium ion, releases the ring-opened pyrimidine and generates an aldehyde. The loss of the DNA base results in an apurinic/apyrmidinic (AP site), which can undergo base-catalyzed elimination of the 3′-phosphate, which results in the final strand cleavage.

9.2.4 Compounds That Cleave DNA after Binding and Generating Free Radicals Are Powerful Drugs Used to Treat Cancer

As seen in the previous section, drugs that work by alkylating DNA destabilize the duplex and ultimately result in DNA damage. Another mechanism to irreversibly inflict damage onto DNA involves molecules that generate free radicals. DNA **strand**

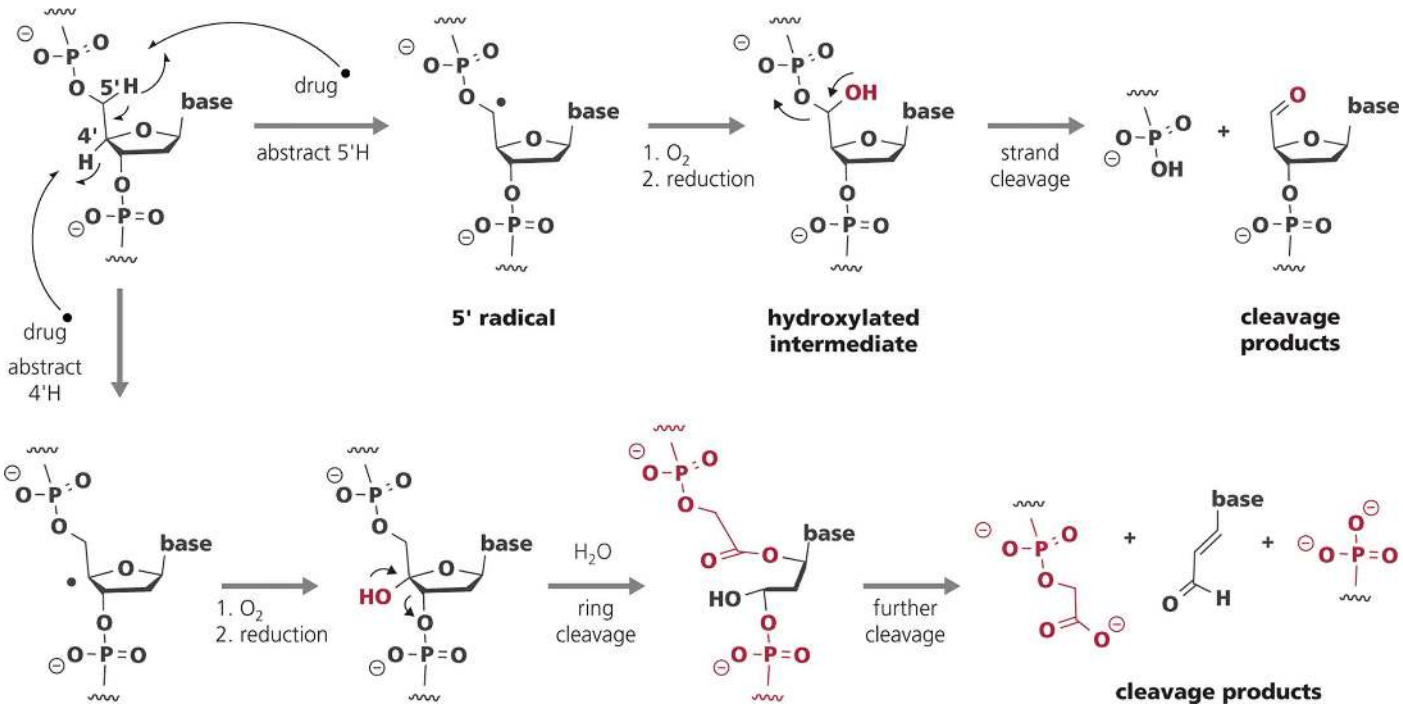


Figure 9.12 Two mechanisms of DNA strand cleavage by radicals. *Top*, drug-derived radical abstracts the C5′ hydrogen. This secondary radical reacts with molecular oxygen and, after reduction, gives an unstable hydroxylated intermediate (hydroxyl group is shown in red). Elimination of the phosphate generates the cleaved 3′ phosphate and an aldehyde, thereby cleaving the strand. *Bottom*, drug radical abstracts the C4′ hydrogen. This tertiary radical reacts with molecular oxygen and, after reduction, generates the hydroxylated intermediate (hydroxyl group is shown in red). Hydrolysis leads to ring cleavage to generate the ring-opened product. Further cleavage generates a 3′ phosphorylated acetate (red), a nucleobase derivative, and the 5′-phosphate (red).

breakers and **nicking agents** are molecules that result in bonds within the DNA duplex being broken. These molecules bind to DNA, often in the minor groove, and generate a free radical. The radical then abstracts a hydrogen atom from the sugar–phosphate backbone, usually at either C4′ or C5′, leading to a carbon radical, which can react with molecular oxygen. These reactive intermediates react further, ultimately causing the DNA strand to break. While multiple mechanisms have been observed, two examples detailing C4′ and C5′ hydrogen abstraction are shown in Figure 9.12, each resulting in DNA damage.

These lesions may nick the DNA, cutting one of the strands, or subsequent radical propagation may break both strands. Either situation leads to incomplete duplication, transcription, and translation. Nicking and cleaving agents are usually highly potent, and some exhibit a degree of sequence selectivity because they bind tightly to the DNA duplex. Many are intercalators or minor groove binders as well as damaging agents.

One DNA strand breaker that is used clinically is bleomycin (BLM), a glycopeptide antibiotic isolated from *Streptomyces verticillus* in the 1960s as its copper complex. BLM operates by generating a radical intermediate after complexation to iron and oxidation by molecular oxygen. The bis(thiazole) moiety of BLM intercalates into DNA, and five nitrogen atoms (shown in red in Figure 9.13) as well as molecular oxygen form a complex with iron *in vivo*, generating BLM–Fe(II)–O_2. This activated BLM species generates radicals near the DNA backbone and abstracts either a 4′ or a 5′ hydrogen.

The enediyne calicheamicin (Figure 9.14a), an unusual natural product isolated from soil bacteria found in Texas, is another example of a DNA strand breaker. Its mechanism of action highlights the complexity that is often involved in DNA strand breakers and nicking agents (Figure 9.14b). After binding to the minor groove of DNA with some specificity for 5′-TCCT-3′ and 5′-TTTT-3′ sequences, nucleophiles in the cell, such as glutathione, add to the trisulfide, generating a thiolate, which undergoes an intramolecular Michael addition to give the highly strained bridged analog. This structure brings the enediyne in close enough proximity to undergo a

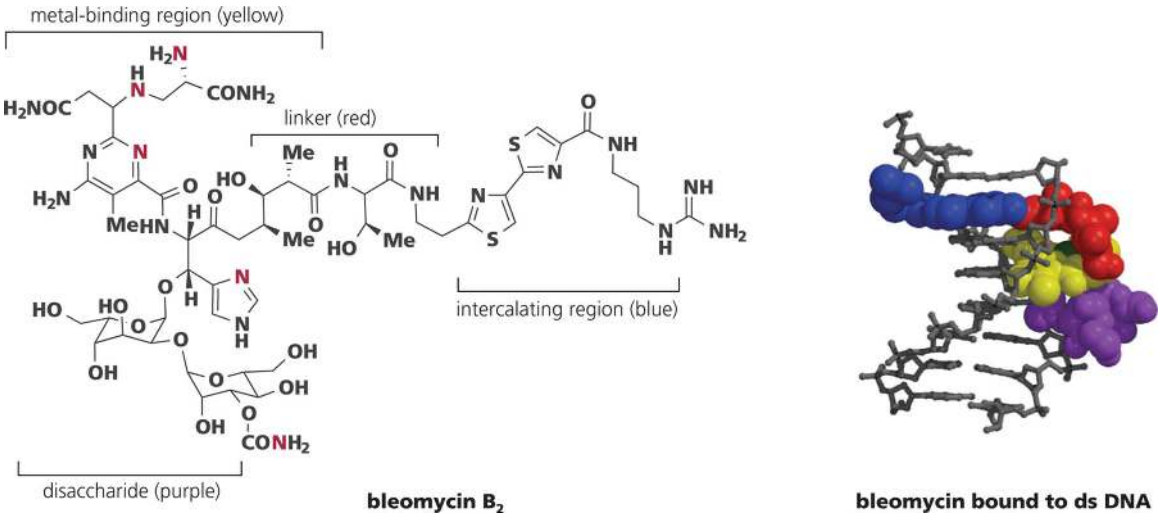


Figure 9.13 **Structure of bleomycin B_2.** *Left*, structure of bleomycin with key interacting regions that correspond to the X-ray structure labelled, and five N atoms that chelate to the metal shown in red. *Right*, X-ray structure of bleomycin bound to DNA duplex (shown in gray ball-and-stick representation). The bis(thiazole) region (blue) intercalates between the base pairs, and a linker (red) orients the metal-binding region (yellow) in the minor groove; the disaccharide (purple) forms additional hydrogen bonding interactions with the backbone.

Source: Adapted from Goodwin, K. D.; Lewis, M. A.; Long, E. C.; and Georgiadis, M. M. (2008). *Proc Natl Acad Sci USA* 105:5,052–5,056. With permission from the National Academy of Sciences.

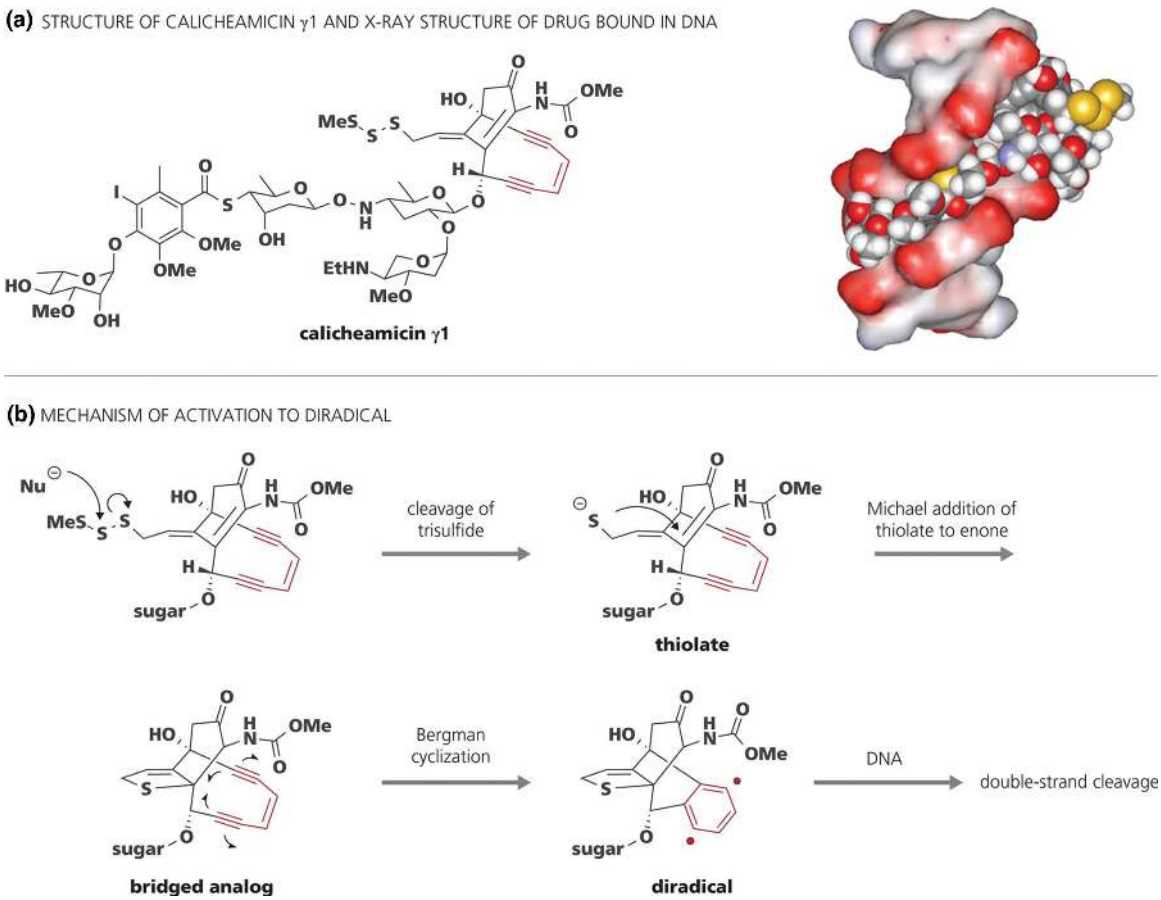


Figure 9.14 **Structure and mechanism of action of calicheamicin γ1.** (a) Structure of calicheamicin γ1, with the enediyne moiety shown in red, and X-ray structure of drug (space-filling atoms: red, oxygen; blue, nitrogen: white, carbon: yellow, sulfur) bound in DNA (surface representation; oxygen, red). Carbohydrate binds in the minor groove, with the trisulfide accessible to nucleophiles. (b) Mechanism of activation of calicheamicin γ1 to generate a diradical. The addition of a nucleophile such as glutathione to the trisulfide generates a thiolate anion, which then undergoes intramolecular Michael addition to the enone. The resultant bridged analog is appropriately oriented to undergo a Bergman cyclization to generate a diradical, which abstracts H-atoms from 4′ and 5′ sites on DNA furanoses, resulting in DNA double-strand cleavage.

Source: From Van Vranken, D, and Weiss, D.A. (2015). Introduction to Bioorganic Chemistry and Chemical Biology, Garland Science.

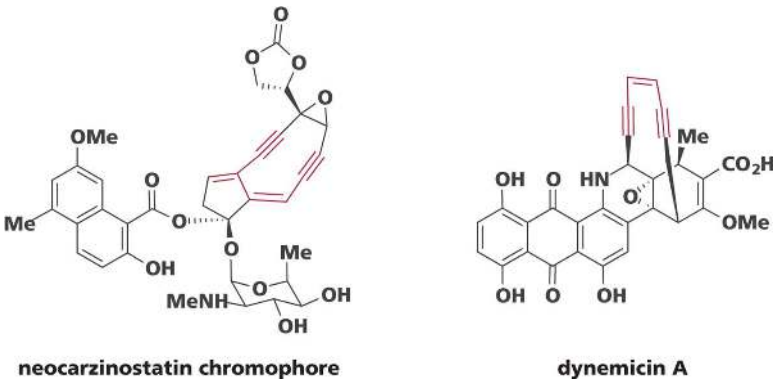


Figure 9.15 Structures of enediyne strand breakers. Neocarzinostatin and dynemicin are other examples of natural product enediynes; enediyne-containing moiety is shown in red.

Bergman cyclization, generating a 1,4-benzene diradical. This reactive species abstracts H-atoms from C5′-positions of pyrimidine nucleotides and the C4′-position of a nucleotide on the complementary strand, causing double-strand cleavage.

Other enediyne natural products, such as neocarzinostatin and dynemicin A, similarly act by the generation of a diradical via Bergman-type cyclization (**Figure 9.15**). Unfortunately, their significant toxicity precludes their clinical use as a single agent. Mylotarg, a derivative of calicheamicin conjugated to an antibody directed to a protein on the surface of certain leukemia cells, was used between 2000 and 2010 to treat acute myelogenous leukemia (AML), withdrawn, and was recently reapproved. The area of conjugating antibodies to highly toxic agents continues to evolve, with a number of these entities currently undergoing clinical trials. They will be further described in Chapter 10.

Drugs that target DNA work through various and sometimes complex mechanisms. Intercalation and minor groove binding are important components to many of them. In addition, a number either make a covalent bond to DNA or cause strand cleavage. Designing selective inhibitors that will target either a specific DNA sequence or a specific DNA structure remains a challenge, although insights into DNA structure may lead to improved specificity in these types of drugs. The majority of drugs that target DNA are anti-cancer agents, with a few being used to treat serious immuno-inflammatory disorders, such as multiple sclerosis, but they are usually accompanied by significant side effects.

9.3 RNA AS A DRUG TARGET

In contrast to DNA-targeted agents, there are fewer approved small drugs that bind directly to RNA. This is perhaps surprising, since RNA is involved in numerous biological processes, such as protein synthesis and transcriptional regulation, that are likely to be modulated in the disease state, while the function of DNA is more limited (although with the exception of tRNA and rRNA, RNA molecules are transient, unlike DNA). Furthermore, the potential for small molecules to bind directly to RNA and distort its structure, or to disrupt RNA–protein binding events, offers numerous opportunities for drug design. However, to date, success in this arena has been limited to targeting the ribosome with traditional small molecules derived from natural products, sequence-specific nucleotide drugs called antisense oligonucleotides, and drugs that work via modulating RNA splicing.

9.3.1 Compounds Targeting Bacterial Ribosomal RNA Inhibit Protein Synthesis and Are Useful Anti-Infective Agents

Many antibacterial drugs work by interfering with bacterial protein synthesis. While they have distinct binding sites and diverse structures, their common mechanism involves binding to rRNA. Ribosomes are multicomponent complexes that contain both protein and RNA. In bacteria, the 70S ribosome is composed of two parts, the 50S and 30S subunits. In the process of protein synthesis, which takes place on the

ribosome, tRNA carries the three-base code for an amino acid into the A site (acceptance site) of the 30S subunit. One class of antibacterial drugs that binds to rRNA is the tetracyclines (**Figure 9.16**). These drugs interact with the 16S rRNA of the 30S subunit of the bacterial ribosome and prevent tRNA from binding at the A site. Several binding sites of this class have been identified, but the location where these small molecules bind with the most frequently is located in a hydrophobic pocket near the A site. This pocket is formed by several helices of the rRNA, including helix 34 (H34), H31, H44, and H18. As seen in **Figure 9.16**, the 1,3-ketone-enol motif on the hydroxyl-rich side of tetracycline participates in numerous interactions with RNA, mostly binding to the phosphates in the RNA backbone, including some mediated through coordination with a Mg^{2+} ion. It has long been known that a magnesium ion is required for tetracycline binding, and SAR studies have shown that little variation is tolerated on the hydroxyl-rich side of the molecule.

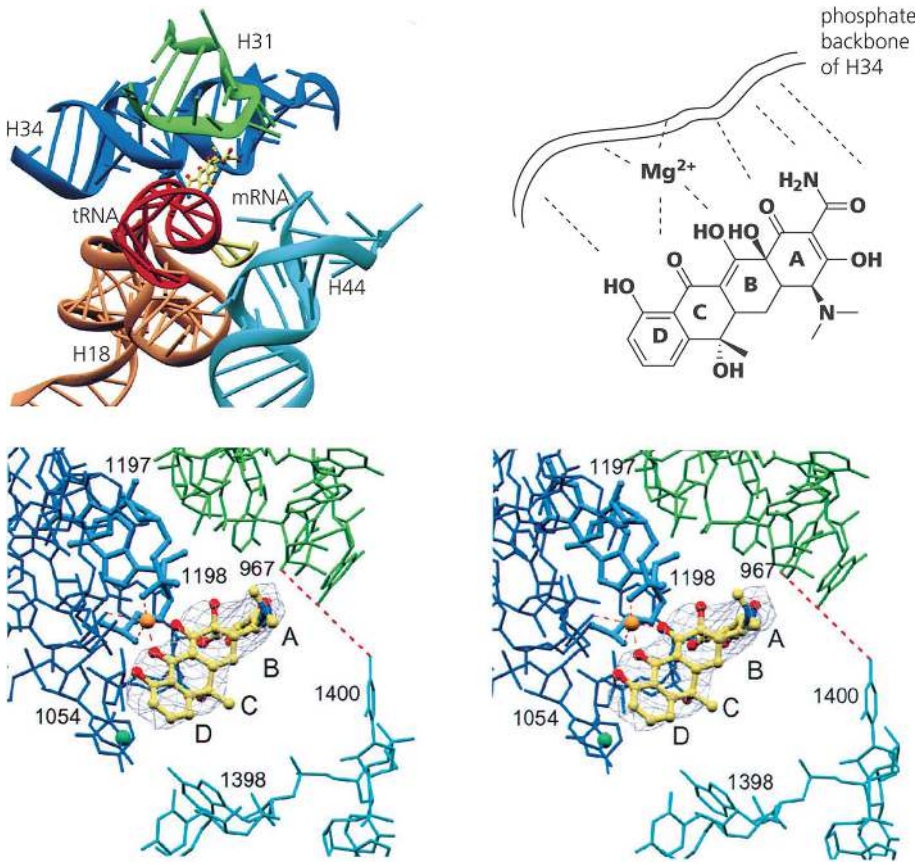


Figure 9.16 Structure of tetracycline bound to rRNA. *Top left*, X-ray structure of tetracycline (yellow, with oxygen atoms in red and nitrogen atoms in blue) in the A-site binding pocket, near tRNA (red). Major binding is to helix 34 (H34) (dark blue), with H31 (green), H44 (light blue), and H18 (orange) forming the pocket. *Top right*, structure of tetracycline in the same orientation shown in the top left, with interactions to phosphate backbone of H34 indicated by dotted lines. Mg^{2+} is required for binding and coordinates to both tetracycline (A, B, C, and D rings labeled) and the rRNA. *Bottom*, stereodiagram showing close-up of the binding site, with coordination to Mg^{2+} (orange circle); colors same as top figures.

Source: From Brodersen, D. E.; Clemons, W. M. Jr.; Carter, AP.; et al. (2000) *Cell* 103:1,143–1,154. With permission from Elsevier.

Other classes of antibacterial protein synthesis inhibitors that bind to the ribosome include aminoglycosides and erythromycins. These are discussed in detail in Chapter 12.

9.3.2 Antisense Therapies Targeting RNA Are Effective in Treating Genetic Diseases and Infections

Antisense therapy is used to treat disorders resulting from a disease-causing gene or an essential gene from an infectious agent. It works by either blocking the synthesis of proteins or, alternatively, by promoting the production of a protein that is deficient or non-functional. As such, knowledge of the sequence of the mutant gene or a gene essential for an infectious agent's replication is required. On the basis of that sequence, an oligonucleotide or modified analog with a complementary sequence (thus the term *antisense*) is synthesized that will bind to the mRNA produced by that gene. In some examples where the goal is to turn off protein production, since the

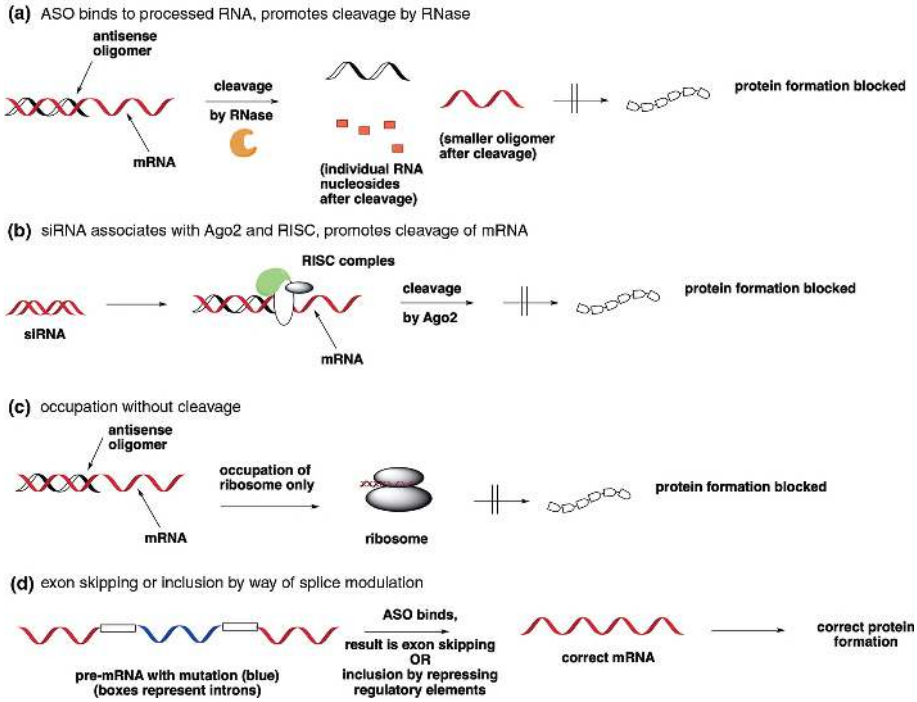


Figure 9.17 Antisense mechanisms. (a) Cleavage-dependent mechanism where the antisense oligomer (shown in black) binds to a specific section of mRNA. RNase H cleaves the segment of RNA bound to the antisense RNA into individual nucleotides, thereby releasing the remaining fragment of RNA and the antisense oligomer. This stops further protein production, preventing translation to protein. (b) Another cleavage-dependent mechanism where siRNA associates with the RISC complex and target mRNA, resulting in cleavage by Ago2. (c) Occupation-only mechanism, where the antisense oligomer physically blocks entry to the ribosome and translation. (d) Exon skipping or inclusion involves splicing modulation of pre-mRNA, where introns are removed. Binding of ASO results in exon skipping of a mutant or may involve repression of regulatory elements, with both resulting in the production of a functional protein.

mRNA is no longer capable of translating the gene into the protein, the specific protein the mRNA codes for is not synthesized, thereby effectively inhibiting the protein. Alternatively, depending on the specific gene mutation, antisense molecules can promote the synthesis of proteins by splicing modulation using exon skipping or inclusion. In this manner, a mutant protein may be "corrected," or regulatory mechanisms may be repressed, allowing for the production of a specific protein.

This approach allows highly precise targeting, since the antisense agents (termed antisense oligonucleotides, or ASOs) are typically in the range of 15–20 base pairs and therefore recognize very specific sequences of mRNA and also prevents nonspecific binding to other mRNAs. However, this strategy also has some challenges, such as the efficient delivery of the agent to the site of action. While antisense drugs are not quite small molecules, many of the traditional medicinal chemistry principles, such as design of modified analogs for selectivity and physical property optimization, have been applied to the discovery and development of these agents.

All ASOs bind to mRNA, and their activity can be divided into two general mechanisms: cleavage-dependent or occupation-only (**Figure 9.17**). About half of the ASOs approved and in clinical development work through the **cleavage-dependent mechanism**, with most relying on the action of the enzyme **RNase H** (**Figure 9.17a**). This endonuclease normally mediates the hydrolysis of the 3′-O-phosphate bond of the RNA strand in RNA–DNA heteroduplexes, releasing intact DNA. In the case of antisense molecules, RNase H recognizes the RNA–antisense duplex, hydrolyzes the portion of the RNA sense strand bound to the antisense molecule, and releases the intact antisense oligomer and the smaller, remaining RNA strand. Drugs acting by this mechanism include mipomersen and fomiversen. Another mechanism involving cleavage relies on small interfering RNAs (siRNAs). These are naturally occurring 20- to 25-nucleotide segments of double-stranded RNA whose typical function is to interfere with gene expression.

Synthetic siRNAs, such as the drug patisiran, used to treat hereditary TTR amyloidosis, are first processed to cleave overhanging nucleotides at the end of the RNA, which allows it to bind to the RNA-induced silencing complex (RISC). Once bound to the RISC, the duplex is separated into two single strands, one of which is released, while the other remains bound. The bound single strand derived from the synthetic siRNA can now bind to a complementary mRNA sequence, in the case of patisiran, to

the mRNA that produces TTR, where the RISC can act, eventually leading to the degradation of the mRNA by the enzyme Ago2 (Figure 9.17b).

Antisense oligonucleotides that act by an **occupancy-only mechanism** work by blocking translation (Figure 9.17c). Although none are yet approved, they are designed to bind to (or next to) the translation initiation region of an mRNA. In the process, they block or change movement of the ribosome down the transcript. By either cleavage or occupation-only mechanisms, the overall result is interference with the formation of a specific protein. Alternatively, a mutant protein, or a protein whose production has been repressed, can be corrected by **exon skipping** or inclusion. This involves splicing modulation of pre-mRNA, where introns are removed (Figure 9.17d). Binding of ASO results in exon skipping of a mutated region or may involve repression of regulatory elements, with both resulting in the production of a functional protein. Examples of drugs acting through this mechanism include etiplersen and nusinersen.

In order to design antisense drugs, several issues that plague natural oligonucleotides need to be addressed. These include poor bioavailability, instability due to degradation by nucleases (which cleave the phosphodiester bonds), and rapid clearance. Because the oligomers are charged, their cell permeability is poor. Sequence specificity requires that a typical modified oligonucleotide contain between 8 and 50 nucleotides, with most averaging approximately 20; therefore, the molecular weights are quite high (~7,000). At the same time, the oligomers must display enough of the same structural features as natural oligonucleotides in order to be recognized by the complementary RNA strand as well as the relevant enzymes, like RNase H or Ago2. To address all these requirements, the design of antisense therapeutics has included modification of the phosphate backbone and the ribose units, as well as the nucleobases.

The first approved antisense drugs, and the dozens of ASOs undergoing clinical trials, incorporate backbone modifications, with phosphorothioate (PS) being one of the earliest and most widely used replacements for the cleavable phosphodiester linkage. These analogs replace the nonbridging oxygen of the phosphate with sulfur, resulting in compounds with increased stability to nucleases that also elicit RNase H cleavage. Others include N3′ → P5′ thiophosphoramidates, phosphorodiamidates, and peptide nucleic acids (Figure 9.18a).

The most widely studied base analogs are the 5-methylpyrimidines (such as thymine) and the unnatural 5-propynylpyrimidines (Figure 9.18b), which exhibit more efficient base stacking, resulting in improved stability of the RNA–oligonucleotide duplex. While duplexes that incorporate the 5-propynyl analogs are even more stable than 5-methyl ones, they are toxic *in vivo*. Examples of modified ribose analogs include the morpholino replacement found in many phosphorodiamidates (Figure 9.18a) and methoxyethyl (MOE)– and fluoro-substituted ribose oligonucleotides (Figure 9.18b), as well as the bridged, locked nucleic acids (LNA). These latter two modifications exhibit increased stability to nucleases and increased affinity to RNA.

The first ASO drug, fomivirsen, was approved in 1998 to treat cytomegalovirus infections, although its marketing has been discontinued in the United States. Fomivirsen is a phosphorothioate oligonucleotide containing 21 bases, with no other modifications of the base or backbone. Subsequently, more complex second-generation antisense drugs have become available, such as mipomersen, which is used to treat familial hypercholesterolemia. Mipomersen is described in more detail in Case Study 2 in this chapter. Eteplirsen, approved by the US FDA with some controversy in 2016, targets a mutation found in about 15% of patients with muscular dystrophy. It works by exon skipping and contains a phosphordiamidate linkage and replaces the ribose sugar with a morpholino group (PMO) (see Figure 9.18a). Another example is nusinersen, used to treat spinal muscular atrophy (SMA) in infants. As of 2025, there are 14 FDA-approved antisense oligonucleotides, used to treat diseases such as Duchenne muscular dystrophy and transthyretin amyloidosis. Other antisense agents are undergoing clinical trials, including those to treat diseases ranging from Ebola and Marburg viral infections to motor disorders and neurodegenerative diseases, such as ALS. It is likely that future antisense therapeutics will have to be similarly optimized with a variety of modifications at multiple sites of the nucleotide.

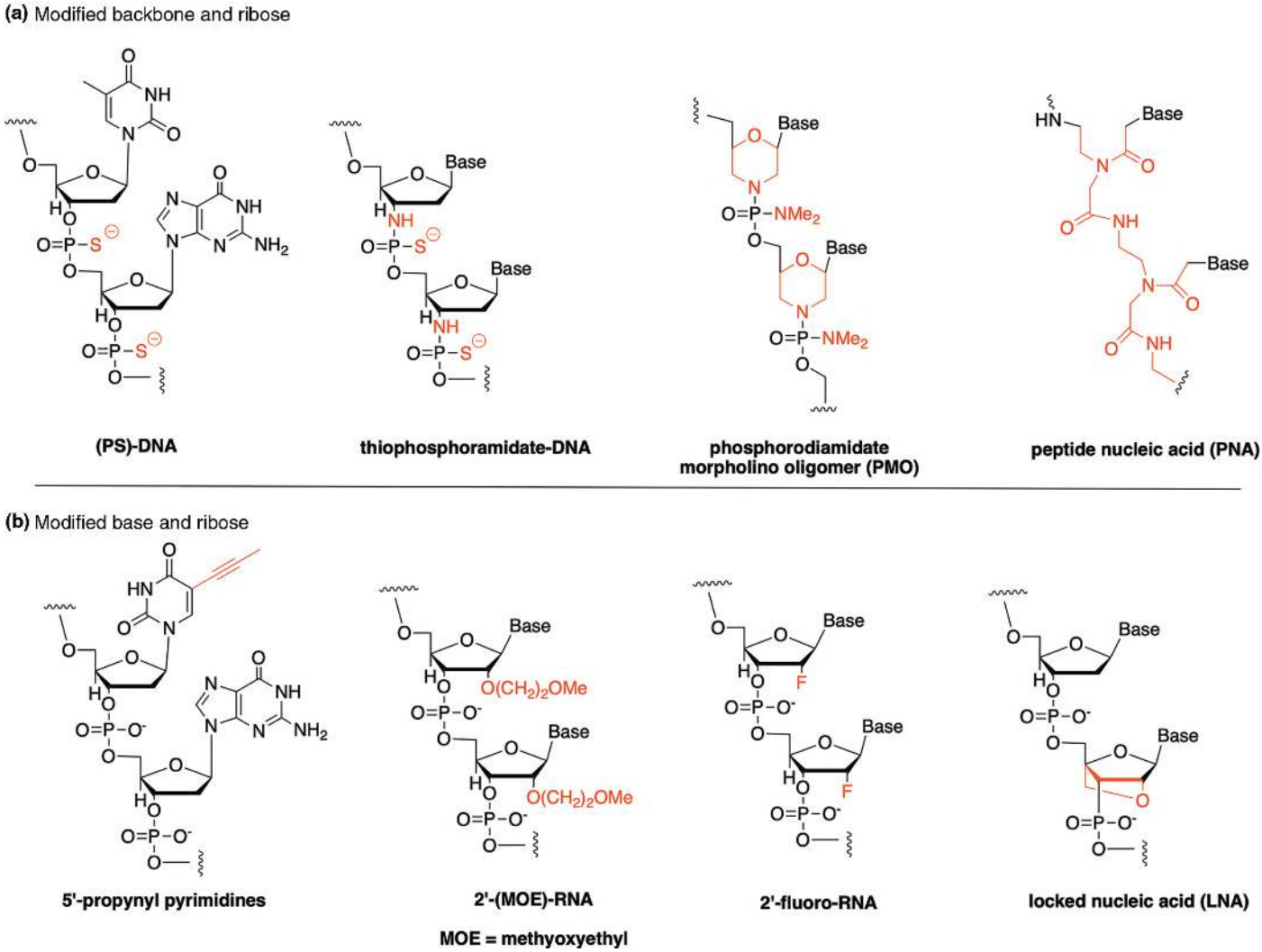


Figure 9.18 Structures of modified oligonucleotides. (a) Examples of structures with modified backbones and/or ribose, including phosphorothioate, thiophosphoramidates, phosphorodiamidates, and peptide nucleic acids. (b) Examples of structures containing modified bases and/or ribose, such as 5-propynyl thymine, 2′-methoxyethyl, 2′-fluoro, and locked nucleic acid (LNA) analogs. Modified regions are shown in red.

9.3.3 Small Molecules Can Bind to RNA and Modify Splicing

As described earlier, ASOs that act to modify splicing have found clinical success, yet there still remain challenges to their therapeutic use. Therefore, recent efforts have been aimed at identifying small molecules that have similar effects. It is known that a number of genetic diseases are caused by mutations that disrupt the RNA splicing process, resulting in a lack of functional protein production. Spinal muscular atrophy (SMA) is one such genetic disorder, where the protein survival motor neuron (SMN) in these patients is shortened and, therefore, non-functional. High-throughput screening led to the identification of small-molecule starting points that could increase the amount of functional protein produced as well as decrease the amount of mutant protein via splicing modulation. Further optimization to improve selectivity and decrease side effects led to risdiplam, approved in 2020. Risdiplam (**Figure 9.19**) acts as a molecular glue that stabilizes the interactions between the spliceosome protein U1 snRNP and adenine in the bulged 5′ splice site, a transient double-strand RNA structure. The result is repair of the splice site and restoration of functional SMN. Based on NMR studies of an earlier analog, SMN-C5, the protonated amine forms an

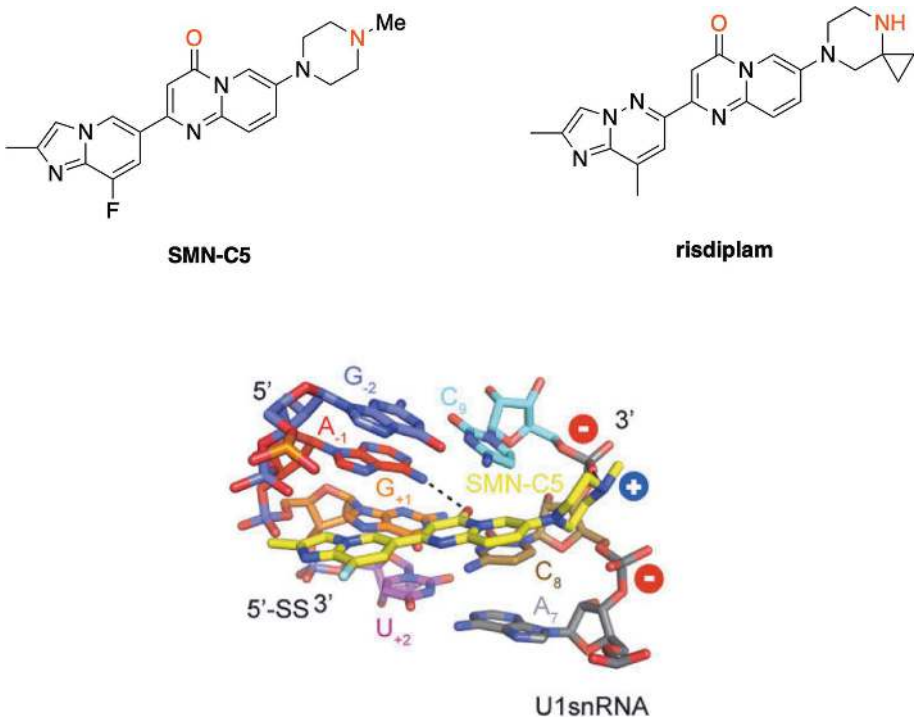


Figure 9.19 Structures of SMN-C5 and risdiplam with key features involved in RNA and U1snRNA binding in red (carbonyl, amine); NMR-based structure of SMN-C5 (yellow) bound to duplex RNA (pdb 8R62). Dotted line indicates biomolecular interaction between oxygen atom in SMN-C5 and adenine in RNA.

Source: From Malard, F.; Wolter, A. C.; Marquevielle, J.; et al. (2024). *Nucleic Acids Research* 52:4,124–4,136. With permission from the Oxford University Press.

electrostatic interaction with the RNA phosphate backbone, and the carbonyl interacts with an adenine base within the RNA duplex. Other work toward identifying and optimizing small-molecule splice modulators has continued, with several molecules in advanced clinical trials.

As seen in Section 9.5, drugs that target RNA directly are far fewer in number than those that target DNA. While several classes of antibacterial drugs have been found to target bacterial ribosomal RNA, these were found from phenotypic screening rather than through rational design. In contrast, antisense drugs are designed to target specific sequences and are optimized for improved physical and pharmaceutical properties. These RNA-targeted agents (the antibacterials that bind to the ribosome, as well as antisense oligonucleotides) affect protein synthesis, as do the small-molecule splice modulators. Designing selective drugs that will target a specific RNA sequence is just beginning to be successful.

9.4 SUMMARY

Approximately 1% of drugs target DNA and RNA, far fewer than those that target receptors, enzymes, and nontraditional targets. Most DNA-targeted drugs are older therapies that are most frequently used to treat cancers. They work through multiple mechanisms, but all require binding to DNA. Some drugs form strong noncovalent interactions, others alkylate DNA, and a third group abstracts radicals from the DNA chain, resulting in breakage of the DNA strand. While still widely used, they all suffer from serious toxicity due to lack of selectivity and specificity. In contrast, antisense therapy targets specific RNA sequences. This approach has been investigated since the late 1990s and has recently come to fruition, with several novel compounds on the market and many in clinical trials for a variety of conditions. Medicinal chemistry strategies to design metabolically stable and permeable oligonucleotides have been an important component to this approach. More recently, small molecules that bind to RNA have been developed which modify RNA splicing, leading to the production of functional proteins, a strategy particularly useful for certain genetic diseases. The potential to design highly selective compounds to target rare genetic disorders is an exciting opportunity for medicinal chemists.

CASE STUDY 1 DOXORUBICIN AND ANALOGS

The first approved chemotherapy drugs, the nitrogen mustards and folic acid derivatives, were introduced in the 1940s to treat cancer. The realization that drugs rather than surgery alone could be used to treat cancer spurred the search for novel anti-tumor agents worldwide. At about the same time, with the discovery of penicillin came the recognition that microbes can produce medicinally active agents. These two findings suggested that natural products, including those from both microbes and plants, might be sources of anti-cancer agents.

The first anthracycline reported was β-rhodomycin II, which contains the fused tetracyclic core common to all anthracycline antibacterial drugs. It was one in a mixture of red compounds isolated from the soil bacterium *Streptomyces purpurascens* in 1951 by H. Brockmann and K. Bauer (**Figure 9.20**). Independently, the C-10-aglycon rhodomycin B was identified in an Indian soil sample by Federico-Maria Arcamone at Farmitalia Research Laboratories, in Italy, in the late 1950s. These red compounds

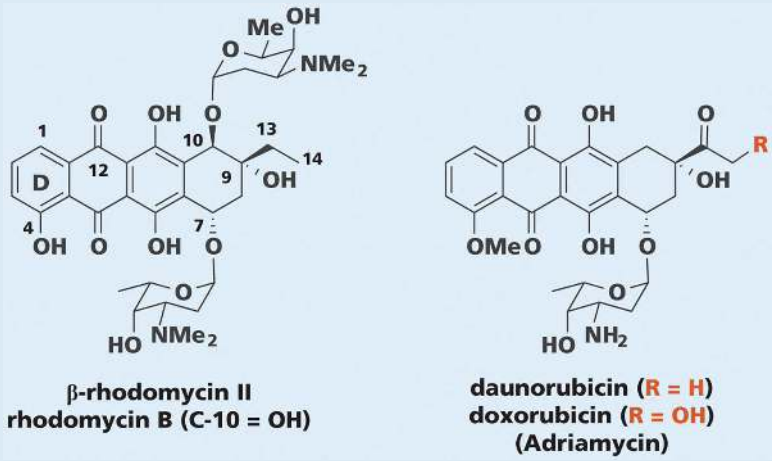


Figure 9.20 Structures of early anthracyclines. β-rhodomycin II, rhodomycin B, daunorubicin, and doxorubicin. All contain the anthracycline-defining fused tetracyclic core, with the D ring and numbering scheme indicated.

exhibited both antibacterial and anti-tumor activity in preliminary studies but were too toxic in animal models to be pursued as anti-cancer drugs.

In the early 1960s, the first clinically useful anti-tumor anthracycline was isolated independently at Farmitalia and Rhone-Poulenc. Daunomycin was derived from an Italian soil sample found near Castel del Monte that contained the microorganism *Streptomyces peucetius*. Farmitalia named the compound they isolated daunomycin, derived from the word *Daunii*, the name of a pre-Roman tribe. A group at Rhone-Poulenc isolated the same compound and named it rubidomycin, from the word *rubis*, for *ruby*. Both groups began preclinical testing and determined that they were working with the same compound, which was renamed daunorubicin. The compound proceeded rapidly to clinical trials, where it exhibited excellent activity against leukemias but also showed significant cardiotoxicity. This side effect results in irreversible damage to the heart muscle with long-term use. Despite this liability, daunorubicin became an important anti-cancer drug. The Farmitalia group induced mutations into the microbe used to produce daunorubicin, creating a variant strain that produced a new compound, doxorubicin (also known as Adriamycin), the C-14 hydroxy analog of daunorubicin (see **Figure 9.20**). It entered clinical trials in 1968 and was approved in the United States in 1974.

Although there is only one change between daunorubicin and doxorubicin, the replacement of a hydrogen at C-14 with a hydroxyl group, the anti-tumor spectrum for doxorubicin includes a variety of carcinomas and sarcomas for which daunorubicin is ineffective. The success of doxorubicin led to a search for additional analogs with broader profiles and reduced cardiotoxicity. Most analogs are variants at the C-14 position, the amino sugar (daunosamine), or the terminal aromatic D ring.

Syntheses of over 1,000 anthracyclines have been reported, with only a few exhibiting an anti-tumor and toxicity profile that would allow advancement into the clinic. The carbohydrate analogs epirubicin and esorubicin were synthesized at Farmitalia (Figure 9.21). Epirubicin has only one change from doxorubicin: the inversion of the 4′-hydroxyl on the carbohydrate from the axial to the equatorial isomer. It has the same anti-tumor profile but reduced cardiotoxicity, which may be attributed to changes in metabolism and elimination. Esorubicin, lacking the hydroxyl on the carbohydrate, is even less cardiotoxic. A possible explanation is that it does not generate oxygen-containing radicals, which may contribute to cardiotoxicity. Amrubicin and valrubicin are newer carbohydrate analogs that were approved in the early 2000s. Amrubicin contains a simplified carbohydrate, replaces the C-9 hydroxyl with an amine, and contains an unsubstituted D ring. Valrubicin is a semisynthetic analog derivatized as

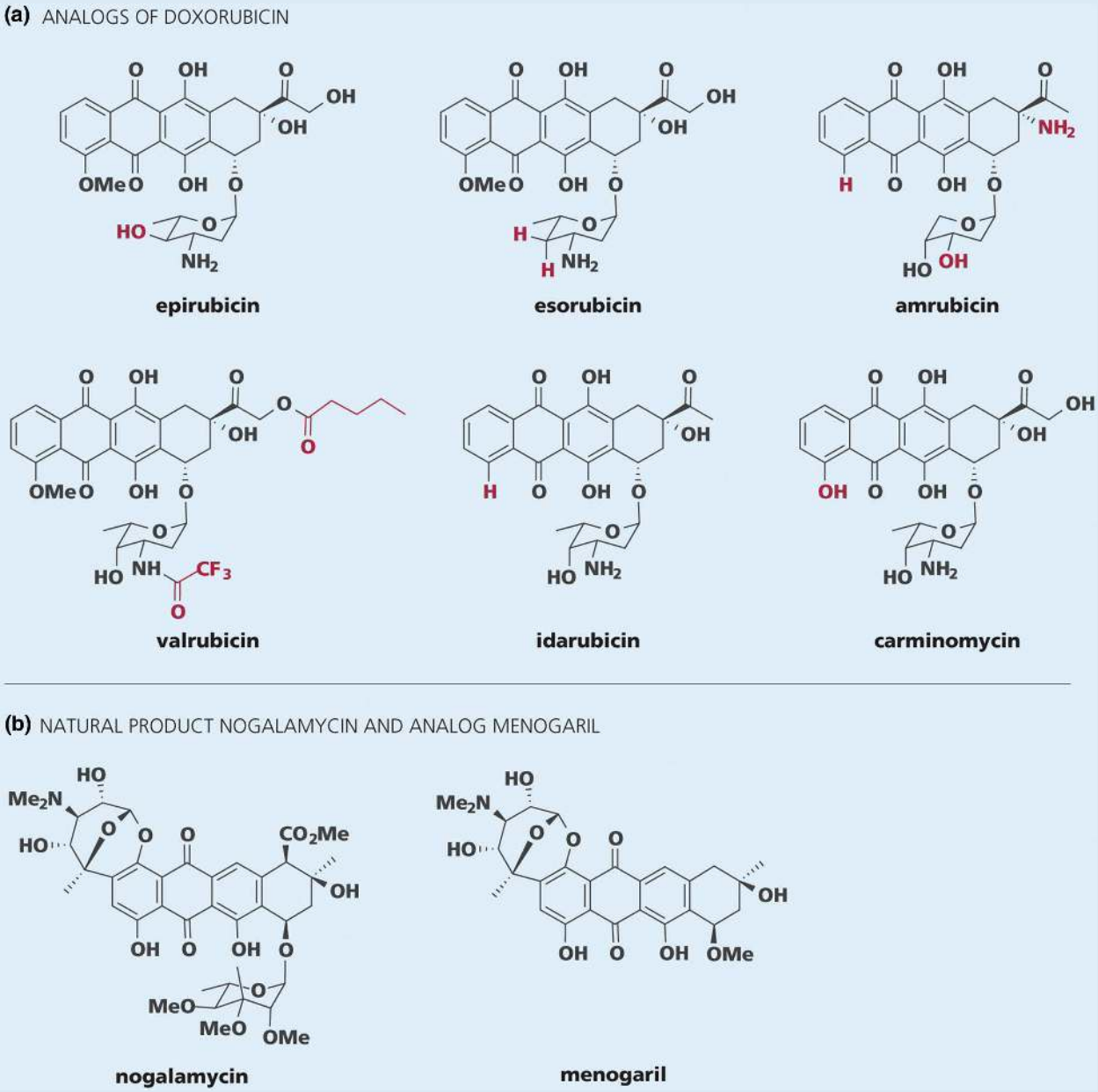


Figure 9.21 Structures of anthracyclines. (a) Structures of synthetic and semisynthetic analogs of daunorubin and daunomycin. Epirubicin, esorubicin, amrubicin, and valrubicin contain modifications in the carbohydrate residue; amrubicin also replaces a hydroxyl group with an amine and eliminates the D ring hydroxyl group. Idarubicin and carminomycin are modified on the D ring. Structural changes compared to doxorubicin are shown in red. (b) Structurally unique natural product anthracyclines nogalamycin and menogaril. Neither of these compounds is clinically approved.

an ester and amide. D ring analogs include idarubicin, a close analog of daunorubicin in which the C-4-position in the D ring is unsubstituted, and carminomycin, the C-4-OH analog of doxorubicin. One advantage of both idarubicin and carminomycin is increased oral bioavailability.

A related anthracycline natural product, nogalamycin (see Figure 9.21b), isolated from *Streptomyces nogalater*, contains an unusual bicyclic modified D ring. Menogaril is a synthetic analog of nogalamycin with a simplified A ring pattern. It is more potent and less cardiotoxic than nogalamycin. Although they are structurally interesting, neither nogalamycin nor menogaril has been approved as a drug.

The mechanisms of action of the anthracyclines include intercalation, topoisomerase II (topo II) inhibition, and radical production. Structure–activity relationships are summarized in Figure 9.22.

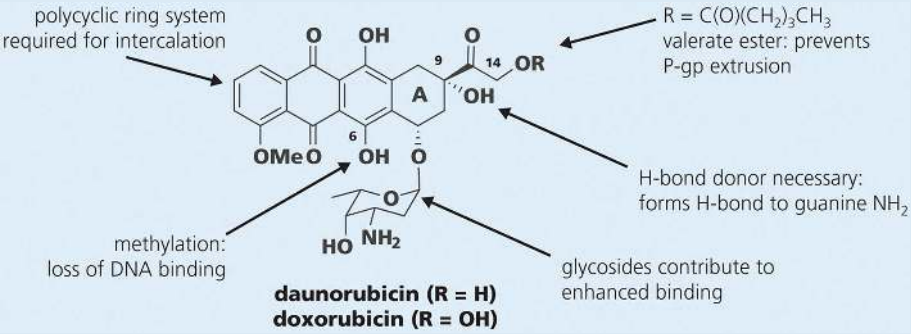


Figure 9.22 Structure–activity relationships of doxorubicin analogs. Essential functional groups include an H-bond donor at the C-9 position to bind to the primary amine in a guanine residue at the intercalation site, the hydroxyl group at C-6, and the planar polycyclic ring system. The presence of glycosides stabilizes DNA binding, with more carbohydrate units providing tighter binding. Ester groups at C-14 are tolerated and can mitigate P-gp efflux.

The C-9 OH or another hydrogen bond donor is necessary for activity and is proposed to form a hydrogen bond to guanine 2-amino groups at the intercalation site. The C-9-amino synthetic analog amrubicin maintains activity as a topoisomerase II inhibitor. Methylation of the hydroxyl at C-6 results in the loss of DNA binding and topoisomerase II–mediated cleavage. One resistance mechanism is efflux mediated by P-glycoprotein (P-gp), and analogs such as the 14-valerate of doxorubicin both increase lipophilicity and prevent P-gp efflux. The glycoside has an effect on stabilizing the DNA–drug complex, and an increase in the number of attached sugars increases the binding to DNA. The polycyclic planar ring system is required for intercalation into DNA.

Doxorubicin is an example of an older drug, discovered in the era when most anti-cancer drugs were identified through screening of natural sources. There are currently several anthracycline anti-cancer drugs marketed, and while there are limitations due to toxicity, they are an important part of the anti-cancer armamentarium.

CASE STUDY 2 DISCOVERY OF MIPOMERSEN

Mipomersen was the second antisense oligonucleotide to be approved by the FDA (2013). It is an inhibitor of the production of apolipoprotein B-100 (apoB-100) and is used to treat hyperlipidemia in patients at risk for cardiovascular disease. The protein apoB-100 is embedded in a lipid membrane of low-density lipoprotein–cholesterol (LDL-C) particles, which have a core of triglyceride and cholesteryl esters (Figure 9.23).

Cardiovascular disease is the leading cause of death in the United States, with one-third of all deaths related to the disease. A major contributing factor is coronary artery disease, which is linked to high levels of LDL-C and apoB-100. Treatment for high levels of LDL-C include the statins, which lower cholesterol levels but are not effective or tolerated in all patients. There are two forms of apolipoprotein B, apoB-48 and apoB-100. ApoB-100 is the form found in lipoproteins synthesized in the liver, and about 85% of apoB-100 synthesized in the body becomes part of LDL particles.

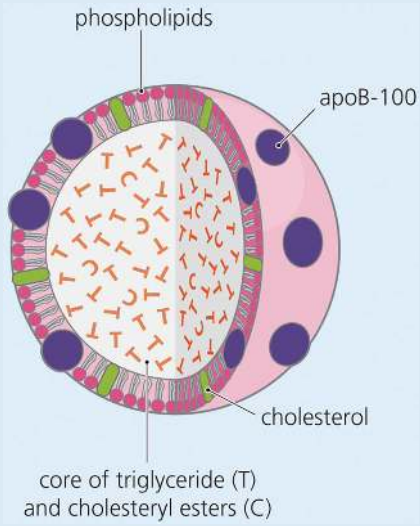


Figure 9.23 Rendition of low-density lipoprotein cholesterol particles. The protein apoB-100 is embedded into a lipid membrane that surrounds a core of triglyceride and cholesteryl esters.

A novel approach was proposed to lower the levels of apoB-100 by the use of antisense oligonucleotides to block the production of the protein. The second-generation phosphorothioates had been shown to have good pharmacokinetics, with wide tissue distribution, and were chosen as the basis for the drug.

A program was initiated at Isis Pharmaceuticals (now named Ionis) in the early 2000s. A number of phosphorothioate 2′-methoxyethyl (MOE) oligonucleotides were prepared, targeting various regions of the apoB-100 gene. Candidate antisense oligomers were evaluated for their ability to reduce levels of apoB-100 mRNA in mouse hepatocytes. Approximately 25% of the oligomers were active, with ISIS 147764 and ISIS 147483 being the most potent, showing 90% inhibition of production of apoB-100 mRNA at 150 nM. ISIS 147764 and ISIS 147483 targeted two different sequences found in the coding region for murine apoB-100:

ISIS 147764: **5′-GTCCC**TGAAGATGTC**AATGC-3′**

ISIS 147483: **5′-ATGTC**AATGCCACAT**GTCCA-3′**

The analogs developed are mixed oligomers. They contain the 2′-MOE-modified furanose only on the ends of the oligomers (shown in boldface type) that flank a sequence containing the natural RNA monomers, creating a **gapmer**. This strategy avoids the known limitation of using all 2′-modified sugars that results in reduced RNAse H activity but maintains the benefit of improved nuclease stability. Preclinical studies on ISIS 147764 in mice showed a dose-dependent reduction in apoB-100 mRNA, as well as a reduction of apoB-100, cholesterol, and LDL-C levels. A long half-life of 11–19 days predicted a long duration of action in humans, and there were no adverse toxicological results. The application of the strategy of using mixed oligomers similar to ISIS 147764 led to the inhibitor of human apoB-100, ISIS 301012 (mipomersen).

Mipomersen is a 20-residue antisense molecule that binds to human apoB-100 mRNA (positions 3249–3269) and results in RNase H–mediated degradation of the mRNA, therefore inhibiting translation of the protein. While the entire backbone is phosphorothioate-modified, the specific structure is considered a gapmer since the termini of the sequence are different (modified riboses) from the internal residues (natural ribose). Mipomersen has the following sequence:

G-mC-mC-T-mC-dA-dG-dT-dmC-dT-dG-dmC-dT-dT-dmC-**G-mC-A-mC-mC**

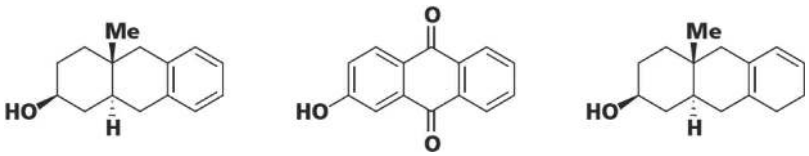
The five terminal bases at each end (shown in boldface type) contain 2′-*O*-(2-methoxyethyl) riboses, the central core contains deoxyriboses, and all cytidines are methylated at C5 (mC = 5-methylcytidine).

Pharmacokinetic studies showed rapid tissue distribution to the liver and the kidney, with slow elimination occurring over weeks. This is a desirable profile and allows for infrequent injections. Metabolism is by endonucleases, resulting in metabolites of 7–14 bases. Since there is no effect on CYP450, mipomersen is potentially free of drug–drug interactions.

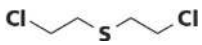
Currently, mipomersen is approved as a weekly injection to treat homozygous familial hypercholesterolemia (HoFH), a rare but fatal orphan disease. It is characterized by a mutation in the LDL-C receptor gene that results in highly elevated levels of LDL-C and early death (average life span = 33 years).

9.5 REVIEW QUESTIONS

1. Describe two mechanisms by which drugs bind to DNA.
2. Name three classes of DNA alkylators, and the electrophiles they generate, and draw the products of their reaction with DNA.
3. Why are some products of DNA alkylation unstable? (O6-G alkylation is chemically stable.)
4. Which of the compounds shown has characteristics of an intercalator? Explain your answer.



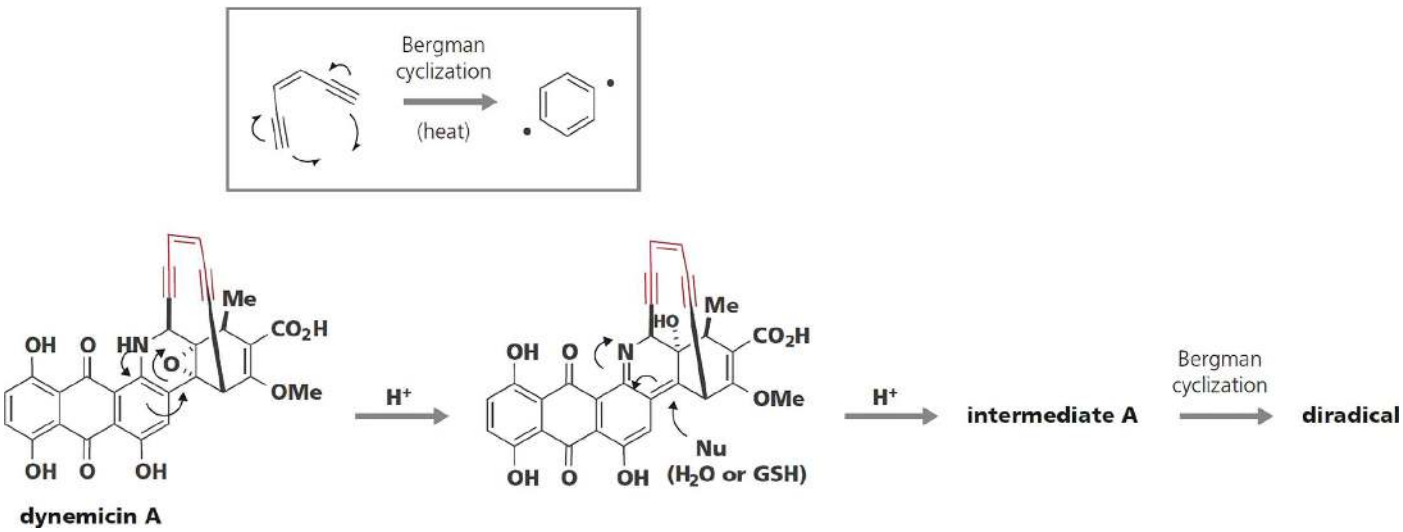
5. Explain why many drugs that directly affect DNA also exhibit significant toxicity.
6. What is the difference between a compound that nicks DNA and a compound that causes strand breakage? Give an example of a DNA strand breaker.
7. Shown is the structure of a sulfur mustard. Draw the mechanism of its reaction with DNA.



8. Briefly describe two methods that are used to detect binding of small molecules to nucleic acids.
9. What are some of the challenges in antisense drug design?
10. Explain why drugs that work by the alkylation of DNA are often prodrugs or require bioactivation.

9.6 APPLICATION QUESTIONS

1. ISIS 3521 was an antisense oligonucleotide that was in clinical trials for the treatment of lung cancer. Research this compound and answer the following:
   a. What gene is targeted, and why is it targeted?
   b. How is this oligonucleotide chemically modified?
2. Hairpin polyamides, such as those designed and studied by Peter Dervan, bind to DNA with sequence specificity. Research the basic structure, the mechanism of action, and the basis for the sequence specificity.
3. Chromomycin A3 is a natural product that binds to DNA. Research its mechanism of action and the structure of the bound complex.
4. Enediynes all undergo the Bergman cyclization *in vivo* to give diradicals. The mechanism of this cycloaromatization, studied by Robert Bergman in the 1970s, is shown here. Dynemicin A intercalates into DNA, and in one proposed pathway, bioreduction results in ring opening of the epoxide. Then, after attack of a nucleophile, either an H_2O or a glutathione (GSH) molecule, it forms an intermediate that undergoes the Bergman cyclization.



a. Identify the portion of dynemycin A that would intercalate into DNA.

b. Draw the product of nucleophilic addition (intermediate A), write a cyclization mechanism, and draw the resulting diradical.

5. Tumor resistance mechanisms to DNA alkylating agents include activity of DNA repair enzymes. Research two mechanisms for DNA repair of alkylated DNA and the work that has been done to circumvent these resistance mechanisms.

6. Research the natural product duocarmycin A. Draw its structure. What is its mechanism of action?

7. Some anti-cancer drugs that target DNA eventually cause cancer themselves. Explain this effect.

8. Compare the structure of phosphorothioate-modified DNA (PS DNA) in **Figure 9.18** to the regular DNA backbone. What changes are imparted in PS DNA by the introduction of sulfur?

9. Research the structure of camptothecin. What part of the structure might intercalate?

10. Epirubicin (see **Figure 9.21**) has a shorter half-life than doxorubicin due to metabolism to the 4′-glucuronide, leading to more rapid clearance. Draw the glucuronide metabolite.

11. Research new approaches to targeting RNA by small molecules that bind to specific RNA secondary structures. Identify three different specific secondary structures, and find examples of small molecules that bind to each.

FURTHER READING

DNA Drugs

Dervan PB (2001) Molecular recognition of DNA by small molecules. *Bioorg Med Chem* 9:2215–2235.

Di Leva FS, Zizza P, Cingolani C et al. (2013) Exploring the chemical space of G-quadruplex binders: discovery of a novel chemotype targeting the human telomeric sequence. *J Med Chem* 56:9646–9654.

Frederick CA, Williams LD, Ughetto G et al. (1990) Structural comparison of anticancer drug–DNA complexes: adriamycin and daunomycin. *Biochemistry* 29:2538–2549.

Fuertes MA, Alonso C & Perez JM (2003) Biochemical modulation of cisplatin mechanisms of action: enhancement of antitumor activity and circumvention of drug resistance. *Chem Rev* 103:645–662.

Kelland L (2007) The resurgence of platinum-based cancer chemotherapy. *Nat Rev Cancer* 7:573–584.

Pogozelski WK & Tullius TD (1998) Oxidative strand scission of nucleic acids: routes initiated by hydrogen abstraction from the sugar moiety. *Chem Rev* 98:1089–1107.

Oppenheimer NJ, Rodriguez LO & Hecht SM (1979) Structural studies of "active complex" of bleomycin: assignment of ligands to the ferrous ion in a ferrous–bleomycin–carbon monoxide complex. *Proc Natl Acad Sci USA* 76:5616–5620.

Shinomiya M, Chu W, Carlson RG et al. (1995) Structural, physical, and biological characteristics of RNA–DNA binding agent N8-actinomycin D. *Biochemistry* 34:8481–8491.

RNA Drugs

Bennett CF & Swayze EE (2010) RNA targeting therapeutics: molecular mechanisms of antisense oligonucleotides as a therapeutic platform. *Annu Rev Pharmacol Toxicol* 50:259–293.

Bouton L, Ecoutin E, Nalard F & Campage S (2024) Small molecules modulating RNA splicing: a review of targets and future perspectives. *RSC Med Chem* 15:1109–1126.

Chopra I & Roberts M (2001) Tetracycline antibiotics: mode of action, applications, molecular biology, and epidemiology of bacterial resistance. *Micro Mol Biol Rev* 65:232–260.

Dhuri K, Bechtold C, Quijano E et al. (2020) Antisense oligonucleotides: An emerging area in drug discovery and development. *J Clin Med* 9:1–24.

Actinomycins

Lown JW (1993) Discovery and development of anthracycline antitumour antibiotics. *Chem Soc Rev* 22:165–176.

Weiss TB, Sarosy G, Clagett-Carr K et al. (1986) Anthracycline analogs: the past, present, and future. *Cancer Chemother Pharmacol* 18:185–197.

Mipomersen

Crooke RM, Graham MJ, Lemonidis KM et al. (2005) LDL cholesterol in hyperlipidemic mice without causing hepatic steatosis. *J Lipid Res* 46:872–884.

Crooke ST & Geary RS (2013) Clinical pharmacological properties of mipomersen (Kynamro), a second generation antisense inhibitor of apolipoprotein B. *Br J Clin Pharmacol* 76:269–276.

Ricotta DN & Frishman W (2012) Mipomersen: a safe and effective antisense therapy adjunct to statins in patients with hypercholesterolemia. *Cardiol Rev* 20:90–95.

Yu RZ, Kim T-W, Hong A et al. (2007) Cross-species pharmacokinetic comparison from mouse to man of a second-generation antisense oligonucleotide, ISIS 301012, targeting human apolipoprotein B–100. *Drug Metab Dispos* 35:460–468.

PART III

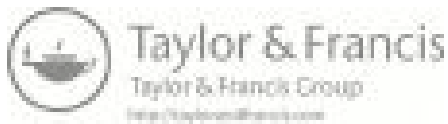
SELECTED THERAPEUTIC AREAS

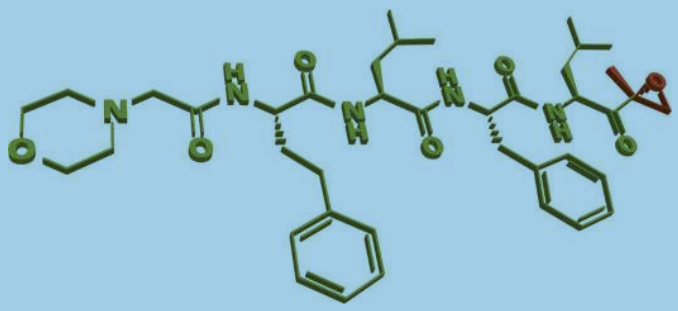
CHAPTER 10
Anti-Cancer Drugs

CHAPTER 11
Antiviral and Antifungal Agents

CHAPTER 12
Antibacterial and Antiparasitic Drugs

CHAPTER 13
Drugs Acting on the Central Nervous System


Taylor & Francis
Taylor & Francis Group



Anti-Cancer Drugs

10

LEARNING OBJECTIVES

- Explain the genetic changes that underpin the development of cancer.
- Describe why targeting cell division is an effective anti-cancer strategy.
- Understand the mechanism of action of different drugs that target enzymes involved in DNA replication.
- Explain the successful strategies for modulating the toxicity of certain anti-cancer agents.
- Identify various approaches to targeting signaling pathways.
- Describe several anti-cancer drugs that work on epigenetic targets.

10.1 INTRODUCTION

While there are hundreds of drugs used to treat cancer, it is still the second leading cause of death in the United States, with >600,000 deaths attributed to this disease each year. Each of the >100 types of cancer, usually classified by the origin of the cancerous cell, is quite distinct, as are each one's symptoms, risk factors, outcomes, and mortality rates. However, they are all characterized by the following:

- Uncontrolled growth of cells
- Invasion and destruction of neighboring tissue
- Migration, invasion, and proliferation at a location or organ remote from the primary site (**metastasis**)
- Evasion of the immune system

It is commonly believed that multiple successive genetic changes at the cellular level occur to cause cancer. A basic form of this theory was described in the 1950s by Carl Nordling, and it was more explicitly detailed in 1971 by Alfred Knudson. Knudson's **two-hit hypothesis** was a unifying model to explain how a specific cancer (retinoblastoma) developed both in individuals carrying cancer susceptibility genes and those who carried normal genes. In this simplified model, two hits, or two mutations to a gene, are required before cancer develops (**Figure 10.1**). In an individual who inherits a mutant gene at birth, only one rare event (mutation) is required before tumor founder cells develop and cancer results. In individuals who have normal genes, two rare mutations are required. The current view of cancer has further evolved to describe that multiple hits or genetic changes within a cell accumulate and allow the cell to acquire the various traits that characterize a cancerous cell. Some avoidable external factors that cause genetic changes include tobacco use, radiation or sunlight exposure,

DOI: 10.1201/ 9781003465607-13

infections, and exposure to certain chemicals. Other genetic changes can be the result of unavoidable factors, such as inherited genes.

At the center of the progression of cancer is dysregulation of the cell cycle, the process by which cells divide (Figure 10.2). All replicating cells progress through specific stages of a cycle, regulated by checkpoints, that includes the phases G1 (first gap phase), S (DNA synthesis phase), G2 (second gap phase), and M (mitosis or cell division). At checkpoints, a cell will proceed through the next phase of the cycle only if DNA is undamaged and cells are normal. If DNA or cells are abnormal, they are marked for programmed cell death, known as apoptosis. When functioning correctly, this process ensures that only healthy, non-mutant cells replicate.

Genes and the proteins they encode tightly regulate the cell cycle, both positively and negatively. **Proto-oncogenes** are normal genes that promote cell division, while **tumor suppressor genes** oppose it. These genes and their gene products work in concert to ensure that cells replicate appropriately. Most of the uncontrolled cell growth characteristic of cancer occurs as a result of mutation of these proto-oncogenes to form **oncogenes**, which then code for mutant proteins, leading to a loss of control of the cell cycle. Alternatively, mutations to tumor suppressor genes result in the loss of function of the encoded proteins that, when functional, block progression through the cycle or promote apoptosis of abnormal cells. Mutations to either oncogenes or tumor suppressor genes lead to dysregulation of the cell cycle and to uncontrolled cell growth and proliferation, the hallmarks of cancer.

Cancer treatment usually involves the removal of the tumor by surgery or destruction of the tumor with radiation and/or drug treatment (chemotherapy). Drug treatment may stop cell growth and proliferation directly, or it may stop the signals that promote it. In many cases, therapy involves a combination of these approaches. First-generation cancer drugs were cytotoxic agents that directly targeted the process of DNA synthesis, replication, and cell division. These drugs, working through a variety of mechanisms, are lethal to dividing cells, regardless of whether the cells are cancerous or not. For this reason, these early drugs exhibit significant toxicities related to this lack of selectivity, and many of their side effects (such as hair loss, nausea, and immunosuppression) can be directly attributed to their effects on dividing but non-cancerous cells.

Facilitating this early phase of cancer drug discovery was the relative ease of identifying potential therapeutics through the use of assays that evaluated a compound's ability to be either **cytostatic** or **cytotoxic**. Cytostatic compounds stop the proliferation of cells, while cytotoxic compounds kill the cells. Many drugs were discovered by such phenotypic assays, which are blind to a specific mechanism. In fact, the mechanism of action of these drugs was often understood only after their therapeutic anti-cancer activity was established. Many drugs from this first generation of anti-cancer agents worked by targeting DNA, either its biosynthesis or function, or through direct binding. Others target processes and proteins, such as tubulin, involved in cell division.

Historically, natural products have been a valuable source of anti-cancer drugs. This may be explained by two factors. First, cell-based anti-proliferative and cytotoxicity assays are simple and can be readily applied to track the activity in natural product extracts and the subsequent assay-guided fractionations necessary to isolate purified products. Second, organisms produce such natural products as a defense mechanism; therefore, they are optimized for toxicity and cell permeability, two features essential for anti-cancer activity. Well-known examples of natural product anti-cancer drugs discovered this way include vinblastine, paclitaxel, and doxorubicin (Figure 10.3). These compounds and their mechanisms of action were described in detail in earlier chapters: See Section 8.3.3, Case Study 2 in Chapter 1, and Case Study 1 in Chapter 9, respectively.

More recently, the field of cancer biology has developed a much deeper understanding of the mechanisms that are involved in tumor initiation, progression, and metastasis. As such, cancer drug discovery and medicinal chemistry have undergone a tremendous shift. Finding very potent, broadly anti-proliferative or cytotoxic agents is now considered a less-desirable approach to developing cancer drugs. Instead, newer drugs target either oncoproteins or tumor suppressors that have been mutated or cancer-specific changes and processes. In addition, advances in diagnostic tools can identify which patients are most likely to benefit from certain drugs and

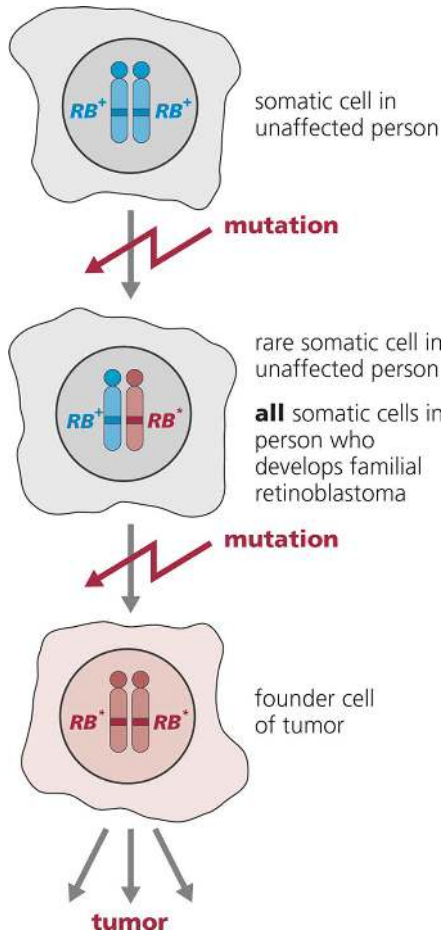


Figure 10.1 Illustration of Knudson's two-hit theory to explain the relationship between hereditary and nonhereditary forms of retinoblastoma. Normal somatic cells carry two undamaged retinoblastoma genes (RB^+; shown in blue). Two successive mutations are required to produce a cell that contains a mutated gene (RB*; shown in red) that results in a tumor. In contrast, in an individual with familial retinoblastoma who inherits a damaged gene (RB^+, RB*), only one mutation to produce the cancer-causing cell is required. Both individuals develop cancer, but the path to acquiring two hits varies.

Source: Adapted from Strachan, R., and Read, A. (2010). *Human Molecular Genetics*, 4th ed. Garland Science.

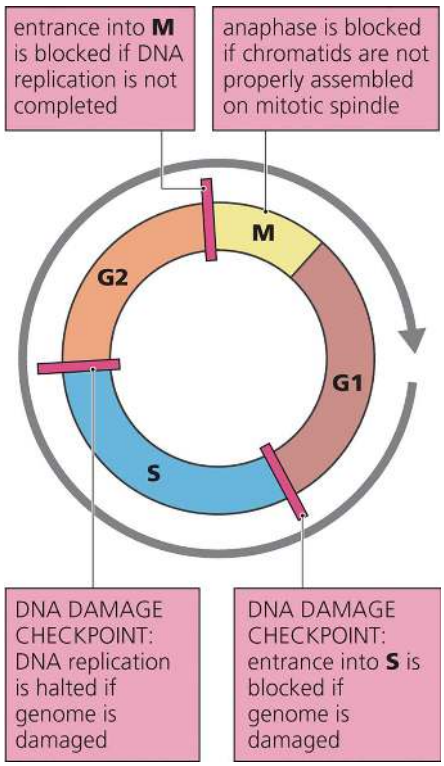


Figure 10.2 The cell cycle and its regulation by checkpoints. Cells enter G1 (first gap) phase, progress to S (synthesis) phase, followed by a shorter G2 phase, and then cell division occurs in M (mitosis) phase. Checkpoints (shown as pink lines) prevent cells with damaged DNA, incomplete replication, or improperly assembled chromatids from progressing to the next phase: These occur toward the end of G1 phase, during S phase, toward the end of G2 phase, and during M phase.

Source: Adapted from Weinberg, R. A. (2013). *The Biology of Cancer*, 2nd ed. Garland Science.

Figure 10.3 Structures of vinblastine, paclitaxel, and doxorubicin. These compounds are examples of anti-cancer drugs derived from natural product sources.

which have refractory tumors. As one example, the characterization of breast tumors as either dependent on (ER-positive) or independent of (ER-negative) the estrogen receptor dictates which treatment is offered and is most likely to be effective. Another example is the antibody trastuzumab (Herceptin) that is used to treat certain breast and stomach cancers characterized by overexpression of the ERBB2/HER2 oncoprotein; only patients whose tumors express this genetic marker benefit from treatment with trastuzumab. Finally, some newer examples take advantage of the concept of **synthetic lethality**, which occurs when two changes are required to cause cell death. Inhibitors of the enzyme poly ADP-ribose polymerase (PARP) cause DNA damage that cannot be repaired in tumors containing BRCA mutations which lack the homologous repair (HR) pathway. In cells that are not mutated, the effects of PARP can be overcome, making therapy with these drugs only effective in certain tumors.

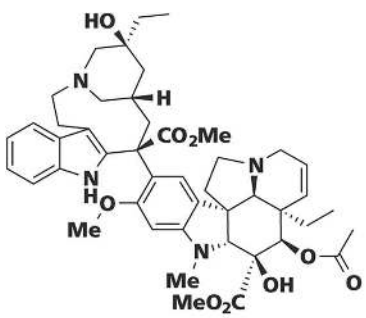
10.2 DRUGS TARGETING DNA REPLICATION AND MITOSIS

Many older anti-cancer agents target either DNA biosynthesis or its replication. As cancer cells are typically rapidly dividing, this approach is often highly effective for various types of cancer and during all stages of cancer. However, these drugs affect all dividing cells, such as immune cells and those lining the intestinal tract. In some cases, this property has been applied for use in nonmalignant disorders. Several are used to treat infectious diseases, where they target the organism's ability to replicate. Autoimmune disorders, such as multiple sclerosis and arthritis, which are characterized by the proliferation of immune cells, can also be effectively treated with drugs that target DNA synthesis, replication, and cell division. Despite these applications, these drugs have significant toxicities associated with them based on their indiscriminate effects on dividing cells. In addition, their long-term use can itself lead to cancer.

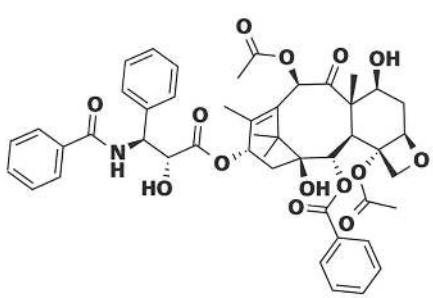
10.2.1 Drugs May Target DNA and DNA Processing Enzymes, Such as Polymerase and Topoisomerase

A variety of mechanisms that target DNA directly and processes that involve DNA have been applied to anti-cancer drug design. Some agents bind to DNA and/or cause lesions in it, thereby making it incompetent for replication. These were described in detail in Chapter 9 and include alkylating agents that damage DNA (such as dacarbazine, cyclophosphamide, and carboplatin), intercalators (for example, daunorubicin, doxorubicin, and mitoxantrone) that bind to the DNA duplex and perturb its three-dimensional structure so that it is not efficiently replicated, and compounds that make lesions in the DNA strand or duplex (like bleomycin and calicheamicin). In general, these compounds are effective, but they cause considerable side effects, such as nausea, bone marrow suppression, and hair loss.

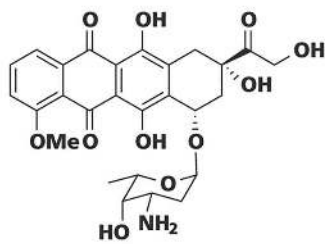
Rather than directly targeting DNA, an alternative cancer drug design strategy is to target DNA processing enzymes. These enzymes use DNA as a substrate or a template to perpetuate the cell's genetic information through replication or transcription



vinblastine



paclitaxel



doxorubicin

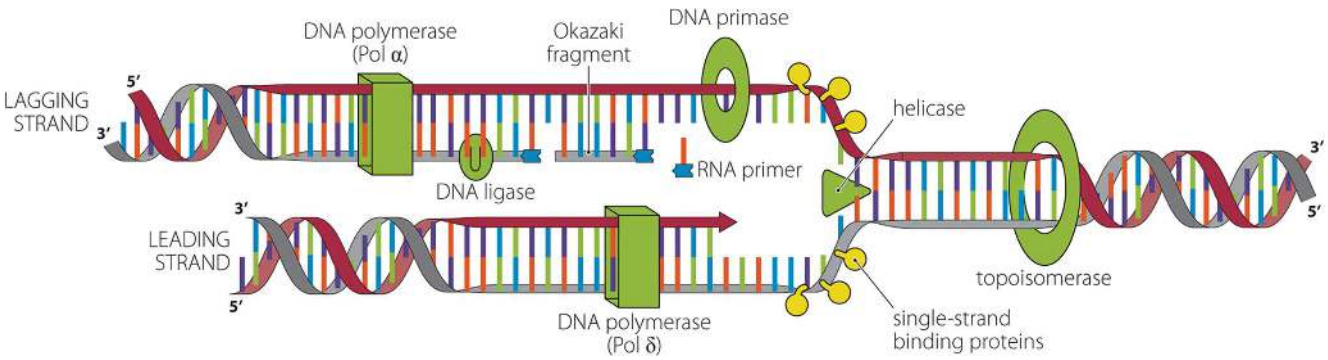


Figure 10.4 DNA replication fork and processing enzymes involved in DNA transcription and replication. Helicase (shown as a green triangle) breaks duplex DNA into single strands, known as leading and lagging strands. DNA primase (shown as a smaller green ring) initiates replication. DNA polymerases, also known as polymerase α and δ (shown as green boxes), catalyze the formation of the new strand, and DNA ligase (shown as a small green oval) links the short Okazaki fragments formed in the lagging strand. Topoisomerase (shown as a larger green ring) relaxes supercoiling of the double strand ahead of the replication fork.

into mRNA, and their inhibition results in cell death. Drugs targeting DNA processing enzymes often carry many of the same liabilities as those that directly bind to DNA. However, improvements in safety can sometimes be realized since the processing enzymes are active only in replicating cells, while DNA is present in all cells. Some key enzymes involved in DNA replication and processing are polymerases (see Section 7.5), topoisomerases, and helicases. While a thorough discussion of DNA replication is beyond the scope of this text, the basic process is outlined in **Figure 10.4**.

During the process of DNA replication, the enzyme helicase splits the double strand into two separate template strands, creating the replication fork. The two template strands are replicated in a process catalyzed by polymerase after initiation by primase. Both strands, the leading strand and the lagging strand, are replicated in the 5′–3′ direction. While the leading strand is replicated continuously, the lagging strand is replicated in short pieces, called Okazaki fragments, which are then linked together by DNA ligase. Ahead of the fork, the tension in the supercoiled strand is relieved by topoisomerase. Both polymerase and topoisomerase are targets of current anti-cancer agents.

Polymerases are multienzyme complexes that are responsible for polymerizing DNA. As described in Chapter 7, they catalyze the formation of a phosphodiester bond between a deoxynucleoside triphosphate and a DNA primer, resulting in elongation of the DNA chain. Of the four forms of eukaryotic DNA polymerases known (α, β, γ, and δ), it is believed that DNA polymerase α plays a major role in the replication of genomic DNA. Cytarabine (1-β-d-arabinofuranosylcytosine, ara-C) is one of the most widely used DNA polymerase inhibitors (**Figure 10.5a**). An analog of the natural nucleoside deoxycytidine (dC) and the HIV drug dideoxycytidine (ddC), it acts through a mechanism identical to that of HIV reverse transcriptase inhibitors, as described in Chapters 7 and 11. It is phosphorylated by deoxycytidine kinase to generate its monophosphate (Ara-C-MP), which is converted to the triphosphate (Ara-C-TP) by the action of other cellular kinases (**Figure 10.5b**). The triphosphate competes with the natural substrate, deoxycytidine triphosphate, for binding to DNA polymerase α and incorporation into the growing DNA chain. Once incorporated, chain elongation is terminated; the replicating cancer cells accumulate in the S phase of the cell cycle rather than proceeding through it, and ultimately, they die. Gemcitabine (see **Figure 10.5a**), a fluorinated nucleoside analog, acts by a related mechanism: It is converted to the corresponding triphosphate and incorporated into the growing DNA strand. However, chain termination does not occur immediately but after one additional nucleotide is added. Fludarabine phosphate (see **Figure 10.5a**) is hydrolyzed to the primary alcohol and then converted to the triphosphate, which inhibits DNA polymerase. Both gemcitabine and fludarabine have additional effects on other DNA processing enzymes, such as ribonucleotide reductase, that contribute to their anti-cancer activity.

The modest numbers of cancer drugs that work through the inhibition of DNA polymerase can be attributed to the lack of selectivity of such drugs. Rational medicinal chemistry approaches to this target are particularly difficult, as they require designing structures that are similar enough to natural substrates to be recognized and phosphorylated by host kinases yet inhibit the DNA polymerase. Furthermore, the triphosphate must effectively compete with the high concentration of natural substrates for binding to the DNA polymerase enzyme, without interfering with the large number

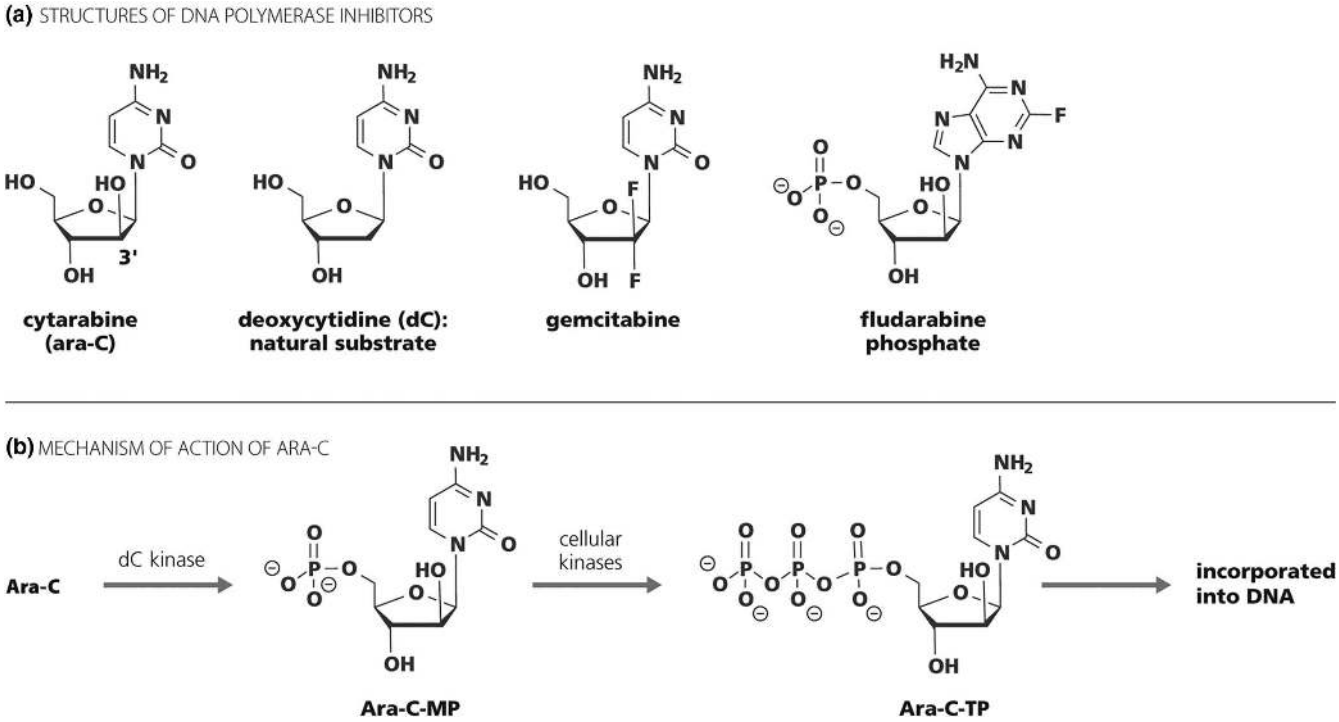


Figure 10.5 Structures of DNA polymerase inhibitors and their mechanism of action. (a) Structures of cytarabine (ara-C), deoxycytidine (dC; the natural substrate), gemcitabine, and fludarabine phosphate. (b) Formation of ara-C-monophosphate (ara-C-MP) is catalyzed by dC kinase; conversion to ara-C triphosphate (ara-C-TP) by cellular kinases is followed by the incorporation into DNA and chain termination.

of other enzymes and proteins that require nucleosides or nucleotides for their activity. In contrast, antiviral approaches can rely on the significant differences between human and viral kinases that can selectively convert nucleosides to the corresponding nucleotides only in virally infected cells, as well as the differences between human DNA polymerases and viral polymerases. As a result, targeting polymerases is a more common approach to antiviral therapy than anti-cancer therapy (see Chapter 11).

Topoisomerases, enzymes essential for DNA replication, transiently break either one (topoisomerase I) or two (topoisomerase II) DNA strands to allow for the unwinding of the helix and its replication, and then they re-form (re-ligate) the phosphodiester bond to restore the DNA tertiary structure (**Figure 10.6a**). An abbreviated cleavage mechanism, shown in **Figure 10.6b**, involves reaction of the enzyme with DNA via nucleophilic attack of a tyrosine residue on the DNA phosphodiester bond and the formation of a covalent topoisomerase–DNA complex (called a cleavable complex) at the 5′-phosphate. A number of anti-cancer drugs specifically target topoisomerase, and several others indirectly inhibit the enzyme.

One well-studied class of topoisomerase I inhibitors is the camptothecins. The parent, camptothecin (**Figure 10.7**), is a pentacyclic alkaloid natural product isolated in the early 1960s by Wani and Wall, the same scientists who discovered paclitaxel. It was isolated from the bark of the "happy tree," *Camptotheca acuminata*, a plant used in traditional Chinese medicine. Like many other early anti-cancer agents, its mechanism of action was determined well after its anti-cancer effects were described. It acts as an uncompetitive inhibitor, binding to and stabilizing the covalent topoisomerase I–DNA complex after initial strand cleavage. It does not bind to either DNA or topoisomerase alone and is termed a **topoisomerase poison** since it renders the essential topoisomerase I a DNA-damaging agent by virtue of stabilizing its covalent complex with DNA. The result, irreversible DNA strand cleavage, occurs during the S phase of the cell cycle and blocks the progression of the DNA replication fork, leading to cell death by apoptosis.

Camptothecin's insolubility, as well as its limited efficacy, prevented its further development. Therefore, drug design efforts focused on improving aqueous solubility, increasing lipophilicity to enhance cellular uptake, and/or stabilizing the required

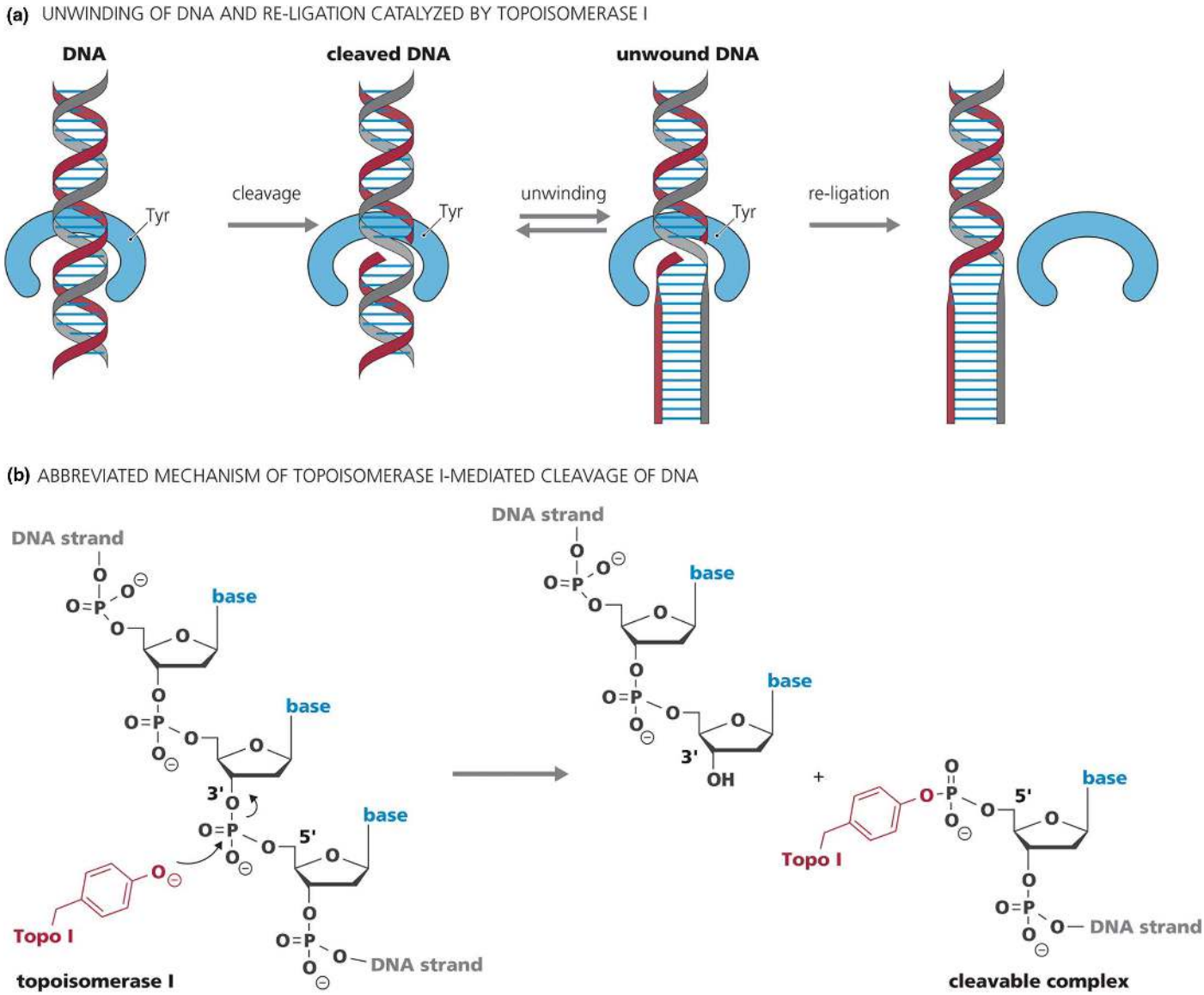


Figure 10.6 Mechanism of topoisomerase I–mediated cleavage and unwinding of DNA. (a) Overview of DNA unwinding. Once topoisomerase I (shown in blue) binds to double-stranded DNA, a tyrosine residue reacts with DNA and cleaves a phosphodiester bond (shown as a gap in the red strand). The helix unwinds, and then, after replication has taken place, topoisomerase I catalyzes re-ligation of the phosphodiester bond. (b) Abbreviated mechanism of cleavage. Nucleophilic attack of tyrosine hydroxyl from topoisomerase I (shown in red) to a phosphodiester bond in DNA (shown in black) results in breaking the phosphodiester bond and cleavage into two DNA segments. One end of the cleaved DNA strand has a 3′-OH, and the other end, termed the cleavable complex, is bound to topoisomerase I through a 5′-tyrosine phosphate ester.

but unstable lactone ring. Three synthetic or semisynthetic analogs, topotecan, belotecan, and irinotecan, are approved for advanced or recurring cancers (see Figure 10.7). Topotecan and belotecan have an amine-containing, water-solubilizing group, at positions 9 and 7, respectively, that does not affect efficacy. Irinotecan is a water-soluble prodrug; the carbamate is cleaved to deliver the active moiety, SN38. A number of other analogs that address lactone instability were evaluated clinically, but none have progressed. Strategies to do so included the incorporation of lipophilic groups (for example, silatecan), which sequester the molecule in the lipid bilayer and thereby protect the lactone from hydrolysis, and ring expansion, as in diflomotecan (see Figure 10.7).

The availability of camptothecin analogs (tecans) by isolation, semisynthesis, and total synthesis, as well as X-ray crystallography of DNA–topoisomerase I–drug complexes (Figure 10.8A), has generated an excellent understanding of the

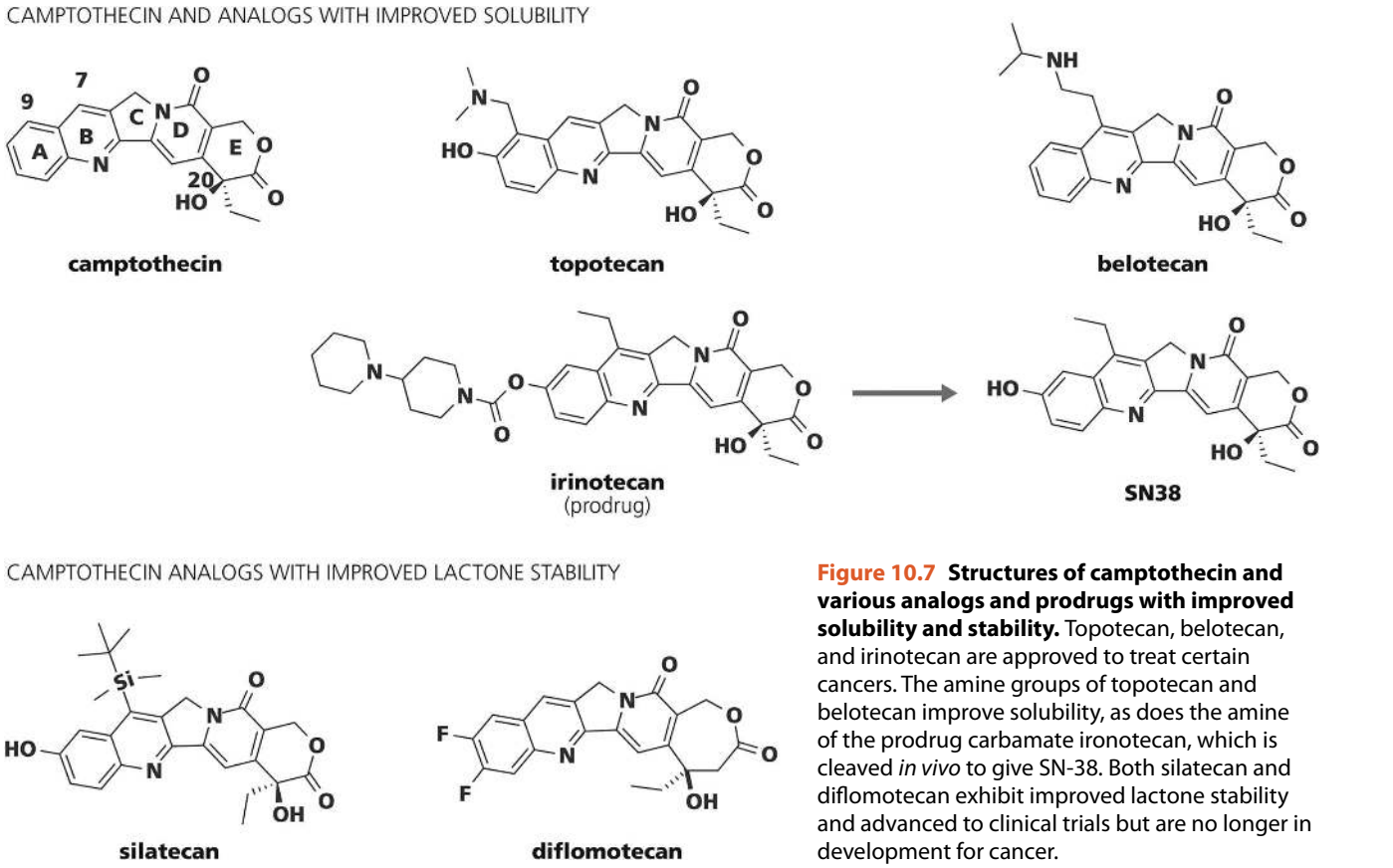


Figure 10.7 Structures of camptothecin and various analogs and prodrugs with improved solubility and stability. Topotecan, belotecan, and irinotecan are approved to treat certain cancers. The amine groups of topotecan and belotecan improve solubility, as does the amine of the prodrug carbamate ironotecan, which is cleaved *in vivo* to give SN-38. Both silatecan and diflomotecan exhibit improved lactone stability and advanced to clinical trials but are no longer in development for cancer.

structure–activity relationships that predicts activity and potency. Essential structural features include the pyridone group, the 20*S*-hydroxyl group, the lactone, and the planarity of the pentacyclic ring system. The rings mimic a DNA base pair and intercalate between the two base pairs that surround the topoisomerase I cleavage site, with carbons 7, 9, and 10 pointed toward the major groove; therefore, modifications (such

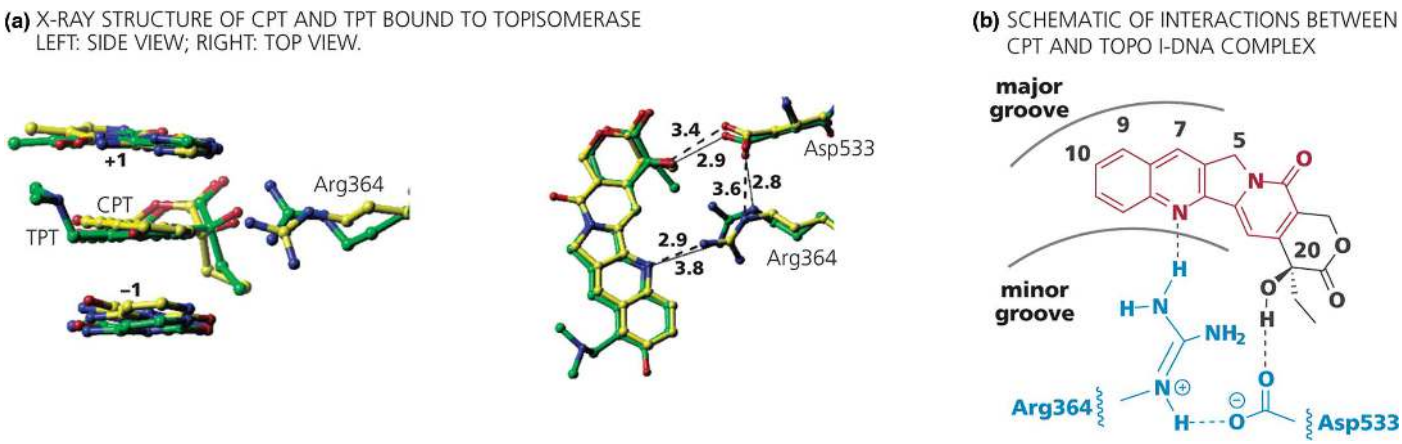
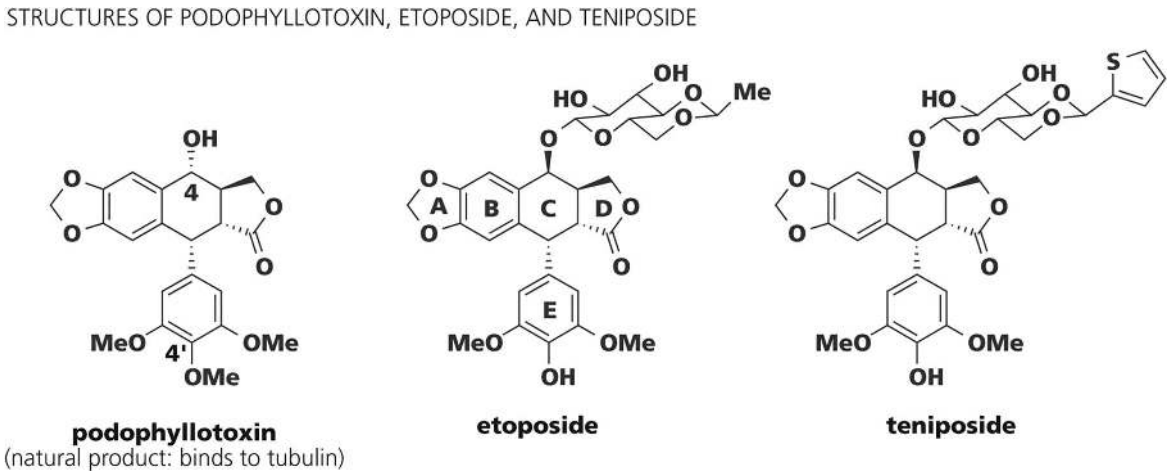


Figure 10.8 Interactions between tecans and topoisomerase I–DNA covalent complex. (a) Overlay of X-ray structures of camptothecin (CPT; shown in yellow) and topotecan (TPT; shown in green) bound to topoisomerase I–DNA covalent complex. *Left*, side view showing intercalation and interaction with Arg364 of protein. *Right*, top view showing Arg364 and Asp533 interactions with B ring nitrogen and OH at position 20. (b) Schematic of interactions between CPT and topoisomerase I while bound to DNA, including H-bond between 20*S*-OH and Asp533 and between B ring N and Arg364 (regions binding to DNA are shown in red; topoisomerase side chains are shown in blue).

Source: Adapted from Staker, B. L.; Feese, M. D.; Cushman, M.; et al. (2005). *J Med Chem* 48:2,336–2,345. With permission from the American Chemical Society.



as those found in the analogs in Figure 10.7) to these regions are tolerated. Specific interactions are highlighted in Figure 10.8b and include the following:

- Intercalation of the planar heterocyclic aromatic core between base pairs, with a preference for binding at a point one site away (+1) from G–C pair sites.
- A hydrogen bond forms between the 20*S*-OH group on the E ring and Asp533 of topoisomerase I (the *R*-configuration of the OH group is not appropriately oriented to make the same interaction).
- A salt bridge forms between Asp533 and Arg364, with a hydrogen bond from the Arg364 side chain to the B ring nitrogen in the minor groove.

Etoposide and teniposide, semisynthetic analogs of the natural product podophyllotoxin (a microtubule binder), are approved anti-cancer agents that act as topoisomerase II poisons and stabilize the DNA–topoisomerase II complex (Figure 10.9). While podophyllotoxin binds to tubulin, epimerization at C4 and demethylation at the C4′ ether give analogs that are topoisomerase poisons. The carbohydrate group, as well as the C and D rings, have been proposed to interact with DNA, while the A, B, and E rings bind to the enzyme. An X-ray crystal structure solved in 2011 of etoposide bound to topoisomerase II and a DNA fragment confirmed some of these interactions. In the complex, two molecules of etoposide bind at each cleavage site (four base pairs apart). Rings A–D intercalate between base pairs, while the carbohydrate protrudes into the minor groove of DNA and the E ring into the major groove. The B and E rings also interact with the enzyme, as does the carbohydrate group.

The tecans and etoposides are thought to exclusively target the topoisomerase enzymes. However, other known anti-cancer agents have been reported to poison topoisomerases as one of multiple mechanisms of action. For example, the anthracyclines mitoxantrone and actinomycin (described in Chapter 9 as DNA-binding agents) are reported to poison topoisomerase II. This effect may be secondary and downstream to their intercalating effects, or it may be an additional mechanism through which they act. Bacteria also have an essential enzyme, DNA gyrase, that has topoisomerase activity. It is the target of quinolone antibacterial drugs, an important class of anti-infective agents described in Chapter 12.

Figure 10.9 Structures of podophyllotoxin and semisynthetic derivatives that are topoisomerase inhibitors. Podophyllotoxin is a natural product and binds to tubulin. Etoposide and teniposide are epimerized at C4 and demethylated at C4′ and are topoisomerase inhibitors.

10.2.2 Drugs May Target Biosynthesis of DNA Building Blocks

Another strategy that directly targets DNA replication involves inhibiting biosynthesis of the pyrimidine and purine nucleotide building blocks of DNA. The rationale for this approach is that, by limiting access to these building blocks, DNA synthesis stalls and cancerous cells can no longer proliferate. Of course, as in all strategies that target DNA replication, noncancerous proliferating cells (such as immune cells and those lining the gastrointestinal or GI tract) are also affected, and toxicity results. Compounds that interfere with the biosynthesis of purine and pyrimidine nucleotides by mimicking the natural substrate and acting as enzyme substrates are known as antimetabolites (see

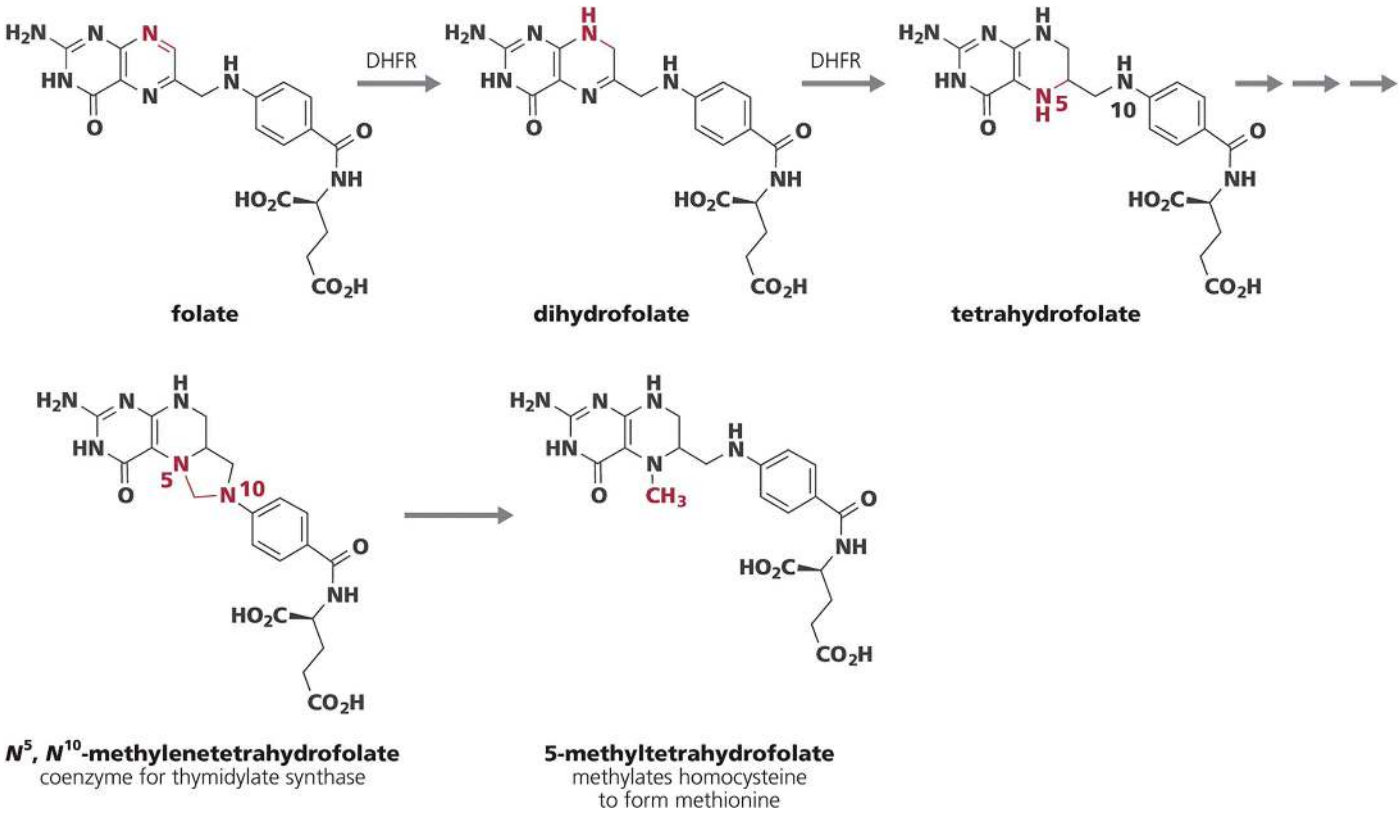


Figure 10.10 Metabolism of folates. Folate is reduced to dihydrofolate, and then to tetrahydrofolate; both reactions are catalyzed by dihydrofolate reductase (DHFR). Tetrahydrofolate is converted, in a number of steps, to N^5,N^{10}-methylenetetrahydrofolate and then to 5-methyltetrahydrofolate. Both are coenzymes required for methylation reactions to biosynthesize nucleotides and certain amino acids.

Chapter 7). Two general tactics have been especially successful. Targeting folate biosynthesis takes advantage of the fact that folate and its metabolites are required coenzymes for the biosynthesis of purine and thymine nucleotides. A second approach relies on inhibiting enzymes that produce specific nucleotides and results in the depletion of that nucleotide and the arrest of DNA synthesis. In both cases, the inhibitors are structurally related to the natural co-factor or substrate, and both approaches represent some of the earliest examples of rational drug design.

Folate, also known as vitamin B9, is essential to DNA synthesis, repair, and methylation. It served as the inspiration for one of the earliest anti-cancer drugs, aminopterin, which was first used to treat leukemia patients in 1947 and was marketed in the United States from 1953 to 1964. Folate is obtained from a number of dietary sources, such as spinach and legumes, and is metabolized to a series of derivatives, including dihydrofolate, tetrahydrofolate, N^5,N^{10}-methylenetetrahydrofolate, and 5-methyltetrahydrofolate (**Figure 10.10**). Several of these products are required coenzymes for the *de novo* synthesis of purine and thymine nucleotides. Therefore, drugs that inhibit enzymes in the folate pathway, known as **antifolates**, inhibit DNA synthesis by eliminating key coenzymes required for the synthesis of nucleotide building blocks.

Aminopterin (**Figure 10.11**), with a structure similar to that of folate, was designed on the basis of Sidney Farber's observation that leukemia patients with folate deficiencies had improved outcomes. Although aminopterin was known to antagonize the production of folate, its actual mechanism of action, the inhibition of dihydrofolate reductase (DHFR), was discovered later. DHFR is one of the enzymes responsible for folate metabolism; it catalyzes the reduction of folate to dihydrofolate and then to tetrahydrofolate in two separate steps (see **Figure 10.10**). Aminopterin therapy reduces the amount of folate metabolites produced, such as tetrahydrofolate and N^5,N^{10}-methylenetetrahydrofolate, a coenzyme required by the enzyme thymidylate synthase (TS). TS is required for the synthesis of thymine nucleotides. By inhibiting the production of its required coenzyme, TS is indirectly inhibited, thymine nucleotides are not synthesized, and DNA synthesis stops.

Methotrexate (see **Figure 10.11**), an N-methylated analog of aminopterin, supplanted its use due to improved efficacy and safety profiles. In fact, methotrexate is

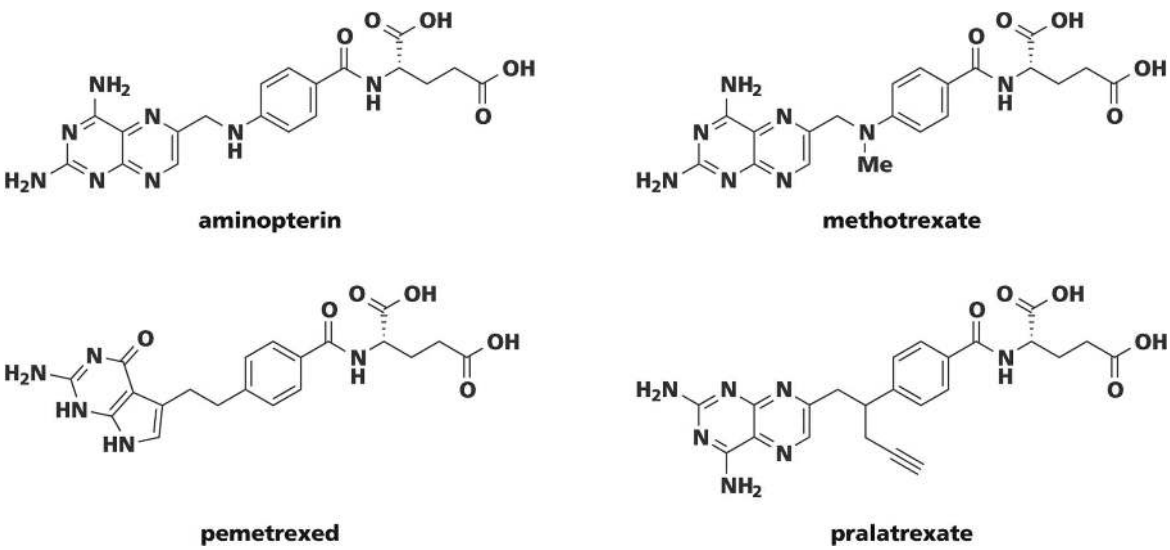


Figure 10.11 Structures of inhibitors of folate metabolism. Aminopterin and methotrexate inhibit DHFR as their primary mechanism of action; more recently approved drugs pemetrexed and pralatrexate have multiple effects on folate metabolism.

still widely used today to treat cancers as well as autoimmune diseases, such as rheumatoid arthritis, Crohn's disease, and multiple sclerosis. The role of DHFR in cell proliferation has also been exploited in the area of infectious diseases, with inhibitors of bacterial folate biosynthesis (including bacterial DHFR) also finding clinical use (see Chapter 12).

Significant efforts went toward understanding how antifolates work and identifying new DHFR inhibitors as anti-cancer drugs. These studies find that, despite the focus on inhibiting one enzyme, various pathways and enzymes are inhibited and an assortment of effects are observed. Furthermore, metabolites of the drugs are also likely to be active. Two examples of drugs with multiple effects on the folate pathway are pemetrexed and pralatrexate (see **Figure 10.11**). The parent drug and its metabolites inhibit several enzymes in the folate pathway as well as other enzymes that use folates as coenzymes. The effects of these drugs are probably further enhanced by their selective uptake or retention into tumor cells compared to healthy cells.

A second approach to inhibiting the biosynthesis of nucleotides is by direct inhibition of the specific enzymes required to produce nucleobases, nucleosides, and nucleotides. Some believe that this strategy offers advantages in efficacy and tolerability that are not possible with broad antifolate approaches, reasoning that folate coenzymes are required for enzymatic reactions beyond just DNA biosynthesis. The multi-step biosynthetic pathway to generate these DNA building blocks requires a series of enzymatic reactions, ultimately leading to the nucleoside triphosphate. Inhibitors often resemble the structure of the natural purine and pyrimidine bases and are further modified and activated to the corresponding nucleoside triphosphate to be effective. Not surprisingly, these inhibitors often inhibit multiple enzymes along the pathway, and even other enzymes that use nucleotides as substrates, such as DNA polymerase. The following section highlights a few examples of drugs that work primarily through the inhibition of purine or pyrimidine biosynthesis.

Fluorouracil (5-FU) is perhaps the mostly widely prescribed inhibitor of nucleotide biosynthesis. It has been used since the 1950s and is effective in a wide array of cancers, particularly colon cancer. The administration of 5-FU results in the inhibition of the enzyme thymidylate synthase (TS), which converts deoxyuridine monophosphate (dUMP) to the thymidine analog dTMP, an intermediate in the eventual synthesis of thymidine triphosphate, dTTP (**Figure 10.12**). The enzyme uses N^5,N^{10}-methylenetetrahydrofolate (partial structure shown in blue) as a coenzyme and as the source of the methyl group introduced at the 5-position of the uridine. 5-FU is converted via the natural biosynthetic pathway to 5-fluoro-2′-deoxyuridine monophosphate (FdUMP), and then, like the natural substrate, it proceeds through the TS reaction pathway,

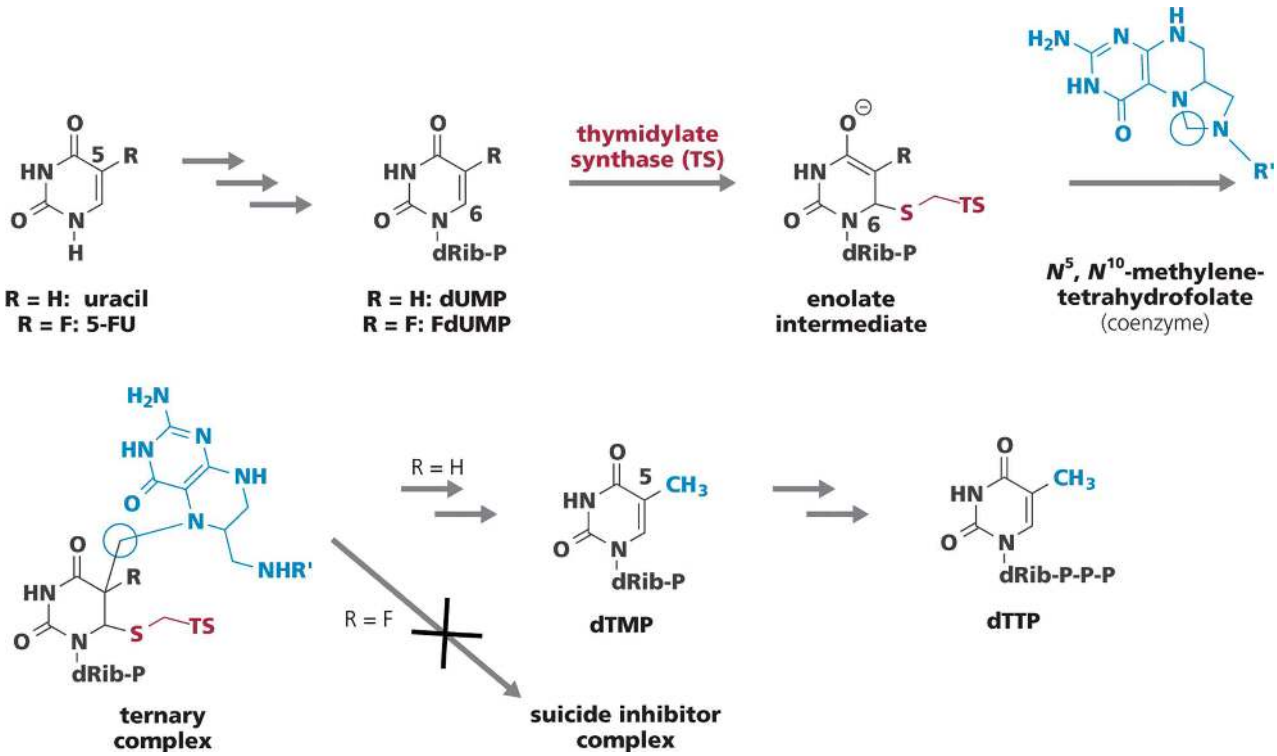


Figure 10.12 Conversion of uracil to dTTP by thymidylate synthase and inhibition by 5-FU. Uracil and 5-FU are converted to their corresponding nucleoside monophosphates, dUMP and dFUMP, respectively (dRib = deoxyribose). The addition of a cysteine of TS (shown in red) to dUMP or FdUMP leads to an enolate intermediate that reacts with N^5,N^{10}-methylenetetrahydrofolate (shown in blue). In the case of dUMP, the ternary complex (R = H) loses a hydrogen atom and eliminates the coenzyme, eventually leading to the methylated product dTMP, which is converted to the nucleoside triphosphate dTTP, which can then be incorporated into DNA. In the case of FdUMP, the presence of the fluorine atom at C5 does not allow elimination of the coenzyme from the ternary complex, and further reaction is blocked, resulting in suicide inhibition of the enzyme.

which involves the addition of a cysteine residue from TS (shown in red) to the 6-position of the uridine moiety to generate an enolate intermediate (see **Figure 10.12**). This enolate adds to the reactive methylene group (circled), present in the coenzyme N^5,N^{10}-methylenetetrahydrofolate, to afford a ternary complex. In the case of the natural substrate (where R = H), an elimination reaction occurs, releasing the coenzyme, and a second elimination (of the TS thiolate) regenerates the α,β-unsaturated system in the pyrimidine ring (dTMP), now containing the additional methyl group (shown in blue). The monophosphate is further phosphorylated to generate the nucleoside triphosphate, dTTP, which can be incorporated into DNA. However, in the case of the inhibitor adduct (where R = F), elimination is not possible due to the presence of the fluorine atom, and no further reaction of the ternary complex occurs. FdUMP acts as a suicide substrate and inactivates the TS enzyme.

As 5-FU is such an important compound in the anti-cancer arsenal, several prodrugs have been developed that release 5-FU, including capecitabine and tegafur (**Figure 10.13**). The metabolism of these compounds is such that 5-FU is selectively released, either in tumor tissue or in the liver, and therefore, these drugs are better tolerated than 5-FU itself. Capecitabine undergoes three separate enzymatic conversions. The first conversion, the hydrolysis of the carbamate, occurs in the liver. The second and third conversions, through the enzymes cytidine deaminase and thymidine phosphorylase, occur primarily in tumors. The aminal ether found in tegafur is converted to 5-FU in the liver by cytochrome P450s, and this slow process avoids conversion in the GI tract, making it orally available.

6-mercaptopurine (**Figure 10.14a**) is another example of an antimetabolite that has found use as an anti-cancer agent. Analogous to 5-FU and the DNA polymerase inhibitors described previously (as well as the antiviral reverse transcriptase

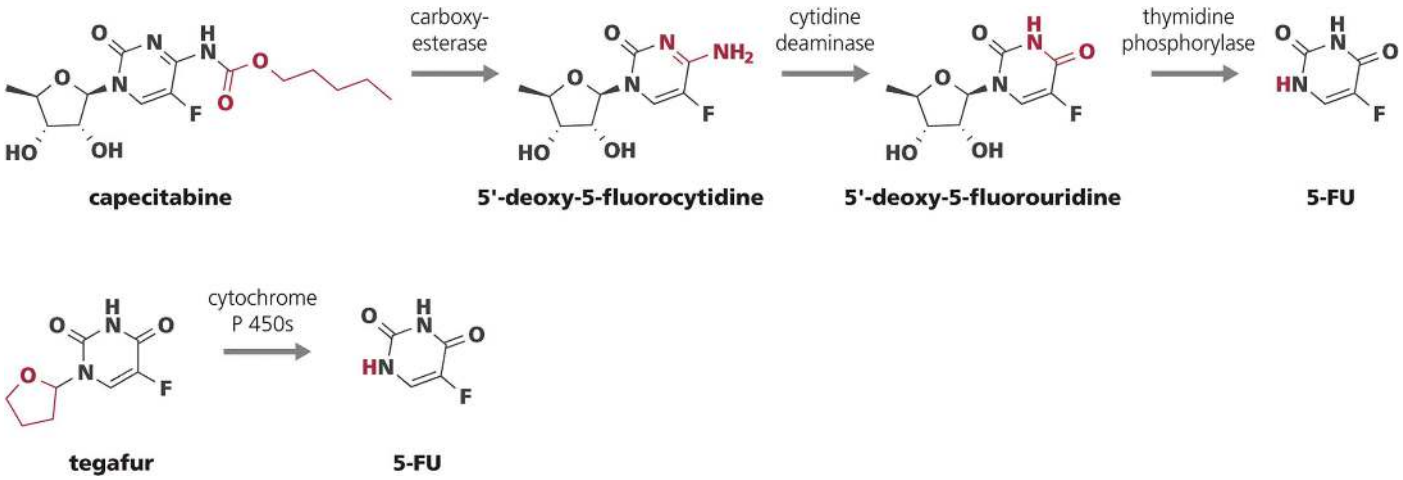


Figure 10.13 Metabolism of 5-FU prodrugs capecitabine and tegafur. Capecitabine is hydrolyzed in the liver by carboxyesterase to 5′-deoxy-5-fluorocytidine. This is then metabolized further by cytidine deaminase in the liver and tumors to 5′-deoxy-5-fluorouridine. The final step to produce 5-FU, catalyzed by thymidine phosphorylase, occurs in tumors. Tegafur is slowly converted to 5-FU in the liver by cytochrome P450s (changes shown in red).

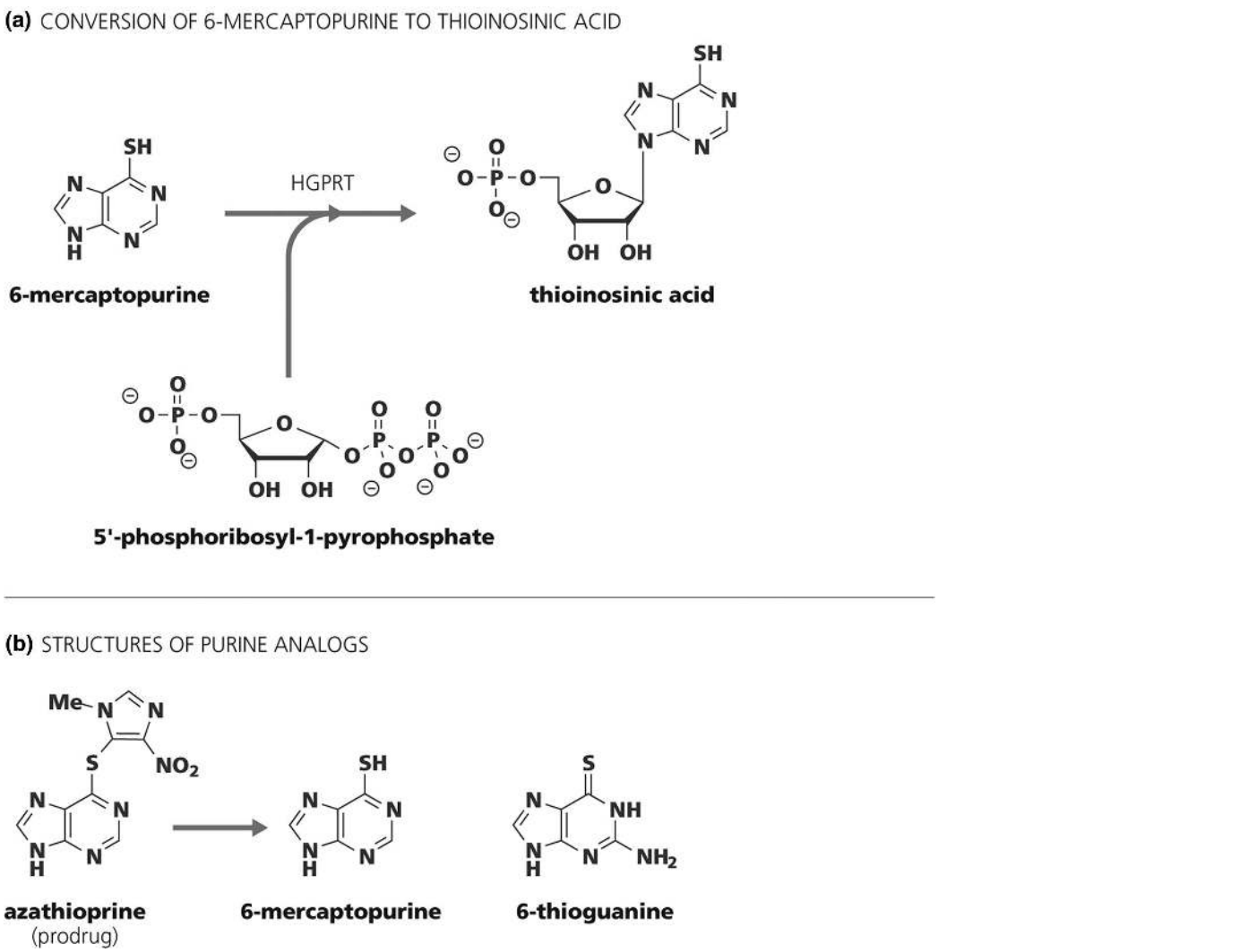


Figure 10.14 Inhibition of purine biosynthesis by antimetabolites. (a) The conversion of 6-mercaptopurine to thioinosinic acid, catalyzed by HGPRT, utilizes 5′-phosphoribosyl -1-pyrophosphate as the source of the ribose unit. (b) Structures of drugs that inhibit purine biosynthesis. 6-thioguanine is related to 6-mercaptopurine, and azathioprine is a prodrug that delivers 6-mercaptopurine.

and polymerase inhibitors described in Chapter 11), the antimetabolite is converted to a nucleotide, which is the active form of the drug. Specifically, 6-mercaptopurine is an analog of adenine, and in a reaction catalyzed by hypoxanthine-guanine phosphoribosyltransferase (HGPRT), it is converted to 6-thioinosinic acid. This molecule inhibits not only HGPRT but also amidophosphoribosyltransferase, both of which are essential enzymes that work very early in the biosynthesis of purine nucleotides. Further metabolism also occurs, as well as inhibition of additional enzyme targets. The end result is the prevention of nucleotide interconversion and the inhibition of *de novo* purine biosynthesis, thus disrupting DNA synthesis.

6-mercaptopurine, as well as 6-thioguanine (see **Figure 10.14b**), a related antimetabolite, represents one of the earliest examples of drug design for cancer therapy. In the late 1940s and early 1950s, George Hitchings and Gertrude Elion, working at Burroughs Wellcome, sought to design anti-cancer agents by mimicking the natural purines and pyrimidines. Their efforts resulted in several cancer drugs, including azathioprine (a prodrug that releases 6-mercaptopurine), 6-mercaptopurine itself, and thioguanine. This work formed the basis for future efforts that resulted in the first drugs for AIDS (for example, AZT). Hitchings and Elion, together with Sir James Black, were awarded the Nobel Prize in Physiology or Medicine in 1988 "for their discoveries of important principles for drug treatment."

10.2.3 Drugs May Target Structural Proteins Involved in Cell Division

During mitosis, structural proteins such as tubulin and its oligomers, microtubules, are responsible for transporting chromosomes to the opposite ends of the cell. Interfering with this process compromises cell division and is an effective strategy to stop cell proliferation. Natural products have been fruitful sources of molecules that interfere with tubulin dynamics, and a number of these have been developed into valuable anti-cancer agents. The structures of these molecules, as well as the mechanisms of action of these modulators of protein–protein interactions, were described in some detail in Chapter 8. The vinca alkaloids vincristine and vinblastine act to destabilize tubulin polymerization and have been used as anti-cancer agents since the 1960s. Eribulin, a synthetic fragment of the marine natural product halichondrin (see also Chapter 1), works through a similar mechanism. Taxol and epothilone represent anti-cancer drugs that are tubulin stabilizers. To date, three taxoids have been approved to treat certain cancers. The epothilones offer some advantages over taxols due to their resistance to P-glycoprotein efflux and their water solubility. In addition, because their structures are somewhat less complex than taxol's, epothilone total synthesis and semisynthetic efforts were more fruitful. These factors contributed to the relatively rapid development of the epothilones compared to that of taxol, and the semisynthetic analog ixabepilone is approved as a cancer therapeutic agent. Ixabepilone contains a lactam moiety rather than the lactone found in natural epothilones. This modification confers considerable improvements in pharmaceutical properties, particularly stability, which is a liability of natural epothilones.

10.2.4 Effects of Anti-Tumor Compounds May Be Optimized by Targeting Delivery Systems

In general, drugs that target cell division, regardless of the specific mechanism, tend to be highly effective yet also highly toxic, because they affect all dividing cells rather than only cancer cells. A solution to this issue of selectivity is to take advantage of characteristics specific to cancer cells. Toward this end, monoclonal antibodies (mAbs) have become important therapies to treat cancers, as well as other conditions. Monoclonal antibodies are manmade proteins derived from animals (usually mice) that have been immunized with an antigen to stimulate antibody production. They may be murine (-omab), chimeric (part mouse/part human, ending in -ximab), humanized (small part mouse fused to human, ending in -zumab), or fully humanized (-umab). Many cancer therapeutic mAbs bind to a protein specifically expressed on cancer cells and result in the destruction of those cells; others work through making cancer cells visible to the innate immune system. Widely used mAbs for cancer treatment include

trastuzumab (targeting HER2), rituximab (targeting CD20), and cetuximab (targeting EGFR). Pembrolizumab and nivolumab target PD-1 receptors on cancer cells. PD-1 allows tumor cells to evade the immune system, and these "immunotherapeutic" mAbs make those antigens visible, allowing the patient's own immune system to destroy the cancer cells. The success of these therapeutics has led to innovative approaches that involve conjugating a compound (known as a payload) through a linker to an antibody that recognizes epitopes selectively expressed on cancer cells (**Figure 10.15**). These are termed **antibody–drug conjugates** (ADCs) and work by first binding to the target antigen on the surface of the cancer cell, then becoming internalized by the cell, and eventually releasing the active small molecule or an active derivative inside the cancer cell, by either degradation of the antibody or cleavage of the linker.

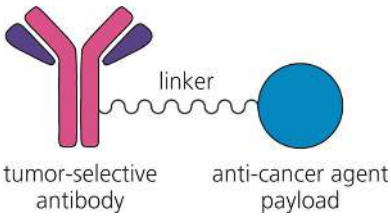


Figure 10.15 General model for antibody–drug conjugate. The ADC consists of a tumor-selective antibody that directs the drug to the tumor, a linker, and an anti-cancer agent or payload that is toxic to the tumor cell.

The design and development of ADCs are complicated by numerous factors, including the need to develop tumor-specific antibodies and to study their effects as independent entities, as well as the lack of understanding of how to design optimal linkers. Furthermore, from a medicinal chemistry perspective, considerable understanding of structure–activity relationships on the small-molecule payload is necessary to identify the most effective site for linker attachment, as are synthetic protocols for derivatization under conditions mild enough to maintain the integrity of the antibody.

Gemtuzumab ozogamicin, consisting of the DNA-cleaving agent calicheamicin (described in Chapter 9) linked to an anti-CD33 antibody, was the first ADC approved. It was used to treat acute myelogenous leukemia (AML), was voluntarily withdrawn in 2010 but re-introduced in 2017. Since then, this area has expanded considerably with >10 ADCs approved as of 2025, and many more in clinical trials. Two examples are highlighted here. Trastuzumab emtansine (**Figure 10.16**) consists of the HER-2 antibody Herceptin linked to several molecules of mertansine. Herceptin is itself a cancer therapeutic used to treat HER-2-positive breast cancer. Mertansine is an analog of the tubulin-destabilizing agent maytansine (Chapter 8), a natural product that underwent

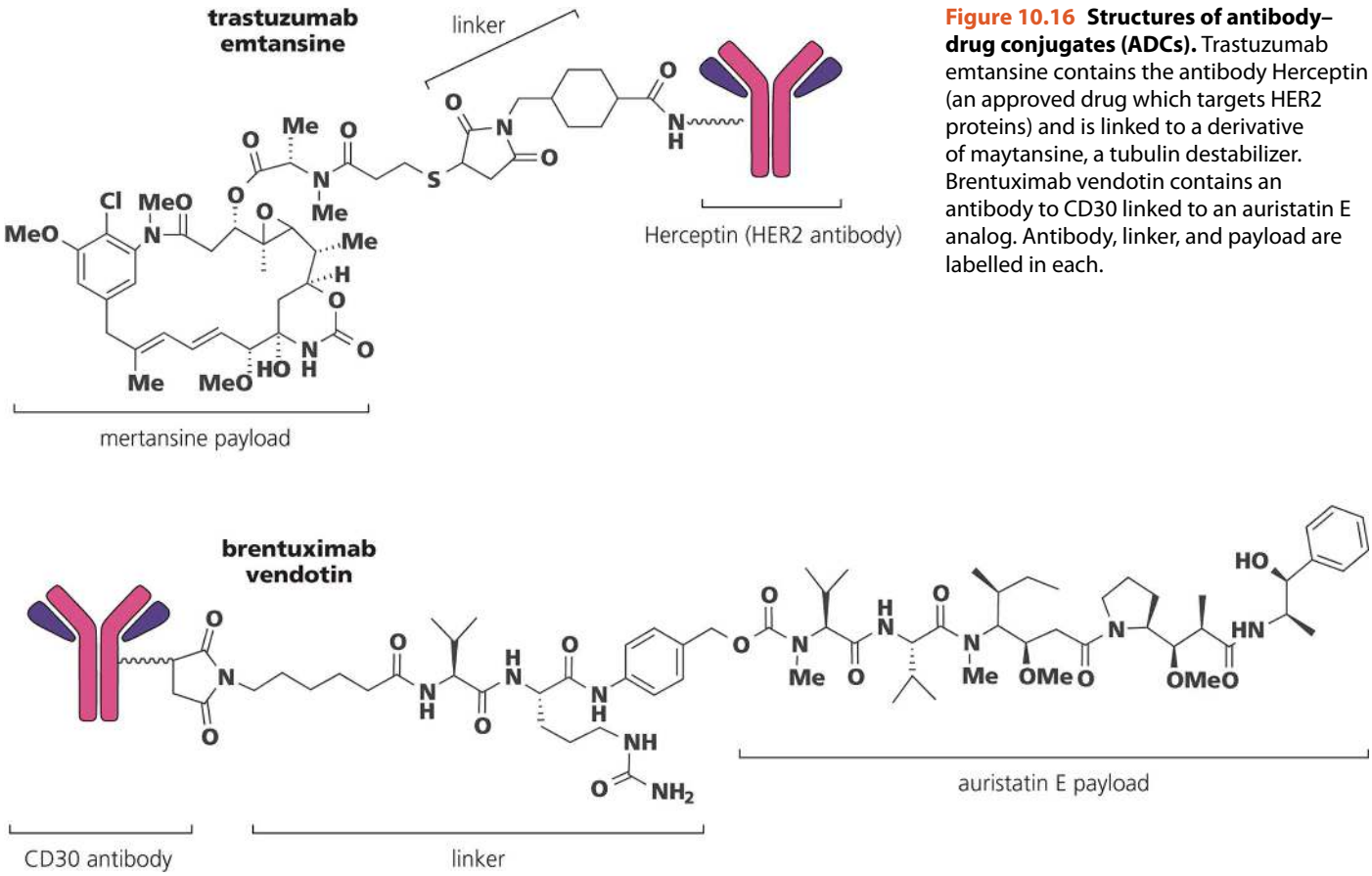


Figure 10.16 Structures of antibody–drug conjugates (ADCs). Trastuzumab emtansine contains the antibody Herceptin (an approved drug which targets HER2 proteins) and is linked to a derivative of maytansine, a tubulin destabilizer. Brentuximab vendotin contains an antibody to CD30 linked to an auristatin E analog. Antibody, linker, and payload are labelled in each.

clinical trials in the mid-1970s but was deemed too toxic for further development. Brentuximab vendotin (see **Figure 10.16**) is an ADC that uses an analog of the natural product auristatin E (see Chapter 8) as the payload, which also targets tubulin dynamics. It is linked to an antibody to CD30 that is expressed on certain lymphoma cells.

While ADC therapy directly delivers anti-cancer agents to cells expressing specific antigens, other delivery systems may also be used for cytotoxic drugs. Some exploit specific differences between cancer cells and normal cells, such as differences in pH, metabolic environment, and nutritional needs, by incorporating small-molecule moieties that are preferentially taken up by one and not the other. Another approach is the use of micelles and liposomes to carry drugs, improving their pharmacokinetics and reducing toxicity. Liposomes have been used to encapsulate drugs such as doxorubicin and taxol, effectively increasing intracellular concentrations of drugs. The use of liposomes takes advantage of the differences in vasculature between tumor cells and normal cells. Blood vessels that supply tumors have gaps in the vasculature and are leaky compared to normal vessels, allowing larger particles such as liposomes to selectively enter them.

10.3 DRUGS TARGETING ONCOGENES AND SIGNALING PATHWAYS

While macromolecules such as DNA, structural proteins, and enzymes are directly engaged in cell division and proliferation processes, many other proteins, particularly enzymes, are indirectly involved. Through tightly regulated signaling pathways, cell proliferation and survival are turned on and off. Numerous proteins in these pathways, often kinases (see Chapter 7), are positive and negative regulators. However, in cancer, the pathway becomes aberrant, and either the growth signal is turned on inappropriately or the signals to stop growth, or for cell death (apoptosis), are missing. In certain cases, these changes have occurred at a genomic level, and the resultant oncogenes or mutant tumor suppressor genes predispose a cell to uncontrolled growth. These cells produce a mutant protein, not found in other cells, that is often hyperactive and insensitive to regulation. In other cases, specific signaling proteins are overexpressed, leading to a similar effect of hyperactivity. The signaling processes are quite complex, with many proteins involved. Some examples include those in the Ras/MAP pathway and the Src/FAK pathway, PI3-kinases, Rho GTPases, and protein kinase Cs (PKCs). Many of these involve phosphorylation pathways, and a simplified picture of how mutant proteins in the process affect cell growth is shown in **Figure 10.17**. In a normal cell, the signal for growth results in a cascade of events, usually kinases phosphorylating other proteins to either activate or deactivate them. Ultimately, certain proteins enter the nucleus to turn on cell division and growth in a highly regulated manner. In aberrant cells, the same pathway is not regulated, and signals to turn on cell division and growth result in a cancer phenotype.

Upon inhibition of these aberrant proteins, the unregulated signaling pathway is attenuated, and the malignant process can be slowed or halted. Furthermore, this approach has the potential to be highly selective for transformations that are involved in certain cancers and lead to fewer side effects. This benefit is in stark contrast to older anti-cancer agents that target all proliferating cells and exhibit significant, sometimes dangerous side effects.

10.3.1 Kinase Inhibitors Target Mutant or Overexpressed Kinases Found in Cancer Cells

As the phosphorylation state often determines whether a protein is active or inactive, kinases (see Chapter 7) play a major role in signal transduction pathways. A number of oncogenes code for kinases, and many more kinases are overexpressed or become mutated in cancers. That role, coupled with their druggability, has made them a frequent target in cancer drug design (see Section 7.3). Finding highly selective kinase inhibitors is challenging due to their highly conserved ATP binding site domain. On the other hand, identification of hits and leads to initiate drug discovery programs

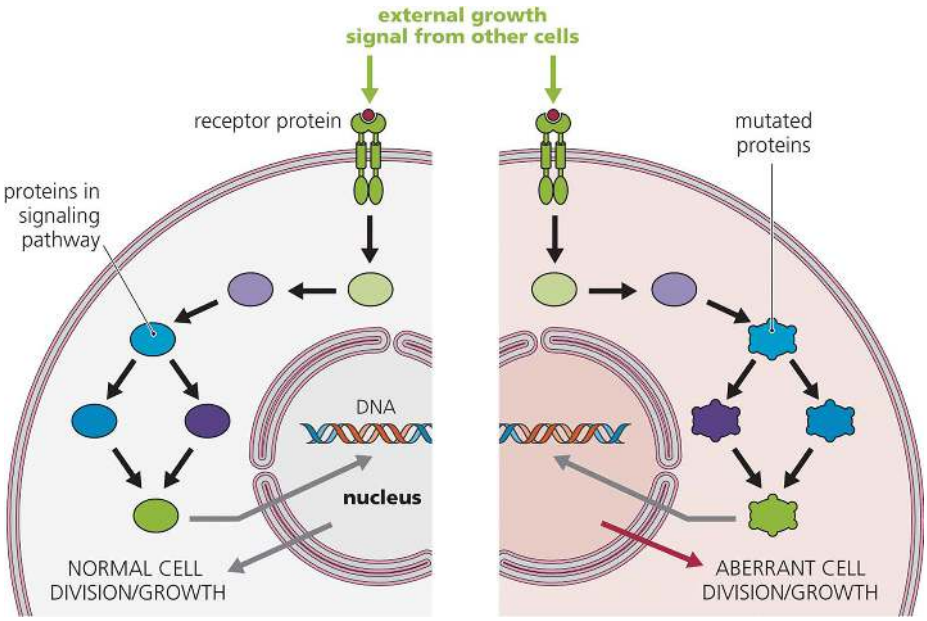


Figure 10.17 Simplified scheme of the effect of mutated proteins on signaling pathways during cell division. Upon receptor activation, proteins in the signaling pathway (depicted as ovals) are activated. In normal cells, these proteins tightly regulate cell division. Mutated proteins lead to aberrant cell division and growth.

benefits from the existence of numerous kinase scaffolds that have already been developed. This section highlights some examples of drugs that target oncogenes and signaling pathways through kinase inhibition not yet covered. More thorough treatment of the design of kinase inhibitors and several examples of anti-cancer drugs that act through kinase inhibition are found in Section 7.3.1. In addition, Case Study 1 in Chapter 7 describes the development of imatinib, the first kinase inhibitor of an oncogene product.

As described in Chapter 7, imatinib was the first example of a drug that was designed to inhibit an oncogenic protein. It inhibits Bcr-Abl, a tyrosine kinase that is the product of a mutated gene. This mutation results in high levels of kinase activity essential for the cellular transformation seen in Philadelphia chromosome–positive chronic myelogenous leukemia (CML). Since the introduction of imatinib, resistance toward it has developed, and second-generation Bcr-Abl inhibitors exhibit activity against these mutants. Unlike imatinib, dasatinib (**Figure 10.18**) binds to the DFG-in/αC helix-out conformation of the kinase and, as a consequence, is also active against another family of oncogenic kinases called Src. Bosutinib (see **Figure 10.18**) is also a dual Src and Bcr-Abl inhibitor and is a member of the quinoline family of kinase inhibitors that originated from screening of a focused library. Nilotinib (see **Figure 10.18**) is a second-generation Bcr-Abl kinase inhibitor that, like imatinib, binds to the DFG-out form of the enzyme. It is a close structural analog of imatinib, and its design was based on optimizing molecular interactions with the enzyme. By reversing the amide (shown in blue) compared to the orientation in imatinib, a more optimal geometry for H-bonds with the kinase was achieved. Dipolar interactions and hydrophobic interactions from the trifluoromethylphenyl group also contribute to improved binding (>10-fold) compared to imatinib.

Ponatinib (see **Figure 10.18**) was designed to address one of the most common resistance mutations observed after imatinib treatment: replacement of the gatekeeper threonine in the Bcr-Abl ATP-binding region with an isoleucine, known as Bcr-Abl(T315I). Many of the second-generation Bcr-Abl inhibitors make an H-bond to the Thr315 residue, and this amino acid change usually results in the loss of a key H-bond between these inhibitors and the alcohol of the Thr315 residue, resulting in their being ineffective against this resistant mutant. The substitution also introduces an unfavorable steric clash between the inhibitor and isoleucine's larger alkyl side chain that does not allow access to one of the important hydrophobic binding pockets adjacent to the ATP site. The small linear ethynyl linkage in ponatinib can accommodate the larger isoleucine side chain and also provides favorable entropic considerations by

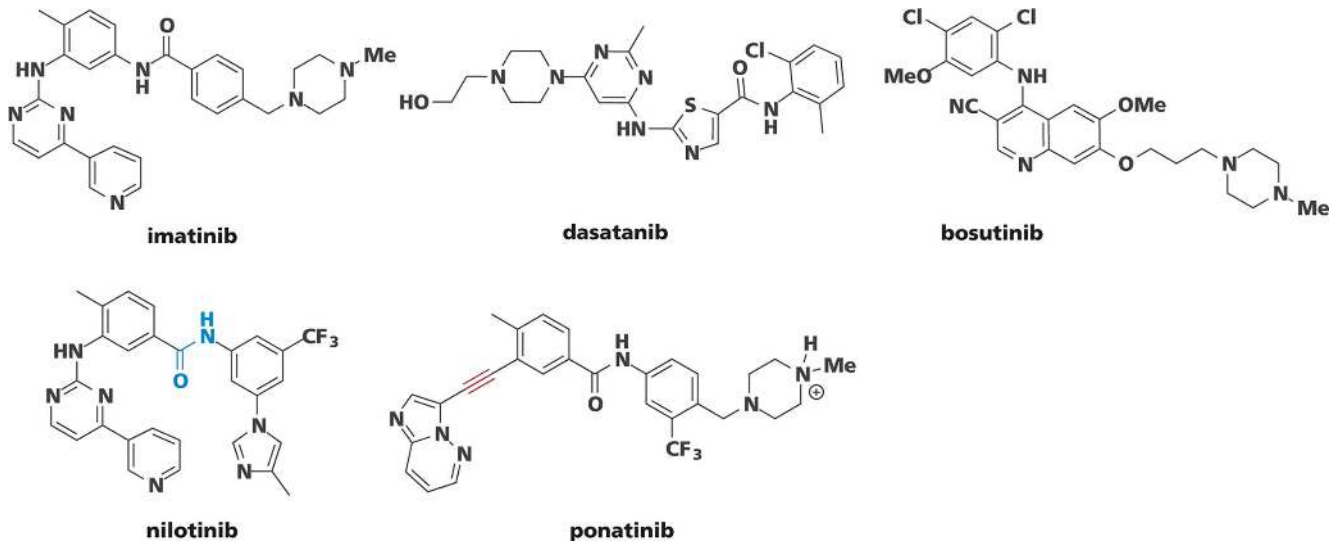


Figure 10.18 Structures of Bcr-Abl kinase inhibitors. Imatinib, dasatanib, bosutinib, nilotinib, and ponatinib are depicted. Imatinib was the first Bcr-Abl inhibitor approved; it binds to the DFG-out conformation. Dasatanib and bosutinib bind to the DFG-in conformation and exhibit broader kinase inhibition effects. The design of nilotinib, a close structural analog of imatinib, was based on optimizing interactions with the enzyme. The amide (shown in blue) is reversed compared to imatinib. Ponatinib, a DFG-out binder, was designed, by virtue of the alkyne (shown in red), to inhibit native as well as drug-resistant Bcr-Abl kinase.

locking the molecule into a favorable binding pose so that it reaches the hydrophobic pocket. Representations of the binding between Bcr-Abl and imatinib, nilotinib, and ponatinib are shown in **Figure 10.19**. These DGF-out binders all occupy the adenine site, a hydrophobic site (occupied by a tolyl group), and the allosteric site revealed in the DFG-out conformation. Common among all three are H-bonds to Met318, Asp381, and Glu286.

The *BRAF* gene, which codes for a serine/threonine kinase, is mutated in numerous cancers. This mutation causes the kinase to become constitutively active and results in an unregulated signal for cell growth. As described in Section 2.6.2, researchers at Plexxikon used a scaffold-based approach to identify novel fragment-sized molecules (MW <350) that inhibited several different kinases. The activities of the molecules were validated by X-ray crystallography using available kinase domains. With novel scaffolds that had broad activity against multiple kinases, optimization for inhibition of a specific mutant of B-Raf (V600E) led to vemurafenib (**Figure 10.20**). Vemurafenib binds to the DGF-in/αC helix-out form of B-Raf, which is stabilized by the oncogenic mutation, and this characteristic is responsible for its selectivity. It is used to treat metastatic melanomas with the V600E B-Raf mutation. Similar B-Raf inhibitors include dabrafenib and encorafenib.

The epidermal growth factor receptor (EGFR) plays a fundamental role in regulating cell proliferation, differentiation, and survival. In some cancers, EGFR is over-expressed, in others there may be mutant forms of the receptor. These are involved in multiple stages of the cancer process, such as malignant transformation, uncontrolled cell proliferation, tumor invasion and survival, and metastasis. The clinical success of monoclonal antibodies to EGFR (for example, cetuximab) provided proof of concept for the potential of small-molecule EGFR inhibitors. EGFR is a receptor tyrosine kinase, and gefitinib and erlotinib (**Figure 10.21**) both inhibit its kinase activity. In this scaffold, the N3 of the quinazoline makes a water-mediated H-bond with the key gatekeeper residue Thr390 in EGFR, while N1 makes an H-bond to the hinge region (see Section 7.3). The ether side chains in both drugs are solvent-exposed and provide improvements in solubility. In the case of gefitinib, the fluoro and chloro substitutions on the phenyl ring afford significant metabolic stability. Osimertinib is a third-generation kinase inhibitor that targets the T790M mutant form, a common mutation found in patients treated with kinase inhibitors that leads to resistance to the drugs. It is a covalent inhibitor that forms a bond to Cys797 through the acrylamide (highlighted in red).

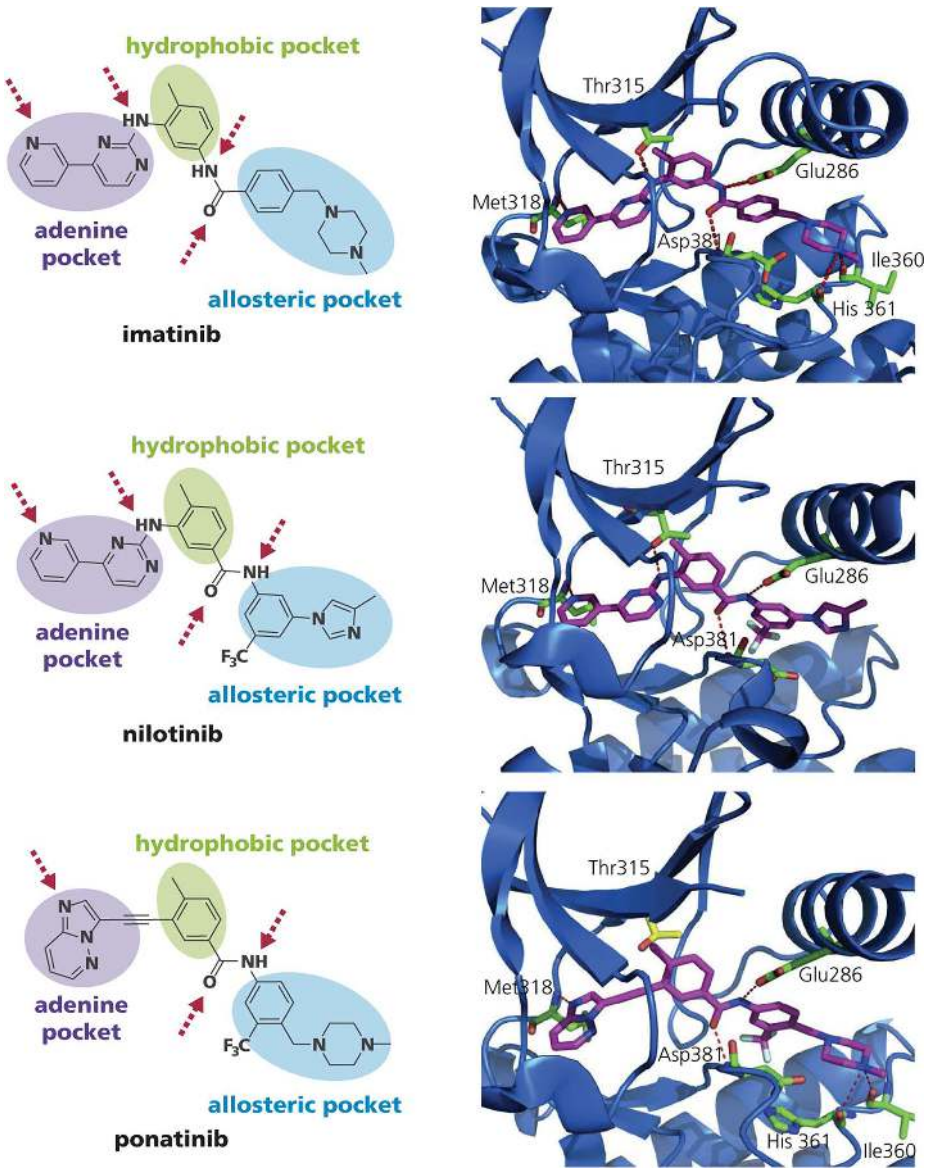


Figure 10.19 X-ray structures of imatinib, nilotinib, and ponatinib showing interactions with Bcr-Abl. All three have components that bind in the adenine, hydrophobic, and allosteric pockets, which are depicted as shaded purple, green, and blue ovals, respectively, in the diagrams on the *left*; all H-bonds are indicated with dashed red arrows. Key interactions common to all three are an H-bond to Met318 in the adenine site, occupation of a hydrophobic pocket by the tolyl group, H-bonds between the amide linker and residues Glu286 and Asp381, and binding by the side chain deep into the newly revealed allosteric site by the aryl-nitrogen heterocycle moiety. The aniline N-H of imatinib and nilotinib make H-bonds to Thr315; this interaction is absent in ponatinib.

Source: Adapted from Wu, P.; Nielsen, T. E.; and Clausen, M. H. (2015). *Trends Pharmacol Sci* 36:422–439. With permission from Elsevier.

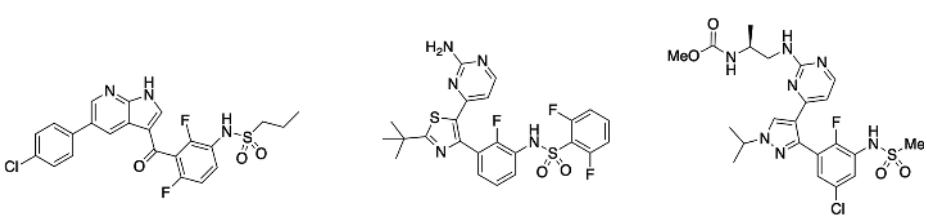


Figure 10.20 Structure of B-Raf inhibitors vemurafenib, dabrafenib, and encorafenib. These inhibit a specific mutant of B-Raf, where a valine is replaced by a glutamic acid (V600E).

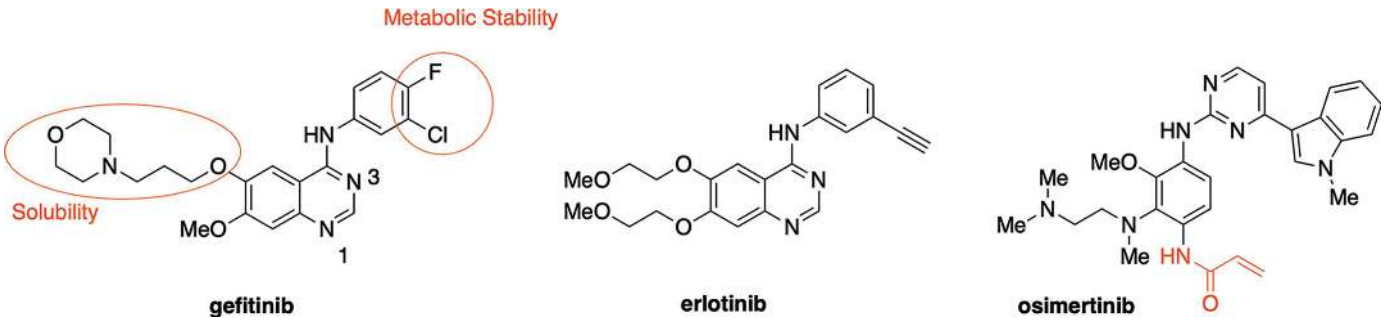


Figure 10.21 Structure of EGFR inhibitors gefitinib, erlotinib, and osimertinib. In both gefitinib and erlotinib, N1 and N3 make H-bonding interactions with the kinase, and the ether side chains (circled in gefitinib) afford improvements in solubility. Halogens on the aryl ring in gefinitib (circled) provide improvements in metabolic stability. Osimertinib is a covalent EGFR inhibitor, with the acrylamide electrophile highlighted in red.

The phosphoinositide 3-kinase/AKT/mammalian target of rapamycin (mTOR) signaling pathway (PAM pathway) is a highly complex network that senses numerous external signals to regulate cell growth and homeostasis. This pathway has evolved to sense the availability of nutrients and put the organism in an anabolic or catabolic state. It therefore plays a role in virtually all cellular functions, and consequently, its dysregulation has been implicated in diseases, such as diabetes, obesity, and cancer. The network is anchored by several different mTOR proteins, which are serine/threonine kinases. The name mTOR (mammalian target of rapamycin) is derived from the natural product rapamycin (sirolimus), an immunosuppressant used in transplant patients, whose target (mTOR) was determined well after its efficacy was discovered. However, there are proteins both upstream and downstream of mTOR itself that are also viable targets in the pathway. Temsirolimus and everolimus are semisynthetic ester and ether analogs of rapamycin used to treat certain cancers. As discussed in Chapter 8, these work through a molecular glue mechanism (Section 8.3.2.2, **Figure 8.15**). A series of small molecules that inhibit the same signaling pathway through binding at the active site of phosphoinositide 3-kinase (PI3K) have also been approved. As one example, idelalisib is an ATP-competitive inhibitor that specifically targets the δ-isoform; it makes H-bonds to the hinge region, as do other kinase inhibitors. However, its propeller shape forces open a new pocket where the small molecule binds. PI3Kδ is in the identical signaling pathway as mTOR but is upstream of it. Others, some of which target different isoforms of PI3K, include leniolisib (δ-isoform), which is approved for an immune condition resulting from a mutation in PI3Kδ and alpelisib (α-isoform-specific) **(Figure 10.22)**. However, serious toxicities have plagued some active site PI3K inhibitors and led to the withdrawal of several of them.

Although compounds like imatinib and vemurafenib were designed to target one oncogenic kinase, subsequent evaluation often finds that other kinases are also inhibited. This characteristic can be a benefit, as these aberrant signaling pathways contain numerous kinases that control cell proliferation and cell survival. Inhibition at multiple nodes may provide multiple points of intervention. Additionally, resistance may be less

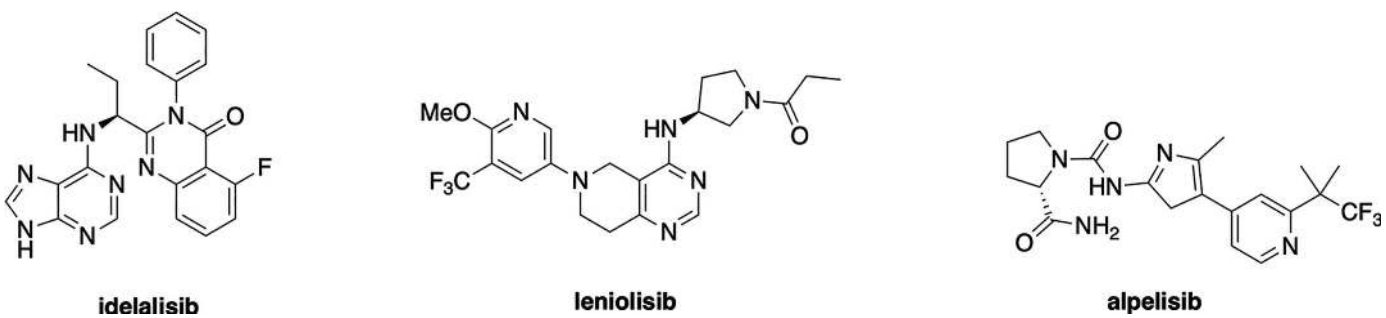
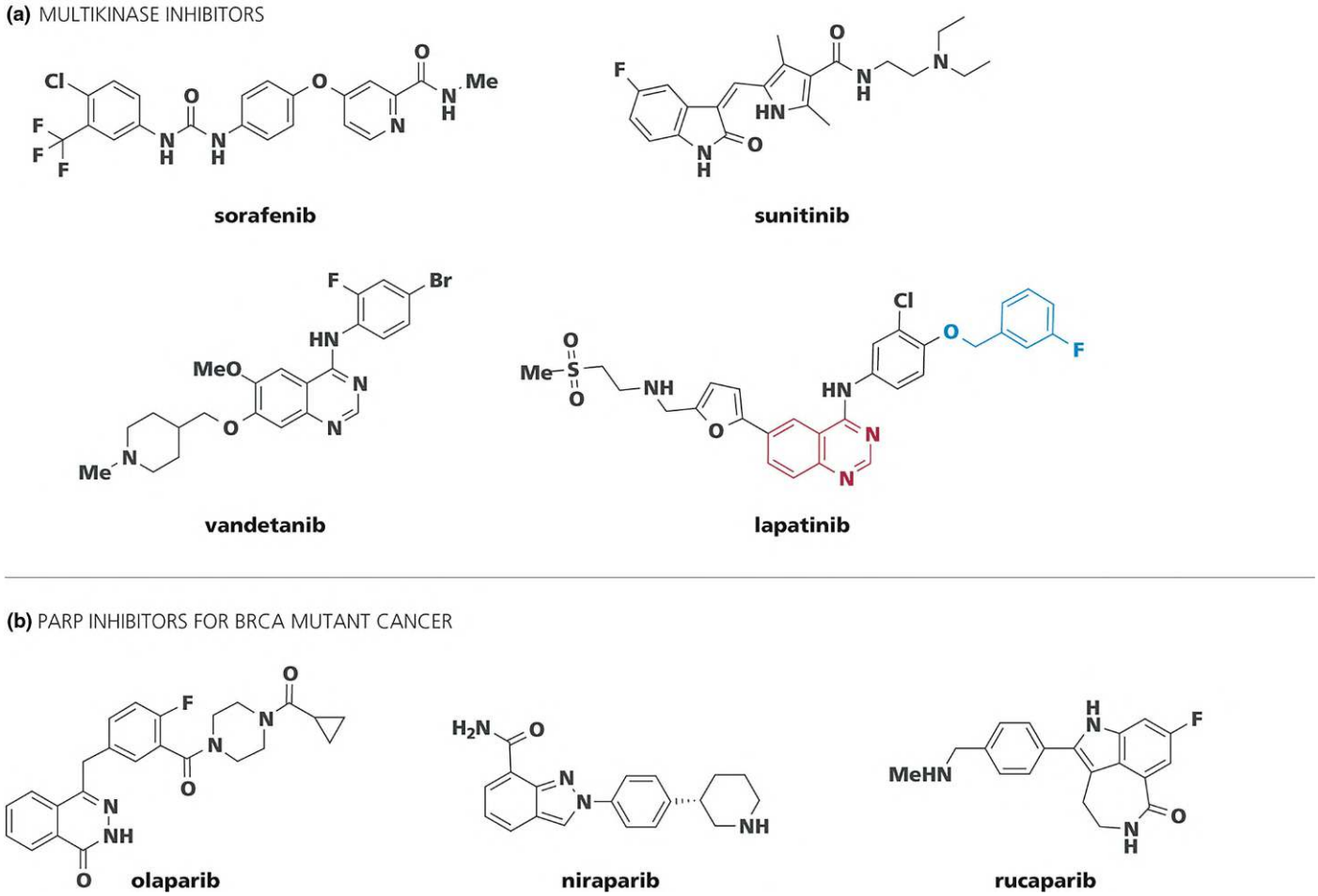


Figure 10.22 Inhibitors of the kinase PI3K. Structures of idelalisib, leniolisib, and alpelisib.



likely to develop with several different targets being inhibited. However, in these cases, a careful balance is required, as kinases are ubiquitous and pan-kinase inhibition is likely to be as toxic as drugs that target DNA or general cell proliferation processes.

Sorafenib (**Figure 10.23a**) was designed to be a selective inhibitor of the serine/threonine kinase B-Raf but was also found to be a tyrosine kinase inhibitor. Sunitinib is a highly promiscuous kinase inhibitor. Vandetanib, a quinazoline, inhibits members of the EGFR, VEGFR, and RET tyrosine kinase families. Unlike the others, lapatinib was designed *a priori* to be a nonselective kinase inhibitor that targets both EGFR and HER2, proteins frequently overexpressed in breast cancer. The starting point for the design of lapatinib was the quinazoline core (shown in red) found in gefitinib and erlotinib, which imparts potent inhibition of EGFR activity. It was found that replacing the fluoro substituent of gefitinib with a benzyloxy group (shown in blue) provided potent HER2 inhibition while retaining inhibition of EGFR. Lapatinib appears to be highly selective for the two kinases it was designed to inhibit. Future efforts are likely to focus on designing multikinase inhibitors with specific profiles.

A different approach to targeting oncogenes and signaling pathways relies on using inhibitors of poly(ADP-ribose) polymerase (PARP). The PARP enzyme is involved in the repair of damaged DNA and helps maintain the integrity of the genome. Its inhibition in normal cells can be compensated for by other repair processes. However, in cancer cells containing certain mutations (for example, *BRCA1* and *BRCA2*), the combination of the mutation and PARP inhibition is lethal. This concept is called **synthetic lethality**, where inhibition of one protein individually does not damage the cell but in combination causes cell death. *BRCA1* and *BRCA2* are tumor suppressor genes that have a role in the activation of DNA repair. Women who inherit these mutated genes are at a significant risk for developing breast and/or ovarian cancer, and a recent focus of drug development aims to target only cells that have mutated *BRCA* genes. The first PARP inhibitor, olaparib, was approved by the FDA in 2014 (see Case Study 1). Two

Figure 10.23 Structures of multikinase inhibitors and PARP inhibitors. (a) Multikinase inhibitors. Sorafenib, sunitinib, vandetanib, and lapatinib are shown. Unlike the others, lapatinib was specifically designed to exhibit multikinase inhibition. Its quinazoline core (shown in red) imparts EGFR potency, and a benzyloxy moiety (shown in blue) is important for HER2 activity. (b) PARP inhibitors for BRCA mutant cancers. PARP is involved in DNA repair, and although not mutated in cancers, its inhibitors are effective in cancers with BRCA mutations. The structures of olaparib, niraparib, and rucaparib are shown.

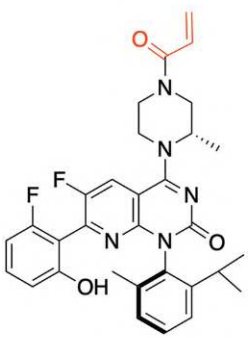
others, niraparib and rucaparib, received approval in 2016 and are being used to treat BRCA mutant breast and ovarian cancers (Figure 10.23b).

10.3.2 RAS Inhibitors Target a Previously "Undruggable" Protein

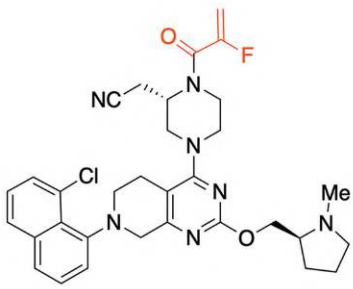
Ras proteins, consisting of isoforms KRAS, HRAS, and NRAS, are small GTPase proteins that tightly regulate cell signaling pathways to direct cell growth, proliferation, and differentiation. They act as binary switches to turn on (when in the GTP form) and off (when in the GDP form) signal transduction. Mutations in *RAS* genes result in the protein being in a persistent "on" state, leading to uncontrolled cell growth and proliferation and, eventually, cancer. It has long been known that RAS mutations are the most frequently mutated oncogenes and drive ~25% of all cancers, and among certain cancers (e.g., non-small-cell-lung cancer [NSCLC], colorectal and pancreatic duct adenocarcinoma), they are believed to be solely responsible. Despite this known link, until recently, mutant RAS was called an "undruggable" target, given that despite decades of effort, little progress was made toward developing small-molecule inhibitors. This lack of success was rationalized as due to a lack of druggable binding sites and its high affinity for GTP, which makes it difficult to develop competitive inhibitors. Furthermore, it was challenging to design selectivity toward mutant Ras and away from wild-type Ras, which was essential for safety.

A breakthrough came in the early 2010s, based on work from the Shokat lab at UCSF. A common mutation is $KRAS^{G12C}$, where a cysteine residue replaces a glycine, resulting in a shift in the equilibrium toward active GTP-bound (RAS-on) state. Fragment-based screening of a library of disulfides led to the identification of a small fragment that selectively bound to $KRAS^{G12C}$, via disulfide formation with the mutant Cys residue at an allosteric site. The disulfide did not react with wild-type KRAS or with other cysteines present in the protein. This work served as the inspiration for the development of new cancer drugs that target RAS. Sotorasib and adagrasib are both allosteric inhibitors that bind in a pocket in the inactive ("off") GDP-bound form of $KRAS^{G12C}$ near the G12C mutation and form a covalent bond to Cys12 (Figure 10.24).

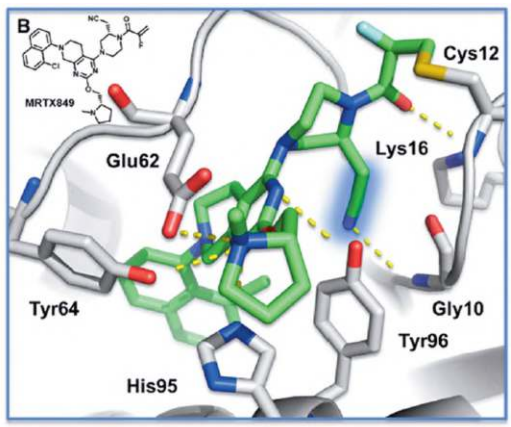
Since 2022, adagrasib and sodorasib have been used to treat NSCLC and colorectal cancer, but resistance develops by mechanisms that increase levels of the GTP-bound form of RAS ("on" form) and reactivate RAS signaling. Other approaches include non-covalent inhibitors, inhibitors that are effective against multiple Ras mutants beyond G12C, and small molecules that target both the "on" and "off" states of RAS. One innovative approach targets the protein–protein interaction between cyclophilin A and GTP-bound RAS via a molecular glue mechanism. These compounds first bind cyclophilin A (CypA), forming a binary complex that binds active RAS proteins with high affinity, ultimately creating a CypA-drug–RAS(GTP) tri-complex that



sotorasib (AMG 510)



adagrasib (MRTX849)



adagrasib X-ray in allosteric pocket of $KRAS^{G12C}$

Figure 10.24 Structures of sotorasib and adagrasib, with X-ray of adagrasib bound to $KRAS^{G12C}$, showing covalent bond to Cys12 (upper right, yellow).

Source: With permission from Chen, H.; Smaill, J. B.; Liu, T.; Ding, K.; and Lu, X. (2020) "Small-molecule inhibitors directly targeting KRAS as anticancer therapeutics." *J. Med. Chem*. 63, 14,404–14,424.

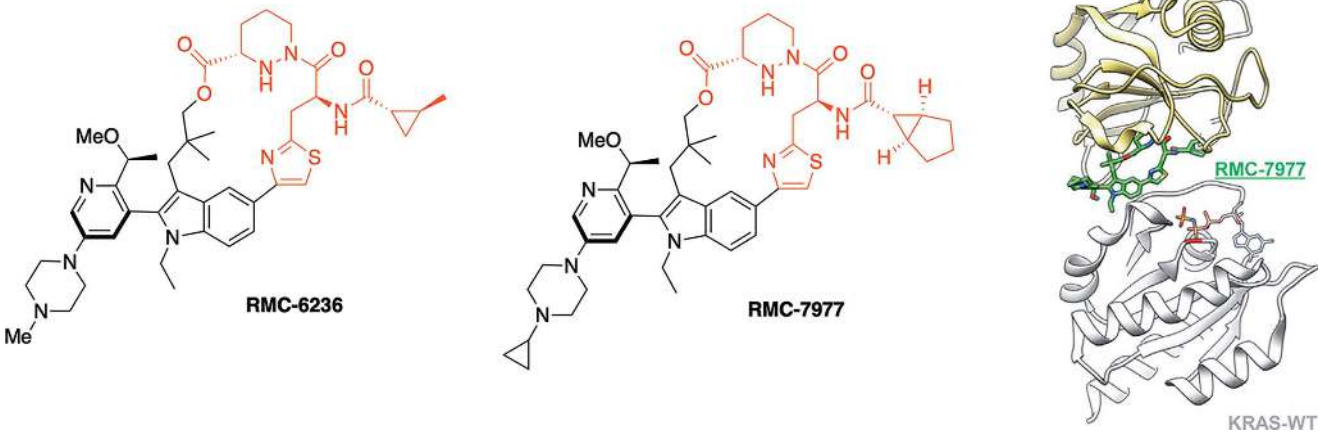


Figure 10.25 Structures of RMC-6236 and RMC-7977, showing tri-complex formation between RMC-7977, cyclophilin A and KRAS (PDB: 8TBF). Regions in red bind to cyclophilin A.

Source: With permission from Su, W., and Hou, X. (2024). *Pharmaceutical Science Advances 2*, 100047.

sterically hinders RAS–effector interactions and impedes downstream oncogenic signaling. Two of these newer RAS modulators, RMC-6236 and RMC-7977, are shown in Figure 10.25, as is an X-ray structure showing the tri-complex.

The drug discovery approach to compounds such as RMC-6236 and RMC-7977 was initially based on sanglifehrin A, a polyketide natural product from *Streptomyces* that binds to CypA. Maintaining the CypA-binding region (red in Figure 10.25), and introducing a macrocycle for conformational constraint, led to the clinical candidates, with promising preclinical results for several types of cancer, including pancreatic duct adenocarcinoma. RMC-6236 has advanced to clinical trials.

10.4 INHIBITION OF ANGIOGENESIS

Angiogenesis, the formation of new blood vessels from existing ones, is critical to support tumor growth, survival, and metastasis. Once they reach a certain size, tumors require a dedicated blood supply to support their rapid and unregulated growth. The idea of inhibiting angiogenesis as a mechanism to control tumor growth and metastasis was proposed by Judah Folkman in the early 1970s. One of the ways in which angiogenesis is mediated is by the action of peptide growth factors, such as vascular endothelial growth factor (VEGF) and basic fibroblast growth factor (bFGF). Binding of VEGF to its receptor, VEGFR, induces a series of conformational changes that results in the activation of its kinase domain. Phosphorylation of various substrates results in the initiation of a signal cascade that activates a number of other enzymatic pathways, many of which may be involved in cancer progression. The enzymatic kinase domain of VEGFR has been targeted with ATP-competitive inhibitors. This approach does not target the tumor itself but the environment it needs to survive and grow. Despite inhibiting a physiological process, VEGFR inhibitors usually exhibit fewer side effects than conventional chemotherapeutics. Currently, a number of marketed kinase inhibitors work through the inhibition of VEGFR; however, they all also inhibit at least a few other kinases, and those effects likely also contribute to their anti-cancer efficacy. Sorafenib and sunitinib (Figure 10.23) inhibit the kinase domain of VEGFR, among the other kinases. The efficacy of pazopanib and axitinib has been attributed primarily to the inhibition of VEGFR (Figure 10.26).

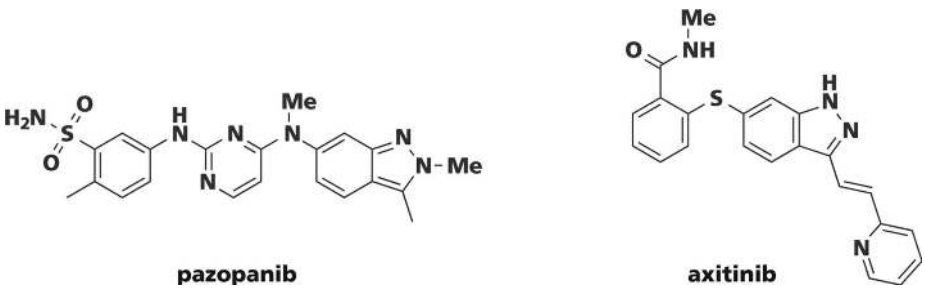


Figure 10.26 Structure of VEGFR inhibitors pazopanib and axitinib. Both drugs are angiogenesis inhibitors that act by inhibiting the kinase activity of the vascular endothelial growth factor receptor (VEGFR).



Figure 10.27 Structures of thalidomide, lenalidomide, and pomalidomide. Lenalidomide and pomalidomide are analogs of thalidomide, and all are angiogenesis inhibitors.

Another example of an angiogenesis inhibitor is thalidomide (Figure 10.27), which has a notorious past. It was marketed in Europe in the late 1950s to treat, among other disorders, morning sickness. Sadly, it caused severe limb abnormalities and significant mortality in the children of mothers who took thalidomide during pregnancy. This tragic example is credited with the development of more stringent regulations in government drug approval processes (see Chapter 5). However, thalidomide remained in use to treat inflammatory complications from leprosy. Interest in the effects of thalidomide resurged in the late 1990s, with reports of activity against multiple myeloma and other blood cancers, as well as inflammatory diseases. In 2006, it was approved with strict controls in combination with another drug to treat multiple myeloma, and later analogs include lenalidomide and pomalidomide (see Figure 10.27).

These compounds are members of a family of drugs called IMiDs (immunomodulatory drugs), and they have both anti-inflammatory and anti-angiogenesis effects. Thalidomide and its analogs potently inhibit the growth of new blood vessels; however, the exact mechanism of action is not well understood. Recently, evidence suggests thalidomide's anti-cancer activity is likely due to multiple effects, including binding to the protein cereblon, which affects protein homeostasis and is part of the E3 ligase complex. As described in Chapter 8, these are also common E3 ligase binding elements in the design of PROTACS.

10.5 DRUGS TARGETING THE UBIQUITIN–PROTEASOME PATHWAY

The proteasome is a multiprotein complex present in all eukaryotic cells. Its function is to degrade proteins that have been tagged via the addition of small proteins called ubiquitin; it is part of a highly complex metabolic process called protein homeostasis that is essential for maintaining cell viability. Proteasomal degradation is one mechanism whereby physiological processes, such as cell growth and gene expression, are controlled and regulated. The inhibition of proteasome activity, however, affects normal cells and transformed cells differently. Normal cells usually undergo cell cycle arrest upon proteasome inhibition, whereas cancer cells typically respond by undergoing apoptosis. Exactly why this difference occurs is still being investigated; however, this difference has been exploited by developing proteasome inhibitors as cancer therapeutics.

The proteasome is a cylindrical structure with a catalytic core, where amide hydrolysis is carried out, and a regulatory site that caps both ends of the core and allows recognition of tagged poly-ubiquinated protein substrates (Figure 10.28).

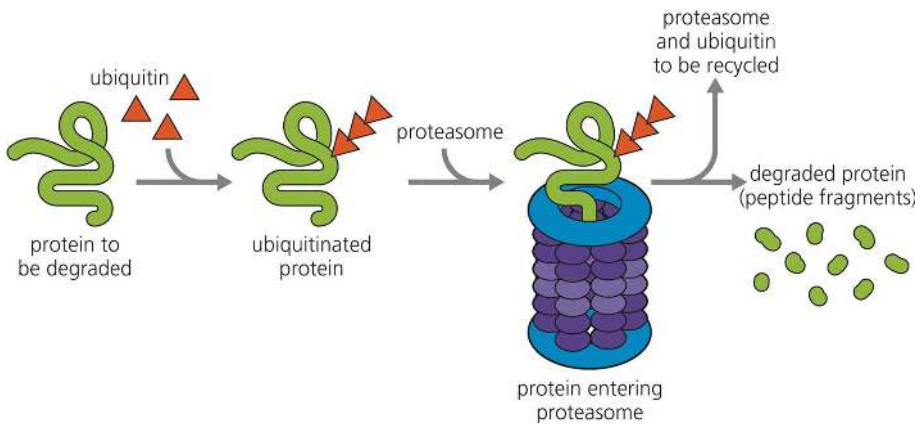


Figure 10.28 Protein degradation by 26S proteasome complex. A protein targeted for degradation (shown in green) is ubiquitinated (ubiquitin, a small protein, is depicted as orange triangles) and then enters the proteasome (shown as a blue-and-purple cylinder). De-ubiquitination and then degradation in the core by sequential cleavage of amide bonds releases the protein fragments.

Source: Adapted from McDonald, R. B. (2013). *Biology of Aging*. Garland Science.

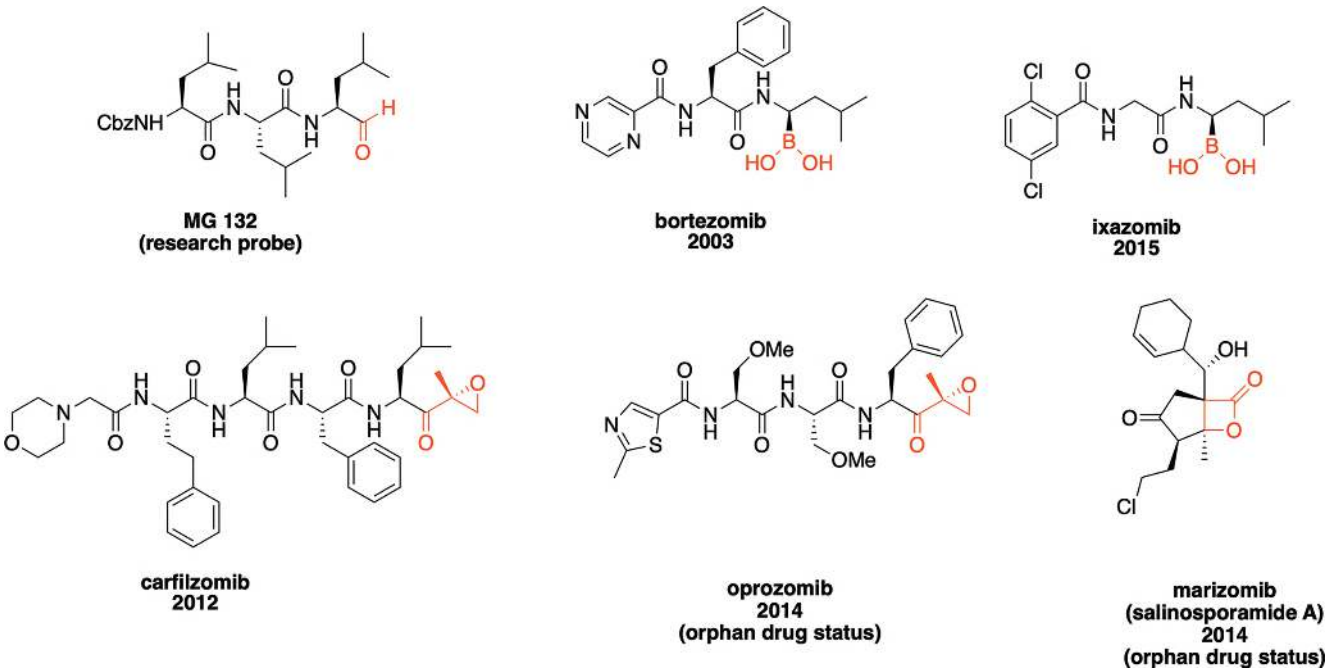


Figure 10.29 Structures of proteasome inhibitors. The warhead portion of each is shown in red. MG132, bortezomib, and ixazomib are reversible inhibitors; carfilzomib, oprozomib, and marizomib are irreversible inhibitors.

The protein substrate is deubiquitinated and enters the core of the proteasome, where it is cleaved, and the resultant peptide fragments are released at the opposite end. Once inside the proteasome, proteins are sequentially processed by multiple proteolytically active sites that are related to serine proteases but contain threonine at the active site (see Chapter 7.4) to eventually generate small peptide fragments with an average length of six amino acids.

The different proteolytic sites within the proteasome exhibit varying substrate specificities, and inhibitors can be generated by applying serine/threonine protease inhibitor design strategies. This approach involves the use of a peptidic moiety that binds to the substrate binding site and the incorporation of a reactive warhead as described in Section 7.4.1 to interact with the active site nucleophile. Aldehydes, vinylsulfones, epoxyketones, and β-lactone warheads have all been incorporated; compounds such as the aldehyde MG132 (**Figure 10.29**) are frequently used as research probes to study the proteasome. In fact, the first proteasome inhibitor approved, bortezomib, was originally used as a probe, but observations of treated cancer cells undergoing apoptosis led to its investigation as an anti-cancer drug.

Bortezomib, which contains the unconventional boronic acid warhead, is a reversible inhibitor of the enzyme (see **Figure 10.29**). Another boronic acid, ixazomib, was approved in 2015 for multiple myeloma and is dosed orally. Carfilzomib, the second proteasome inhibitor approved, is based on the natural product epoxomicin and possesses an electrophilic epoxyketone that irreversibly inhibits the enzyme. A similar drug, oprozomib, was granted orphan drug status in 2014 to treat multiple myeloma. The natural product salinosporamide A (marizomib), also granted orphan drug status, has a β-lactone that reacts with the active site threonine. **Figure 10.30** shows the interactions seen in X-ray co-crystal structures of bortezomib bound to two of the proteolytic sites in the proteasome. Since the proteasome contains multiple active sites, it is likely that the inhibitor occupies several of them simultaneously. The formation of a covalent bond between the threonine hydroxyl and the boronic acid moiety generates a negatively charged boronate tetrahedral intermediate that closely mimics the transition state. The boronate formed is further stabilized by H-bonds to Gly and Thr residues. The peptide residues of the inhibitor make a series of H-bonds to both backbone and side chain carbonyls and amines.

Boronic acids are particularly effective proteasome inhibitors and form a tetrahedral intermediate with the active site threonine that is very slowly reversible. They

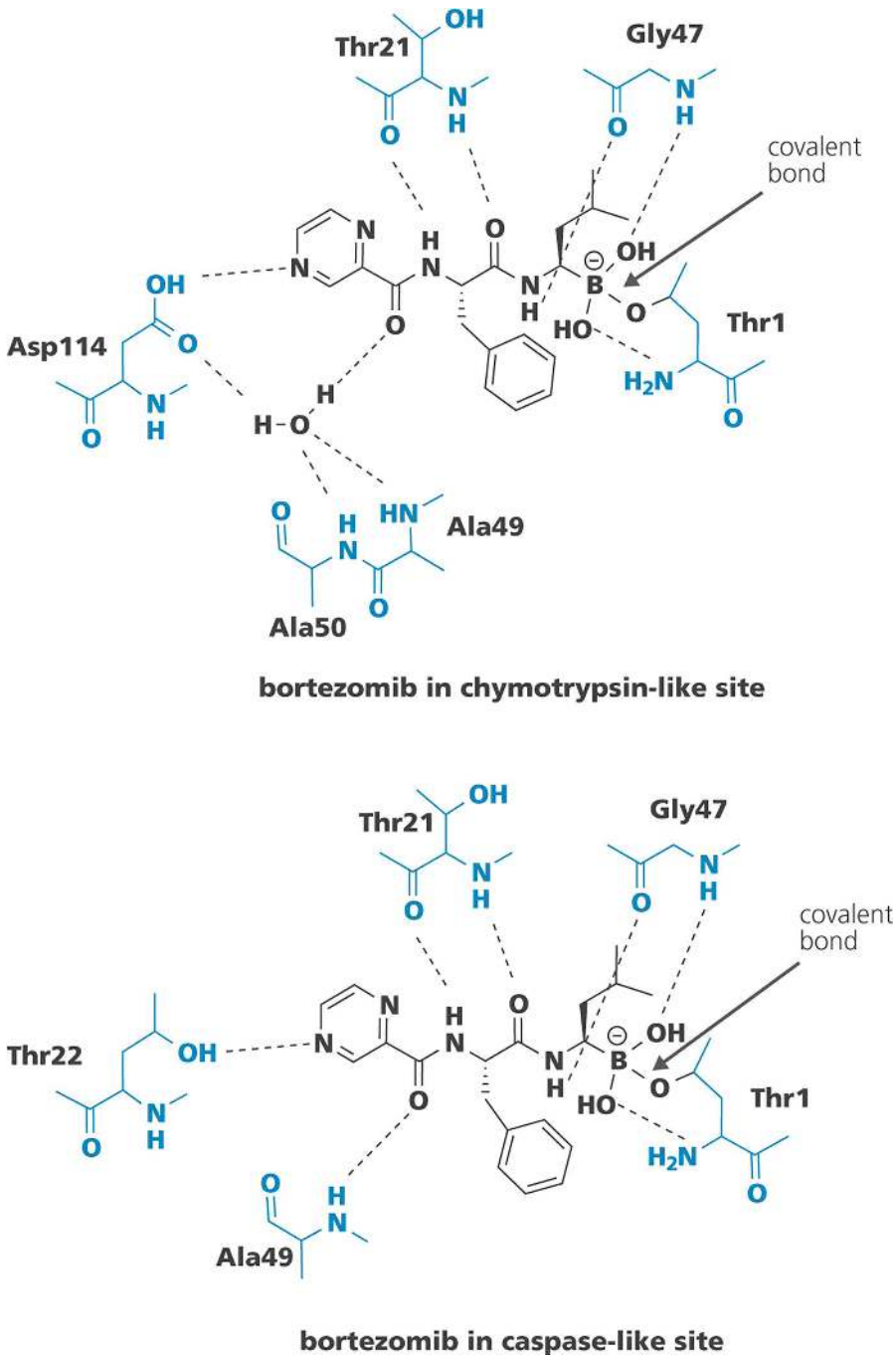


Figure 10.30 Interactions from co-crystal structure of bortezomib bound to two proteasome proteolytic sites. *Top*, interactions of bortezomib bound to chymotrypsin-like site. *Bottom*, interactions of bortezomib bound to caspase-like site. Both structures show a covalent bond between the threonine1 oxygen and boron atom (arrow), as well as key H-bonds between backbone and side chain residues (blue) and bortezomib.

Source: From Groll, M.; Berkers, C. R.; Ploegh, H. L.; and Ovaa, H. (2006). *Structure* 14:451–456. With permission from Elsevier.

appear to be much more selective for serine/threonine proteases over cysteine proteases. In contrast, aldehyde and vinylsulfone warheads inhibit both serine/threonine and cysteine proteases. The stronger interaction between boron and oxygen, compared to that between boron and sulfur, is likely the root of this selectivity. Other attributes of boronic acids include their resistance to metabolic oxidation and efflux.

As described in Chapter 8, PROTACS make use of the ubiquitin–proteasome pathway by directing a specific protein to be ubiquitinated and then guided to the proteasome for degradation. Proteins specifically targeted that are involved in cancers include the estrogen receptor and the androgen receptor, with PROTACS that have advanced to late clinical trials to treat breast cancer and prostate cancer. Examples of other PROTACS targeted for cancer include Bruton's tyrosine kinase (BTK, B cell malignancies), BCL-xl (B cell lymphoma), bromodomain proteins BRD4 (AML), EGFR (lung cancer), HER2 (breast cancer), and BCL-2 (AML)

10.6 DRUGS TARGETING EPIGENETIC PROCESSES

The term **epigenetics** refers to the carefully controlled chemical reactions that occur on the genome and activate or deactivate gene expression without changing the DNA sequence. These reactions modify DNA itself, as well as the proteins that are involved in compacting DNA in the nucleus, and are heritable. Epigenetic modifications are an important aspect of normal physiological processes that allow control of gene expression; however, an individual's epigenetic profile is dynamic. It will change with age and can be influenced by diet and exposure to toxins, among other external factors. Aberrant epigenetic processes may lead to cancer and autoimmune disorders, as well as other conditions.

In the cell, DNA is packaged in nucleosomes. These are densely packed particles containing DNA and eight units of a small, positively charged protein called histone, which serves to protect and compact the DNA. The DNA chain wraps around the histone octamer 1.65 times, and the histone tails extend from the complex, as seen in **Figure 10.31**. The tails typically end in lysine residues; in addition to DNA, these are also sites for modification. When the histone is modified, the process, called chromatin remodeling, exposes the DNA and allows access to the transcription machinery. Some enzymes responsible for these processes are considered epigenetic "writers" and make a covalent modification. These include enzymes that transfer acetyl groups (histone acetyltransferase, HAT) and methyl groups (histone methyltransferase, HMT; protein arginine methyltransferase, PRMT), as well as kinases that phosphorylate. Histone deacetylase (HDAC) and histone lysine demethylase (KDM) are epigenetic "erasers" and remove modifications. Other proteins contain specific recognition domains and include bromodomain, chromodomain, and Tudor domain proteins, which are epigenetic "readers"; they interact with the modified histones or DNA, resulting in gene transcription.

When a normal cell becomes transformed into a cancerous cell, some of the most significant changes are epigenetic ones. Differences in DNA methylation patterns and histone modifications that silence tumor suppressor genes and activate oncogenes are observed. In the context of cancer drug discovery, interfering with the processes of methylation and deacetylation are validated strategies, with others under active investigation.

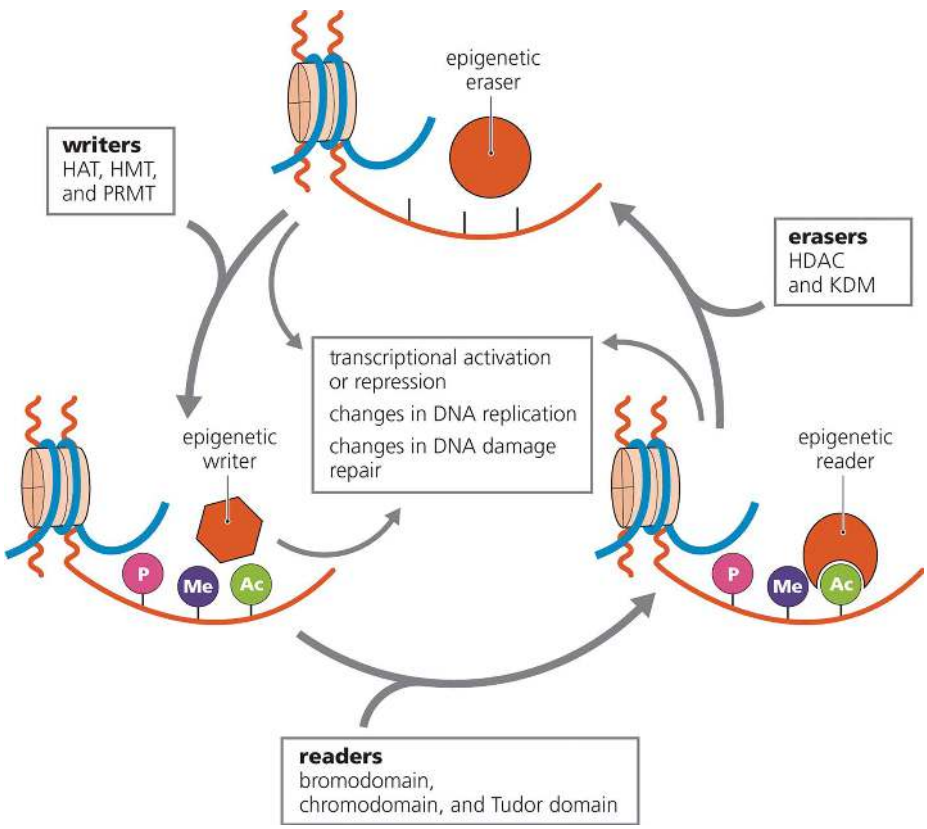


Figure 10.31 Summary of chromatin remodeling. DNA (shown in blue) is wound around histones (orange cylinder) that contain extended tails (shown in orange). Reactions catalyzed by enzymes modify the histone tails to regulate transcriptional activation or repression, DNA replication, and DNA damage repair. Specific reactions of epigenetic writers (HMT, HAT, and PRMT) include methylation (Me), acetylation (Ac), and phosphorylation (P). These changes may be recognized by proteins (such as those containing bromodomains, chromodomains, and Tudor domains) that are epigenetic readers. Epigenetic modifications can be removed by epigenetic erasers, such as HDAC and KDM.

Source: Adapted from Falkenberg, K. J., and Johnstone, R. W. (2014). *Nat Rev Drug Discov* 13:673–691. With permission from Macmillan Publishers Ltd.

10.6.1 Nucleosides with Modified Bases Can Act as DNA Methyltransferase Inhibitors

DNA methyltransferase is an epigenetic writer that modifies the 5-carbon of cytosine at a CG sequence on the DNA backbone. The enzyme uses *S*-adenosylmethionine as a coenzyme and methyl source. As described earlier, the specific methylation pattern of DNA controls gene expression patterns, and aberrant DNA methylation has been strongly correlated with a number of human cancers. 5-azacytidine and decitabine (**Figure 10.32**) are nucleoside analogs that contain the same unnatural base, which is the basis of their activity.

Like most other nucleosides, these analogs are first activated to their corresponding triphosphates. They become incorporated into the host DNA and are substrates for the DNA methyltransferase enzyme. As shown in **Figure 10.33a**, under typical conditions, the cytosine bases in DNA become methylated via methyltransferase enzymes.

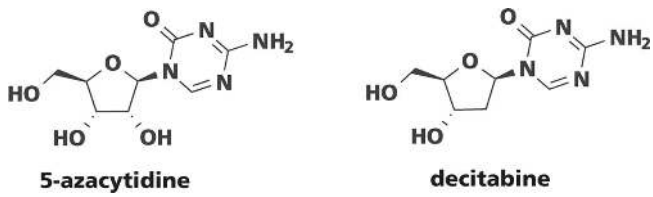
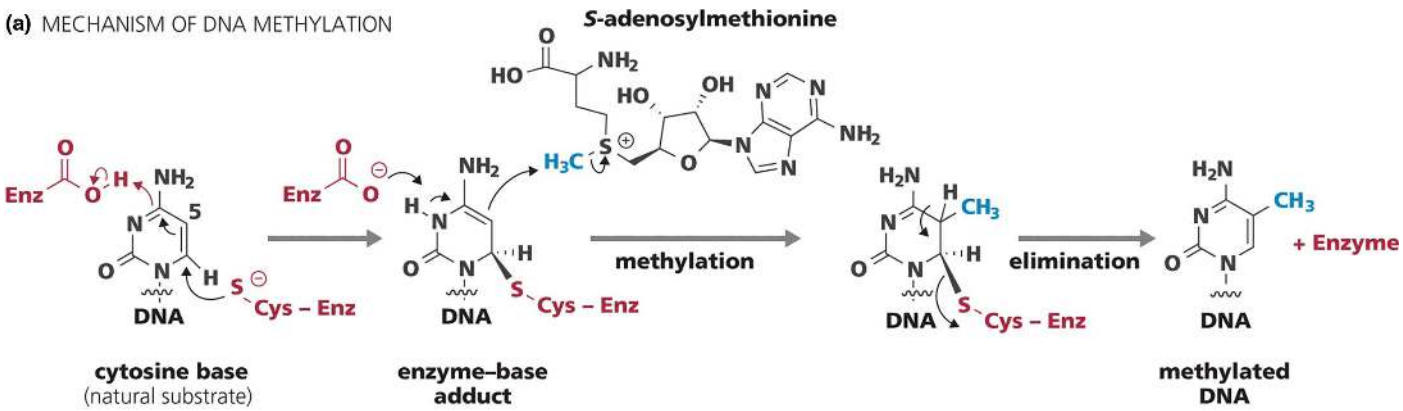


Figure 10.32 Structures of 5-azacytidine and decitabine. Both compounds contain the same unnatural base, which is the basis of the mechanism of inhibition.



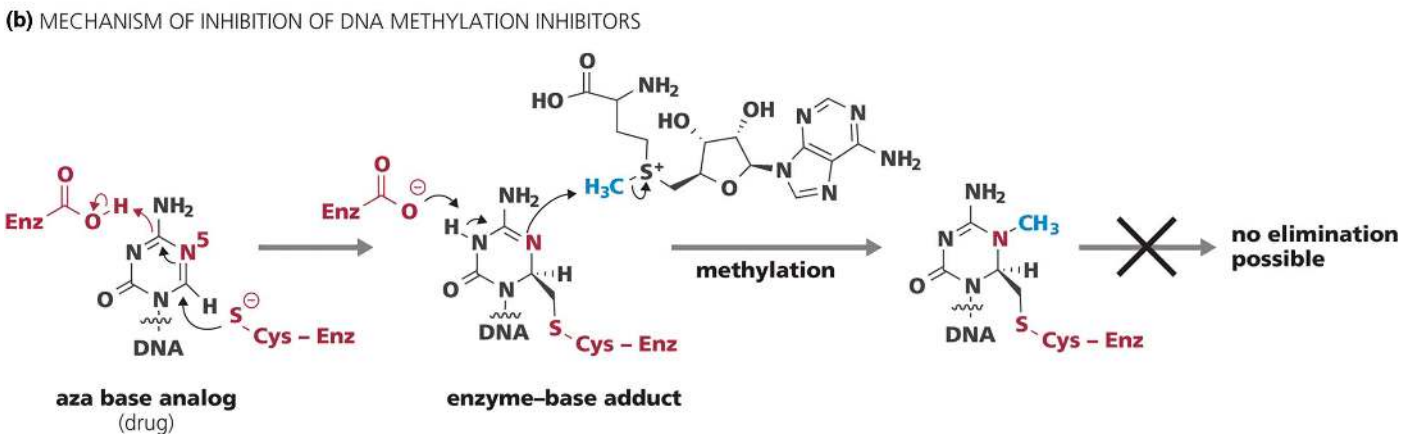


Figure 10.33 Comparison of the mechanism of DNA methyltransferase with natural substrate and aza analog. (a) Mechanism of DNA methylation. A cysteine from the enzyme (red) reacts at the 6-position of the base to generate a reactive enzyme–base adduct. The intermediate reacts with the co-substrate, *S*-adenosylmethionine, to introduce a methyl group at the 5-position. A 1,2-elimination regenerates the enzyme and releases the new methylated base. (b) Mechanism of inhibitor of DNA methylation. The unnatural substrate undergoes the same initial reactions to generate the enzyme–base adduct; however, the 1,2-elimination is not possible, and the enzyme is irreversibly bound to the inhibitor–DNA complex.

The addition of a cysteine residue of the enzyme to the 6-position of the base generates an enzyme–base complex. This activated intermediate reacts with the enzyme's co-substrate, *S*-adenosylmethionine, to transfer a methyl group to the 5-position. In the case of the natural substrate, elimination of the C-5 hydrogen takes place with the enzyme's thiolate as a leaving group to regenerate the active enzyme and the newly methylated DNA. The same initial steps occur when the aza base analog is the substrate (see **Figure 10.33b**); however, the final elimination step is not possible. The inhibitor maintains a stable covalent bond with the enzyme, rendering the enzyme ineffective; the damaged DNA is recognized, and signals for the cell's destruction are initiated. This mechanism of inhibition bears some resemblance to inhibitors of thymidylate synthase described in **Figure 10.12**.

10.6.2 Histone Deacetylase Inhibitors Bind to Zinc in the Enzyme's Active Site

Like DNA, histones are modified by the addition or removal of a small group, such as acetate, phosphate, or methyl. Gene transcription is regulated by a carefully regulated series of acetylations and deacetylations. Acetylation of the lysine-rich amino-terminal tails of histones is carried out by histone acetyltransferases (HATs), and histone deacetylases (HDACs) catalyze the removal of the same acetate groups (see **Figure 10.31**). Both enzymes work in concert to appropriately configure the DNA–chromatin complex so that it is competent for transcription. The inhibition of HDACs disrupts the equilibrium, resulting in changes in gene expression and anti-proliferative effects. In addition, HDACs have non-histone substrates, and inhibiting those processes may also contribute to the anti-cancer efficacy. In fact, the first inhibitors of HDACs were discovered due to their potent anti-cancer effects, and their mechanism was only identified later. Currently, there are several approved HDAC inhibitors that are used to treat cancer; those that act on HATs are an active area of research, but inhibitors have not yet been approved.

The mechanism of HDACs is related to that of zinc metalloproteases (see Section 7.4.1), in that a zinc cation at the active site is required for the activation and hydrolysis of the acetate. The active site consists of a long tunnel whose terminus contains a catalytic Zn^{2+}, which is coordinated to a series of aspartic acid and histidine residues that participate in the hydrolysis process. The substrate (or inhibitor) occupies the tunnel and coordinates to the Zn^{2+}. The inhibitor vorinostat (also known as suberanilohydroxamic acid, or SAHA) is shown bound to HDAC in **Figure 10.34**.

As of 2025, four HDAC inhibitors, vorinostat, belinostat, panobinostat, and romidepsin (**Figure 10.35**), have been approved to treat cancer. Most conform to a common pharmacophore consisting of a cap structure (shown in red), a linker region (depicted in blue), and a zinc-binding group (shown in green). In vorinostat, belinostat, and panobinostat, the zinc-binding group is a hydroxamic acid. In romidepsin, a disulfide (highlighted in pink) is cleaved *in vivo* to generate the zinc-binding thiol. While there is a large family of HDACs, most of the currently approved inhibitors have

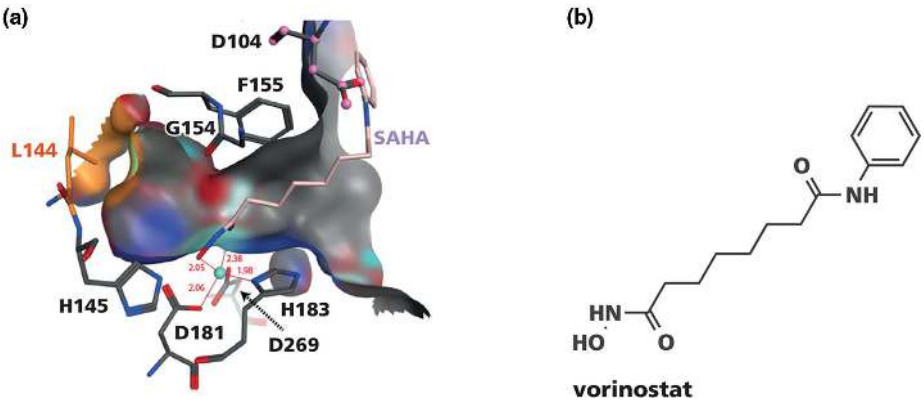


Figure 10.34 Vorinostat binding to HDAC. (a) X-ray structure of vorinostat (SAHA) bound to histone deacetylase-like protein (HDAC2), showing the long narrow enzyme binding pocket and the catalytic zinc ion (light blue) at the active site. The zinc ion coordinates to aspartates 181 and 269 (D181 and D269) and histidine 183 (H183) and to the hydroxamide of vorinostat. (b) Structure of vorinostat.

Source: (a) From Lauffer, B. E. L.; Mintzer, R.; Fong, R.; et al. (2013). *J Biol Chem* 288:26,926–26,943. With permission from the American Society for Biochemistry and Molecular Biology.

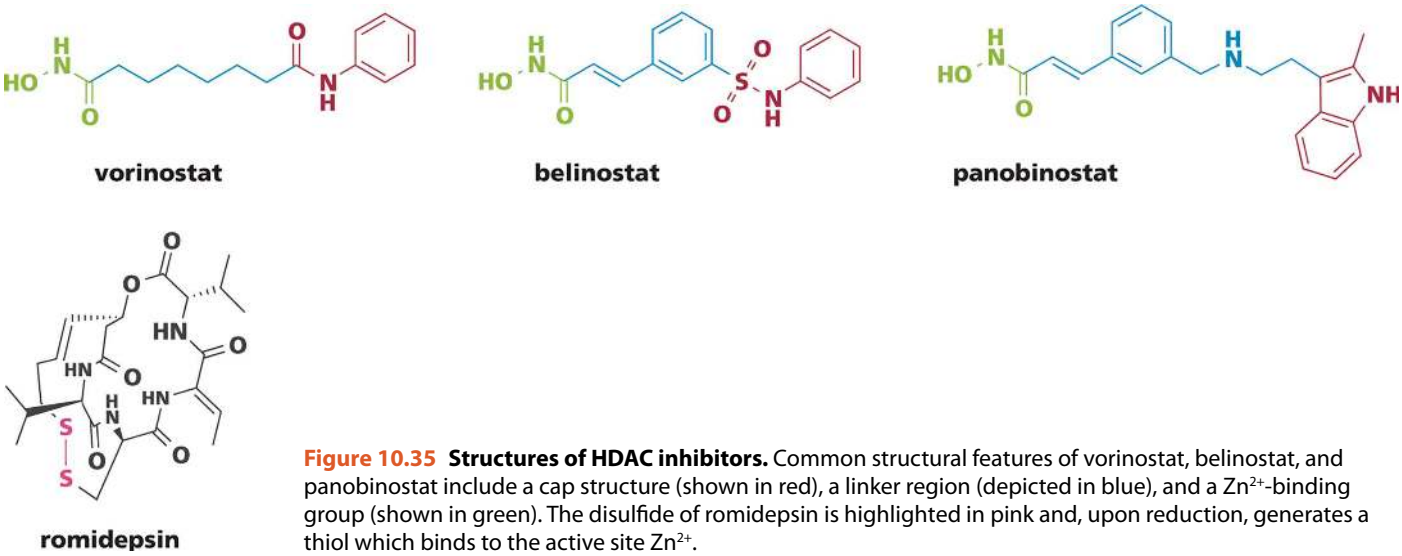


Figure 10.35 Structures of HDAC inhibitors. Common structural features of vorinostat, belinostat, and panobinostat include a cap structure (shown in red), a linker region (depicted in blue), and a Zn^{2+}-binding group (shown in green). The disulfide of romidepsin is highlighted in pink and, upon reduction, generates a thiol which binds to the active site Zn^{2+}.

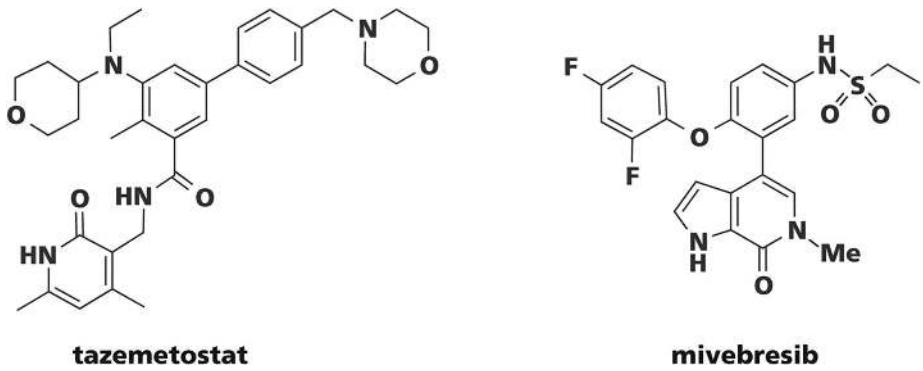


Figure 10.36 Structures of drugs targeting other epigenetic processes. Tazemetostat inhibits the histone methyltransferase EZH2, and mivebresib prevents binding between a bromodomain protein and acetylated histones.

little selectivity for one class over another. The discovery of vorinostat, as well as other HDAC inhibitors, is discussed further in Case Study 1.

While only a few drugs targeting epigenetic processes are approved, work on other epigenetic targets beyond HDACs have also led to drugs and candidates in clinical trials. These include tazemetostat, an inhibitor of the histone methyltransferase EZH2 (an epigenetic writer), approved in 2020, and mivebresib, an inhibitor of the PPI between bromodomain proteins (epigenetic readers) and acetylated histones that advanced to phase II clinical trials (**Figure 10.36**).

10.7 DRUGS TARGETING HORMONE-DEPENDENT TUMORS

In addition to the aforementioned aberrant pathways that can drive the malignant process, some tumors require the presence of hormones to proliferate. In these cases, hormone-dependent nuclear receptors (described in Section 6.4) that control gene transcription are dysregulated. For example, some breast cancers grow in response to the female hormone estrogen. When breast cancer patients are first diagnosed, their tumors are categorized by whether they are estrogen receptor positive (ER+) or negative (ER–), to determine whether their growth is responsive to estrogen. This information dictates specific therapy and, often, prognosis.

In the cases of hormone-responsive tumors, one aspect of therapy usually involves reducing or eliminating the amount of receptor signaling. Therapeutically, this is

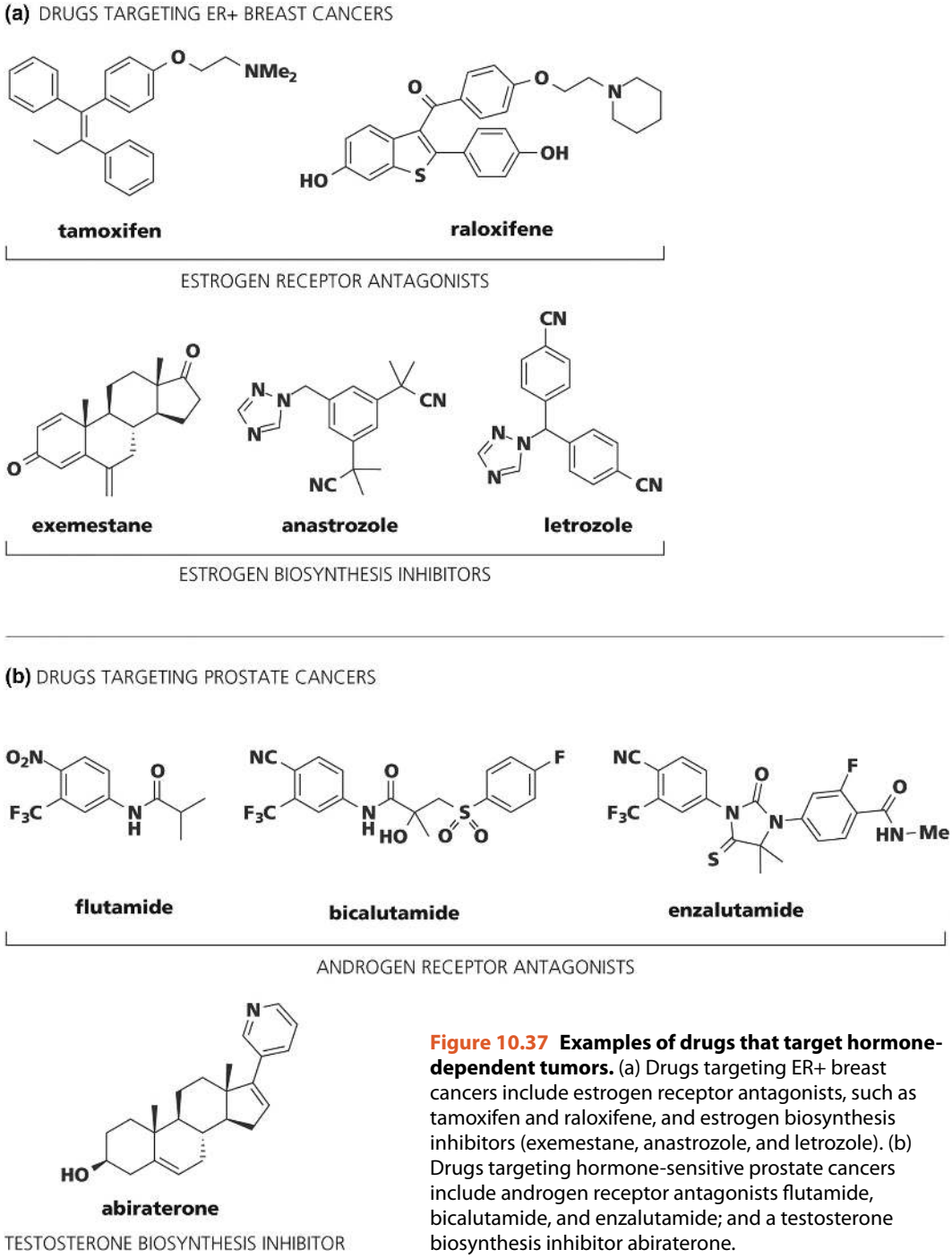


Figure 10.37 Examples of drugs that target hormone-dependent tumors. (a) Drugs targeting ER+ breast cancers include estrogen receptor antagonists, such as tamoxifen and raloxifene, and estrogen biosynthesis inhibitors (exemestane, anastrozole, and letrozole). (b) Drugs targeting hormone-sensitive prostate cancers include androgen receptor antagonists flutamide, bicalutamide, and enzalutamide; and a testosterone biosynthesis inhibitor abiraterone.

accomplished in two ways: using a receptor antagonist or inhibiting the biosynthesis of the natural ligand. Figure 10.37 illustrates a few examples of synthetic agents used as anti-cancer agents, their mechanisms of action, and their structures.

Anti-estrogens, such as tamoxifen and raloxifene (see Section 6.4.1), are antagonists of the estrogen receptor and are frequent components of ER+ breast cancer therapy regimens. Other options include reducing amounts of estrogen through the use of biosynthesis inhibitors, such as the aromatase inhibitors exemestane, anastrozole, and letrozole. Many prostate cancers respond to similar strategies. For example, abiraterone inhibits biosynthetic pathways to testosterone. Flutamide, bicalutamide, and enzalutamide act as androgen receptor antagonists, blocking the effects

of testosterone in prostate cancer. The first examples of PROTACs in clinical trials were those that target the estrogen (ARV-471) and androgen receptor (ARV-110) for cancer treatment (see Chapter 8).

10.8 SUMMARY

Just as there are many biological mechanisms implicated in the development of cancer, there are also many targets for drugs to treat the disease. As of 2025, the US National Cancer Institute lists over 200 chemotherapy agents, and many more if multiple formulations of the same drug or combinations are counted. This number gives some idea of the diversity of compounds that are effective anti-cancer agents. The first chemotherapy drugs were discovered in the 1940s, and until the introduction of kinase inhibitors in the 2000s, many drugs targeted DNA, either directly or by affecting enzymes involved in DNA synthesis. Other early drugs targeted proteins, such as tubulin, that were intimately involved in cell division. The toxicity of these older drugs, due to their limited selectivity, has been mediated more recently by delivery systems, including tumor-directed antibodies, as well as improved formulations. Newer drugs target enzymes and processes that are selective or specific to tumors and lead the way for personalized medicine, where a careful analysis of the patient and the tumor determines the most effective therapy. These small-molecule therapeutic agents include kinase inhibitors, RAS inhibitors, PARP inhibitors, proteasome inhibitors, and drugs that target epigenetic processes, as well as drugs targeting hormone-specific tumors. While it is impossible to cover all anti-tumor drugs in a single chapter, examples of the major classes of drugs and their mechanisms are found here, as well as in Chapter 9.

CASE STUDY 1 DISCOVERY OF VORINOSTAT

Vorinostat, also known as suberanilohydroxamic acid (SAHA), was developed at Columbia University by Ronald Breslow and Paul Marks based on an observation by Charlotte Friend that cells cultured in high concentrations of the common laboratory solvent dimethyl sulfoxide (DMSO) underwent apoptosis. The examination of other small-molecule polar solvents identified amides, such as dimethylformamide (DMF), as having a similar effect. A series of bisamides was prepared, and hexamethylene bisacetamide (HMBA) weakly inhibited the growth of murine erythroleukemia (MEL) cells at 5 mM concentration (**Figure 10.38**). SAR studies on the chain length showed that the 6-carbon linker was optimal, and animal studies with HMBA were initiated. The compound had low toxicity and progressed to phase II clinical trials for acute myelogenous leukemia and myelodysplastic syndrome in 1992. While there were partial remissions in some patients, these were transient, and high doses of the drug were required.

Without knowledge of the target, and based on an assumption that the amides of HMBA may be binding to a metal, the corresponding hydroxamides were synthesized. The analog suberic bis(hydroxamic acid) (SBHA) was twice as potent as the parent bisamide HMBA in causing growth arrest and death of transformed cells. As in the bisamide series, the optimal chain length was found to be six carbons. The next round of optimization came from the decision to maintain one hydroxamide and add a hydrophobic group at the other end. This modification resulted in the synthesis of SAHA in 1996, which was six times more potent than SBHA in cytotoxicity assays. The mechanism of action of these compounds was still unknown at this point; however, a member of the research group, Victoria Richon, noticed the similarity in structure of SAHA to the natural product trichostatin A. Although it had been isolated as an antifungal antibiotic in 1976, trichostatin A was identified as a growth inhibitor in Friend's MEL cells in 1986 and had been identified as an HDAC inhibitor in 1990. In 1998, SAHA was demonstrated to inhibit HDACs at low (50 nM) concentration. This inhibition results in selective apoptosis of transformed (cancer) cells.

In animal models, SAHA inhibited the growth of several types of tumors, including prostate, mammary, and neuroblastoma, with little to no toxicity. In addition, SAHA exhibited synergism with other anti-tumor agents, such as fludarabine, bortezomib, and imatinib. Sponsored by Columbia University, SAHA entered Phase I clinical trials in patients with refractory hematologic and solid tumors and showed little toxicity. At that time, no large pharmaceutical companies were interested in SAHA due to concerns that the hydroxamic acid moiety might cause toxicity due to the nonselective inhibition of Zn^{2+}-containing proteins and enzymes. As a result, the original discoverers formed Aton Pharma, which was then acquired by Merck after promising results from Phase II clinical trials. SAHA, now known as vorinostat, received FDA approval in 2006.

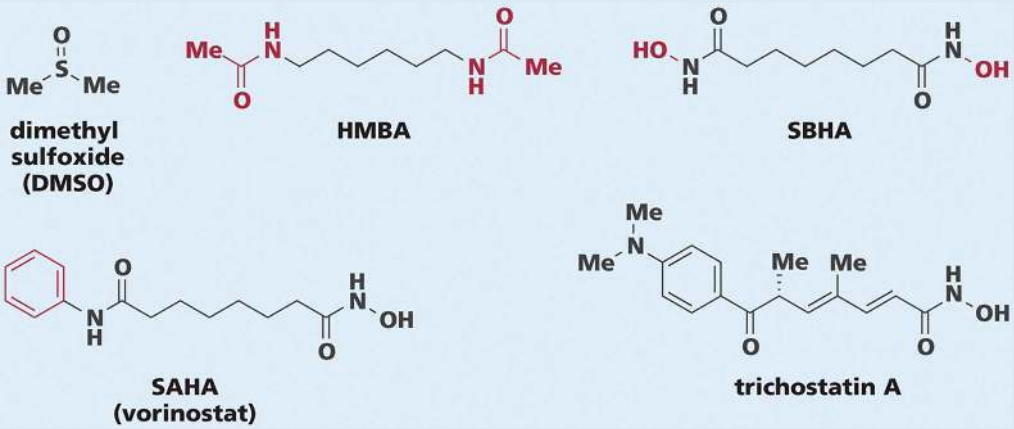


Figure 10.38 Structures involved in the development of vorinostat. DMSO was the initial lead, and SAR studies identified HMBA, which advanced to clinical trials. Further optimization led to the incorporation of hydroxamic acid groups (SBHA) and then desymmetrization to generate SAHA. The structural similarity between SAHA and trichostatin A is shown.

CASE STUDY 2 SELECTIVE PARP1 INHIBITOR SARUPARIB

The PARP1 enzyme (poly(ADP-ribose) polymerase 1 or poly-ADP-ribosyltransferase 1) catalyzes the covalent attachment of ADP-ribose moieties onto itself and other proteins. PARP1 is activated by DNA single-strand breaks (SSB). Zinc finger domains in the PARP1 enzyme bind to damaged sites on DNA, which results in hydrolysis of PARP1's NAD+ co-factor to generate ADP-ribose moieties that are sequentially added to PARP1 itself and various other proteins, resulting in proteins containing long poly(ADP-ribose) tails. The presence of these PARylated proteins recruits specific DNA repair enzymes to the lesion, which bind to and repair the SSBs after release of PARP1 from the DNA. If PARP1 is inhibited, the enzyme is "trapped" onto DNA, resulting in double-strand breaks in DNA rather than repair.

Inhibitors of PARP are particularly effective in cancers that lack the homologous repair (HR) pathway that fixes strand breaks. Due to the lack of the HR repair pathway, some tumor cells, such as BRCA1 and BRCA2 mutants, are more sensitive to the inhibition of PARP, resulting in synthetic lethality, while cells without those mutations have other mechanisms to overcome the inhibition. This difference in sensitivity allows for highly selective therapy.

The first generation of PARP inhibitors were mimics of NAD+ and include olaparib (2014), niraparib (2017), rucaparib (2016), and talazoparib (2018), shown in **Figure 10.39**. In all, there are six approved PARP inhibitors as of 2025, and they are mainly used to treat ovarian, breast, and also prostate cancers. Although effective, these are not selective, inhibiting both PARP1 and PARP2 to varying degrees. A consistent problem is hematological toxicity, which is likely due to the inhibition of PARP2.

Scientists at AstraZeneca initiated a program to identify selective PARP1 inhibitors, with the goal of reducing the hematological toxicity. Discovery efforts used a fluorescence polarization (FP) assay to evaluate selectivity of known PARP inhibitors. Compounds were also assayed in a BRCA2 mutant cell proliferation assay. This effort prioritized FR257531, a quinazolinone, which had 300-fold selectivity for PARP1 but poor pharmacokinetics, as reported by scientists at Fujisawa Pharmaceuticals

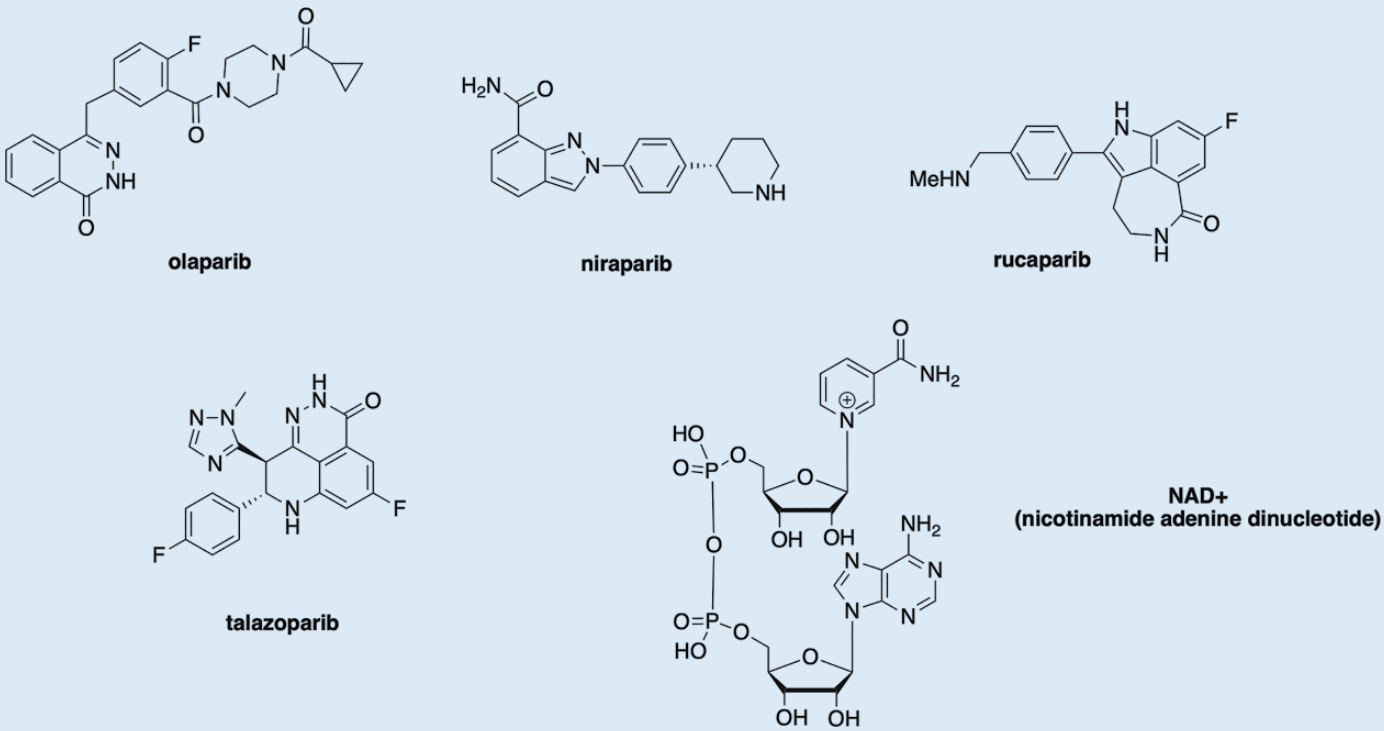


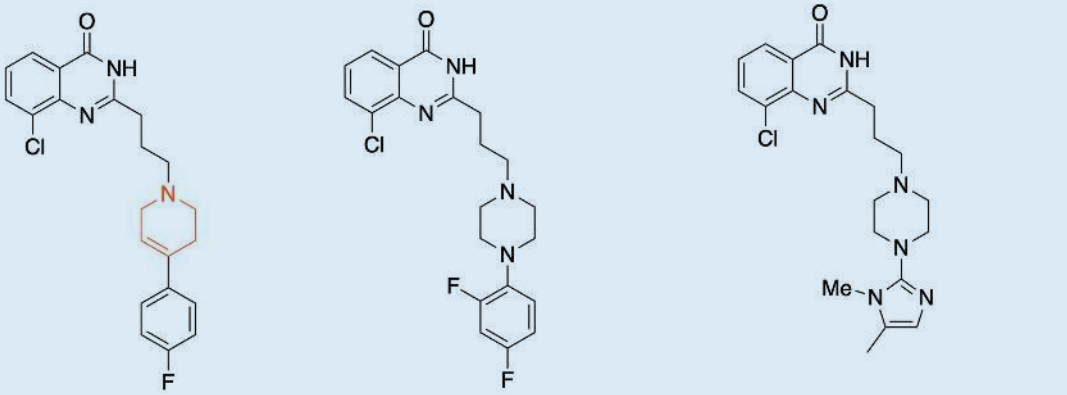
Figure 10.39 **Structures of PARP inhibitors, with NAD+.**

(Figure 10.40). Deficiencies included a relatively high logD, poor stability in liver microsome assays, poor solubility, and potential hERG liabilities. In addition, it had very low potency in the BRCA mutant cell proliferation assay.

Lead Modification

In the first round of modifications, efforts focus on using piperazines to replace the phenyl-tetrahydropyridyl group due to the potential of this moiety to lead to CNS side effects related to Parkinson's syndrome (tetrahydropyridyl in red; Figure 10.40). Two libraries, differentiated by the chemistries used to prepare each, were synthesized incorporating the same 8-chloroquinazolinone core, with piperazines replacing the tetrahydropyridyl group and diversification at the terminal piperazine substituent.

Lead candidates from the piperazine libraries were compounds **1**, with a difluorophenyl group, and **2**, with a dimethylimidazole moiety as the terminal rings. Both were highly selective, but focus was directed at analogs of **2** due to its optimal logD (2.3), as well as its potent activity in the cell proliferation assay. However, liabilities



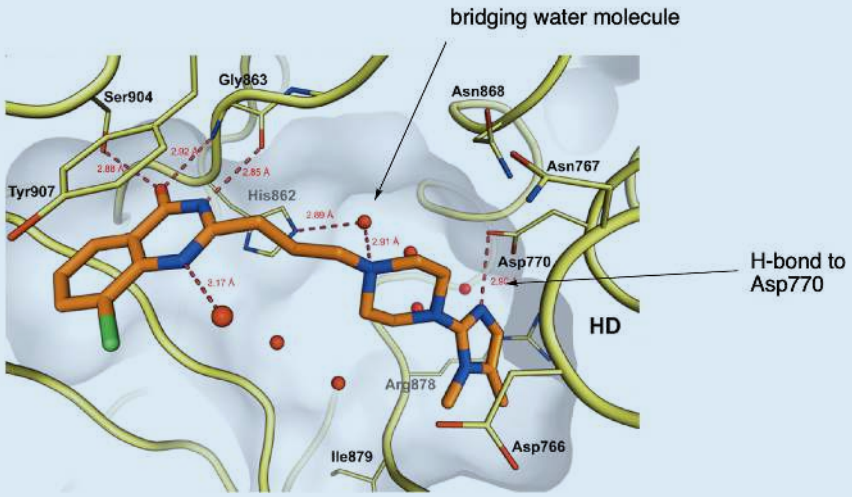


Figure 10.40 Structures of the lead compound FR257531 (tetrahydropyridyl in red) and two candidates from lead modification program. *Bottom*, X-ray structure of compound **2** bound to PARP1, with bridging water molecule and H-bond to Asp770, important for PARP1 selectivity.

Source: With permission from Johannes, J.; Balazs, A.; Barratt, D.; et al. (2021). *J Med Chem* 64:14,498–14,512.

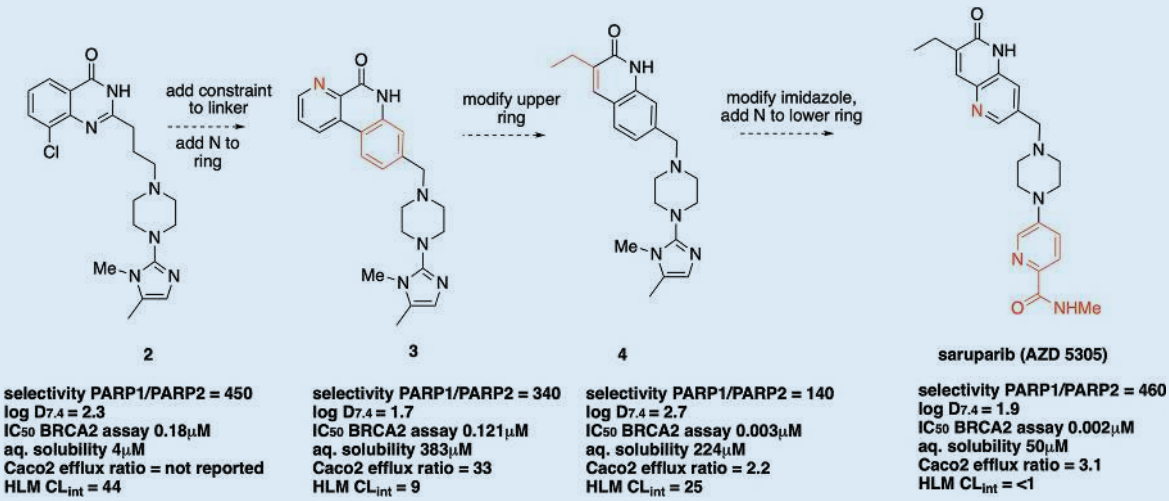


Figure 10.41 **Lead modification progression in the development of saruparib, with important properties.**

still had to be addressed, including hERG activity, high clearance, and low solubility. An X-ray structure of **2** bound to PARP1 shows that the quinazolinone mimics NAD+, as expected. An important feature is a water molecule interacting with the piperazine, which bridges to His862 in the active site and is a likely point of selectivity over PARP2. In addition, the imidazole forms a hydrogen bond to Asp770 and stacks against a helical domain, which also is a region where PARP1 differs from PARP2. In the next round of modifications, the focus was on several goals:

- Reduce logD and lower hERG liability (a problem with basic and lipophilic compounds)
- Improve solubility
- Decrease high clearance (CL_{int}) in human liver microsome (HLM) assays
- Improve potency in anti-proliferation assay
- Achieve low effux in Caco2 assay

Further optimization (**Figure 10.41**) included modifying the propyl linker in compound **2** to incorporate conformational constraint and the introduction of a nitrogen atom and the removal of the chlorine group in the upper ring to reduce lipophilicity, leading to compound **3**. This analog was less lipophilic, with less inhibition of hERG. It also maintained potency in the proliferation assay but still had high efflux in the Caco2 assay. To address this, replacements for the pyridyl ring were investigated, and better solubility was obtained by saturating and truncating the outer ring, to give compound **4**, which maintained good properties, with the exception of high clearance in the HLM assay.

With the upper core optimized, changes to the imidazole ring were investigated. Although a number of other 5-membered ring heterocycles were incorporated, none provided improvements over compound **4**. However, the use of a 6-membered aryl ring afforded promising compounds, with the N-methyl carboxamide maintaining excellent selectivity, anti-proliferative activity, and pharmacokinetic properties. Adding a nitrogen to the aryl ring further lowered the logD and resulted in very high potency and selectivity, lower clearance, with saruparib chosen for clinical trials.

A synthesis of saruparib used ethyl-5-amino-6-methylnicotinate (**5**) as starting material. Benzylic oxidation afforded an aldehyde, with concomitant oxidation of the aryl amine (**Figure 10.42**). A Horner-Wadsworth-Emmons reaction of the aldehyde produced alkene **6**. Hydrogenation, followed by acid treatment to effect ring closure, and then oxidation with DDQ, provided the naphthyridine **7**. The reduction of the ester to an alcohol and reaction with thionyl chloride afforded the chloride **8**. Introduction of the piperazine side chain occurred via displacement of the chloride with the piperidine **9** to afford saruparib.

Saruparib entered clinical trials in 2023 and has been tested in patients with breast, ovarian, prostate, and pancreatic cancers. There are fewer adverse events than with the nonselective PARP inhibitors, allowing a higher dose. In early trials in patients with HR-deficient breast cancers, at a 60 mg daily dose, an objective response rate of 48% was seen, with a median progression-free survival of 9.1 months.

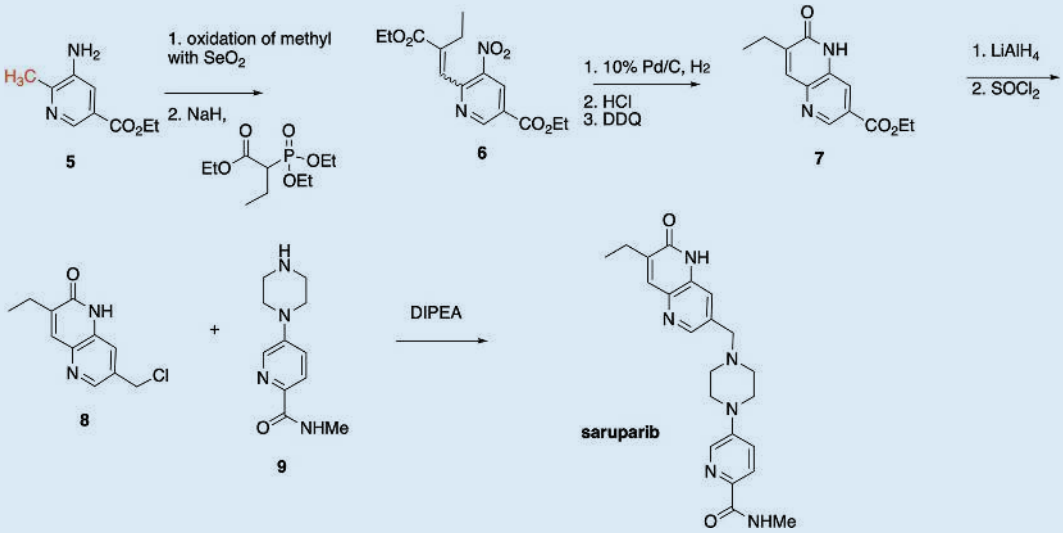


Figure 10.42 **General outline of synthesis of saruparib.**

10.9 REVIEW QUESTIONS

1. Explain why drugs that target DNA directly are intrinsically more toxic than drugs that target oncogenes.
2. Given the interactions from the X-ray co-crystal structure of bortezemib bound to the proteasome proteolytic site shown in **Figure 10.30**, draw the expected structure of MG132 bound to the same sites.
3. Several examples in this chapter were developed by the use of structure-based design. Name one class that is particularly amenable to this approach and explain why.
4. Explain the importance of the ethynyl moiety in ponatinib.
5. Explain why natural products are a frequent source of cancer drugs. List five different cancer drugs (of at least three different mechanisms) that were derived from natural products, and their mechanisms of action.
6. Identify two examples of a suicide inhibitor described in this chapter.
7. Explain why hormone receptor antagonists are useful therapies for certain cancers. Conversely, explain why hormone replacement therapy may be harmful.
8. Explain how the inhibition of PARP is most effective in certain cancers.
9. A number of anti-cancer agents are also used as anti-inflammatory or immunosuppressant agents. Explain why.
10. Review the drug examples in this chapter, and find three where an isosteric replacement strategy was used. Draw the structures of the drugs and the original compounds, and identify the isosteres.

10.10 APPLICATION QUESTIONS

1. Research the drug raltitrexed. Draw its structure and explain its mechanism of action.
2. Hydroxamic acids and thiols are zinc-binding groups found in FDA-approved histone deacetylase inhibitors. Research another zinc-binding group that is found in HDAC inhibitors being evaluated in the research phase.
3. Research antibody–drug conjugates that are in clinical trials. Choose one. What does the antibody target? What is the small-molecule payload, and what is its biological target? Is either component of the ADC an approved drug itself?
4. Crizotinib is a drug that targets an oncogenic fusion protein. Research its structure and its mechanism of action.
5. The compound neocarzinostatin has been used in Japan to treat liver cancer. Research its structure and its mechanism of action.
6. Review the binding of camptothecin to topoisomerase in **Figure 10.8**. Draw a pharmacophore that is consistent with these interactions, and design a novel compound based on this pharmacophore.
7. Research the mechanism of epoxomicin's reaction with the proteasome's active site threonine. Draw the product of that reaction.
8. Review the examples in this chapter, and identify two compounds (targeting different proteins or enzymes) that incorporate functional groups that improved their solubility. Draw the structures with and without the solubilizing groups in ChemDraw (or another program), and calculate the differences in solubility and cLogP for each. How different are the pairs?
9. Review the structures of methotrexate and pralatrexate. Draw a pharmacophore model that is consistent with both. Design two new molecules that will test two different aspects of your pharmacophore.
10. Identify three compounds in this chapter that contain fluorine. From what was described in Chapter 4, the mechanism of action of each chosen drug and the site of the fluorine propose the function of fluorine in each molecule.

FURTHER READING

General References

Alberts B, Johnson A, Lewis J et al. (2002) Chapter 23: Cancer. In *Molecular Biology of the Cell*, 4th ed, pp 1091–1144. Garland Science.

Kamb A, Wee S & Lengauer C (2007) Why is cancer drug discovery so difficult? *Nat Rev Drug Discov* 6:115–120.

Drugs Targeting Topoisomerase

Deweese JE & Osheroff N (2009) The DNA cleavage reaction of topoisomerase II: wolf in sheep's clothing. *Nucleic Acids Res* 37:738–748.

Redinbo MR, Stewart L, Kuhn P et al. (1998) Crystal structures of human topoisomerase I in covalent and noncovalent complexes with DNA. *Science* 279:1504–1513.

Staker BL, Feese MD, Cushman M et al. (2005) Structures of three classes of anticancer agents bound to the human topoisomerase I–DNA covalent complex. *J Med Chem* 48:2336–2345.

Wu C-C, Li T-K, Farh L et al. (2011) Topoisomerase inhibition by the anticancer drug etoposide. *Science* 333:459–462.

Drugs Targeting Signaling Pathways

Chen H, Smaill JB, Liu T et al. (2020) Small-molecule inhibitors directly targeting KRAS as anticancer therapeutics. *J Med Chem* 63:14404–14424.

Glaviano A, Foo ASC, Lam HY et al. (2023) PI3K/AKT/mTOR signaling transduction pathway and targeted therapies in cancer. *Mol Cancer* 22:138.

Su W & Hou X (2024) Targeting active RAS with molecular glue. *Pharm Sci Adv* 2:10047.

Drugs Targeting Microtubules

Jordan MA & Wilson L (2004) Microtubules as a target for anticancer drugs. *Nat Rev Cancer* 4:253–265.

Epigenetics

Blancafort P, Jin J & Frye S (2013) Writing and rewriting the epigenetic code of cancer cells: from engineered proteins to small molecules. *Mol Pharmacol* 83:563–576.

Falkenberg KJ & Johnstone RW (2014) Histone deacetylases and their inhibitors in cancer, neurological diseases and immune disorders. *Nat Rev Drug Discov* 13:673–691.

Special issue on epigenetics. (2016) *J Med Chem* 59:1247–1654.

Proteasome

Borissenko L & Groll M (2007) 20S Proteasome and its inhibitors: crystallographic knowledge for drug development. *Chem Rev* 107:687–717.

Orlowski RZ & Kuhn DJ (2008) Proteasome inhibitors in cancer therapy: lessons from the first decade. *Clin Cancer Res* 14:1649–1657.

Angiogenesis

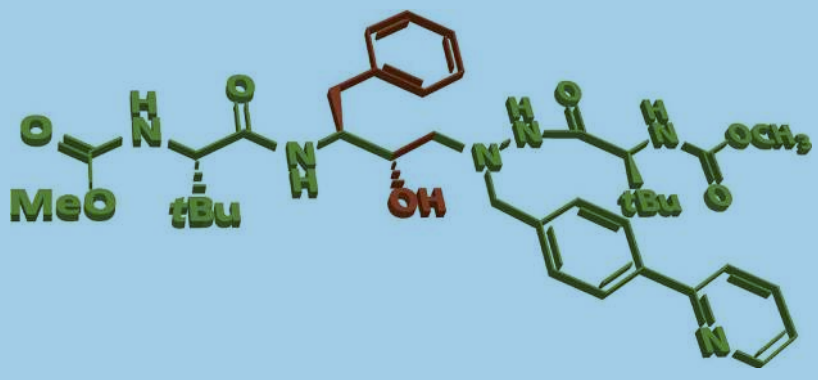
Kerbel R & Folkman J (2002) Clinical translation of angiogenesis inhibitors. *Nat Rev Cancer* 2:727–739.

Case Studies

Johannes J, Balazs A, Barratt D et al. (2021) Discovery of 5-{4-[(7-Ethyl-6-oxo-5,6-dihydro-1,5-naphthyridin-3-yl)methyl]piperazin-1-yl}-N-methylpyridine-2-carboxamide (AZD5305): a PARP1–DNA trapper with high selectivity for PARP1over PARP2 and other PARPs. *J Med Chem* 64:14498–14512.

Marks PA & Breslow R (2007) Dimethyl sulfoxide to vorinostat: development of this histone deacetylase inhibitor as an anticancer drug. *Nat Biotechnol* 1:84–90.

Meinke PT & Liberator P (2001) Histone deacetylase: a target for antiproliferative and antiprotozoal agents. *Curr Med Chem* 8:211–235.



Antiviral and Antifungal Agents

11

LEARNING OBJECTIVES

- Describe common points of intervention for antiviral therapy.
- Identify classes of drugs that target viral enzymes involved in nucleic acid replication.
- Understand the medicinal chemistry approaches taken to discover and develop antiviral protease inhibitors.
- Describe drugs that target viral attachment, uncoating, and budding.
- Identify classes of antifungal drugs
- Explain the mechanism of action of antifungal agents.

11.1 INTRODUCTION

Both viral and fungal infections lead to diseases and conditions that range from mild, such as the common cold or athlete's foot, to deadly, such as AIDS or systemic fungal infections. Until the mid-1900s, there were no safe drugs to treat either type of infection. Historically, however, certain viral infections, such as smallpox and polio, have been prevented with vaccines which, when administered, provoke an immune response so that antibodies are present in case of exposure. Vaccines have been in use since 1796, when Edward Jenner advocated the use of cowpox vaccine to protect against smallpox. Although they are effective against many viruses, for a variety of reasons, vaccines against certain viruses, such as HIV, as well as against fungal infections, have not yet been developed, despite the medical need, and they are not effective once the infection is established.

For the instances when vaccines are not available (or not effective), there are a number of approved drugs to treat viral infections, but far fewer drugs are available to treat fungal infections. These numbers are small, however, when compared to drugs that treat bacterial infections, which are the subject of Chapter 12. There are several reasons for this disparity, but a primary explanation is related to the difference between bacterial, viral, and fungal drug targets and human homologs. Since viruses often rely, to varying degrees, on certain human proteins to replicate, there are only a few viral-specific targets that drugs can affect. The reason for the lack of drugs for fungal infections is different. Fungi are eukaryotic organisms, as are animals. As a result, the potential drug targets for fungi are very similar to the targets found in humans, which leads to difficulties in selectivity. In contrast, prokaryotic bacteria have their own replication machinery, and therefore, drug targets are often quite distinct from mammalian targets, making selectivity between human and bacterial targets more readily achievable.

DOI: 10.1201/9781003465607-14

11.2 ANTIVIRAL AGENTS: COMMON VIRAL STRUCTURES, REPLICATION PROCESS, AND IMPACT

Structurally, all viruses are infectious particles, or **virions**, composed of a nucleic acid containing the viral genome and a protein coat or capsid that is encoded by the viral genome. The viral genome is often associated with proteins, called nucleoproteins, and the combination of the genome and its associated proteins is called the nucleocapsid. In some viruses, the nucleocapsid is helical and the protein subunits, or protomers, form rigid rods or flexible filaments around the nucleic acid. In other viruses, the protomers are arranged in clusters, called capsomeres, that form an icosahedral shell, a polygon with 20 equilateral triangular faces and 12 vertices. Some viruses may be spherical, like the influenza virus (**Figure 11.1**). Some viruses, in addition, have an envelope surrounding the nucleocapsid. This envelope is a lipid bilayer that is derived from a modified host cell membrane with an outer layer of glycoproteins derived from the virus; these glycoproteins are involved in determining which specific cells the virus infects.

The nucleic acid contained in the virion can be either single-stranded (ss) or double-stranded (ds) RNA or DNA. It can be configured in either a linear or circular orientation, and the entire genome may be present either on a single nucleic acid molecule or on several segments of the nucleic acid. Genomes can also be either (+)-sense, which can be translated immediately into protein, or (–)-sense, which must be first converted to (+)-sense RNA before translation. Since viruses are intracellular parasites, they cannot express their gene products or replicate without the help of a living cell; the specific replication strategy depends on the type of genome in the virus. For example, most DNA viruses replicate in the cell nucleus and use host cell DNA-dependent RNA polymerase to synthesize mRNA. The early proteins translated from the mRNA include viral DNA polymerases that are responsible for the ongoing replication of the viral genome. Most RNA viruses replicate in the cytoplasm and have their own polymerases to transcribe RNA into mRNA, since the host cell does not have an enzyme to accomplish this. Retroviruses are a specific type of RNA virus containing a dimer of two identical (+)-single-stranded RNA molecules. The retroviral reverse transcriptase transcribes the viral RNA into dsDNA (reverse transcription), which is then incorporated into host DNA by the viral enzyme integrase (integration). Regardless of the strategy used, once mRNA is synthesized, it is translated into viral proteins, such as viral kinase and viral DNA polymerase, needed for the production of viral progeny.

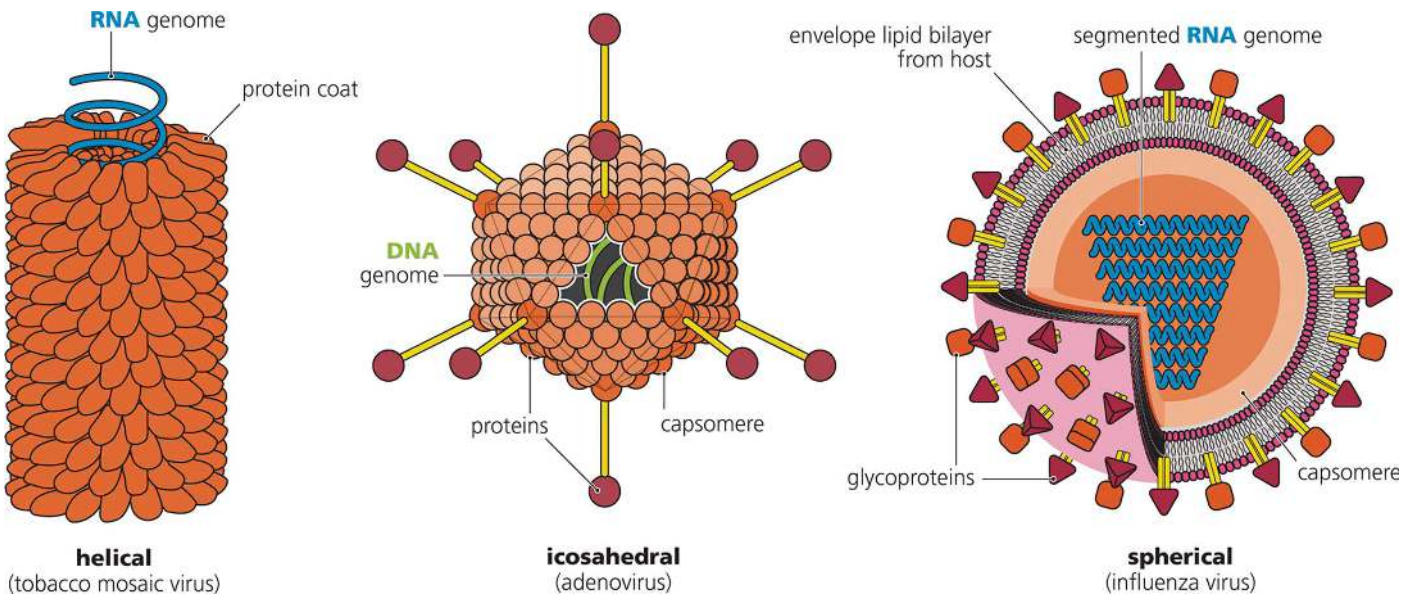


Figure 11.1 Shapes and components of virion structures. Common components are nucleic acid genome (DNA or RNA) enclosed in a protein coat, the capsid. The protein coat can form helical (e.g., tobacco mosaic virus); icosahedral, made up of clusters called capsomers (e.g., adenovirus); and spherical shapes (e.g., influenza virus). Some viruses contain a lipoprotein envelope made up of lipids derived from the host and may have glycoprotein spikes on the surface. The nucleic acid can be present as one strand, or several segmented strands.

While the exact steps may vary from one virus to another, all viruses follow a general replication cycle of attachment, entry, uncoating, transcription, and translation to early (replicative) and late (structural) proteins, replication, virion assembly, budding, and release from the host cell, allowing the new virion to infect additional cells. A detailed process for the life cycle of the influenza virus, an enveloped ssRNA virus, is illustrated in **Figure 11.2**, with an emphasis on possible points for drug action.

The specific steps in the replication cycle are as follows:

- *Attachment.* The virion binds to the host cell by interacting with specific host cell surface receptors.
- *Entry.* The virion may enter the cell by endocytosis or by fusion with the cell membrane.
- *Uncoating.* The virion sheds its protein coat to release the viral genome.
- *Transcription.* Viral nucleic acids are transcribed via mechanisms that are dependent on the specific virus class (DNA, RNA, or retrovirus).
- *Translation.* Proteins are translated from viral mRNA by use of host machinery.
- *Integration.* The DNA of some viruses is integrated into host DNA, resulting in that cell becoming permanently infected (not shown).
- *Replication.* Viral genome is replicated (usually by use of host machinery).
- *Virion assembly.* New viral particles are packaged.
- *Budding and release.* New viral particles bud from the infected cell (enveloped viruses) or cause the host cell to burst, releasing viral particles.
- *Re-initiation.* The process begins again.

Virions can go on to infect other cells but can also exist externally (in the air, in mucus, and in blood) and can infect other individuals. Viruses that are stable in airborne particles, such as influenza, are highly contagious since large numbers of virions are dispersed when an infected individual sneezes or coughs, and others can be infected when they come into contact with a viral particle. Other viruses are spread less easily; for example, those that are transmitted via blood, such as the human immunodeficiency virus (HIV) and hepatitis C virus (HCV). Therapeutic intervention to prevent healthy but at-risk individuals from infection (**prophylaxis**) is a strategy in antiviral

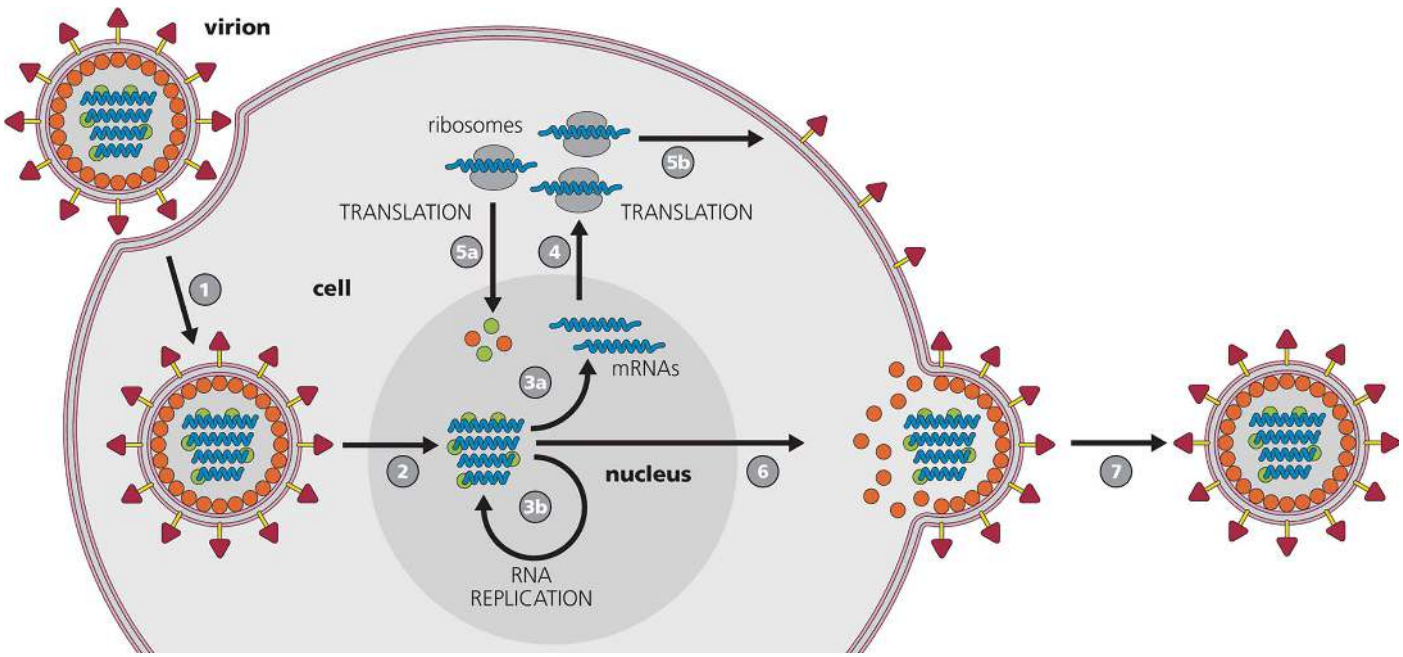


Figure 11.2 Life cycle of influenza viruses. Step 1: Attachment, by interaction between host cell receptors and glycoproteins on the surface of the virion, and entry. Step 2: Uncoating and release of genetic material into the nucleus. Steps 3a and 3b: Replication of genetic material. Steps 4, 5a, and 5b: Translation of viral proteins, followed by virion assembly on ribosomes. Step 6: Budding. Step 7: Release of a new viral particle.

drug discovery that focuses on intervening prior to the integration step. For example, it is routinely used when health-care workers have an occupational exposure to HIV, to treat pregnant HIV-positive women to prevent transmission to neonates, and in individuals at high risk for HIV infection.

The pathological effects of a viral infection vary greatly. Some viruses, such as rhinoviruses, are relatively harmless, and infected individuals recover with no medical treatment. Infection with HCV often causes no acute symptoms but, years later, can result in very serious disorders, such as liver cancer or cirrhosis. Influenza and SARS-CoV-2 (severe acute respiratory syndrome coronavirus) cause severe acute symptoms that can be deadly in certain patient populations. The specific symptoms of a viral infection depend on the type of cell it targets, the fate of that cell, and the life cycle of the virus. Some infected cells are destroyed by the host immune response; some die because the virus overtakes the cellular machinery; some cells burst and die, releasing virions; and other cells replicate with the virus, remaining latent for extended periods of time (years) before the virus reactivates. Some symptoms of a viral infection can be the result of the body's own immune response to the infection (as is the case with the common cold) or because the function of the infected cell is compromised. For example, HIV targets immune cells, and symptoms (immunosuppression) are a consequence of the depletion of those cells. The respiratory symptoms of influenza infection reflect that it targets airway epithelial cells. SARS-CoV-2 targets the angiotensin-converting enzyme 2 (ACE2) receptor, which is found in the heart, lungs, and kidneys. The ACE2 receptor binds to the spike protein of the virus, which facilitates entry. Although mRNA vaccines had been tested in the early 1990s for influenza vaccines, advances in the use of lipid nanoparticles led to the first FDA-approved mRNA vaccines for COVID–19. The vaccines are based on the mRNA of the spike protein of SARS-CoV–2.

Untreated viral infection can cause significant mortality and morbidity. During the 1918 influenza pandemic, it is estimated that one-third of the world's population was infected (~500 million), and between 50 and 100 million people died in a single year. While modern versions of influenza are less severe, the potential for a virulent strain spreading rapidly, in a world where transmission is easier than it was in the early 20th century, is high. HIV, the causative agent of acquired immunodeficiency syndrome (AIDS), is a bloodborne pathogen. As of 2024, over 40 million people are living with HIV worldwide, and over 44 million have died since 1981, when AIDS cases were first described. While today HIV infection can be controlled, the life expectancy of an infected individual in the 1980s was less than two years. Globally, 130–150 million individuals are infected with the hepatitis C virus, and >60% will develop chronic liver disease. In the United States, the major cause of liver transplants is HCV infection, and no vaccination is available. Unlike many other diseases, such as cancer, which has been known for thousands of years, new viral diseases are continually emerging; therefore, the need to develop new therapies is continual. Most recently, in late 2019, the world saw the emergence of the coronavirus SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2), the cause of COVID-19 (coronavirus disease-19), and a worldwide pandemic that spread rapidly and killed over 18 million people within the first two years. SARS-CoV-2 followed an outbreak of SARS-CoV-1 in 2002–2004 and MERS-CoV (Middle East respiratory syndrome) in 2012–2015. While rapid development of vaccines, and several new medications, helped immensely in taming the SARS-CoV-2 pandemic, long-term effects, such as long COVID, still linger.

11.3 ANTIVIRAL DRUG DISCOVERY

While vaccines have been highly effective in preventing and even eradicating certain diseases (e.g., smallpox) caused by viruses, in some instances, no effective vaccine exists (for example, HIV and HCV), and in other cases, such as influenza, infections still occur even though a vaccine is available. Despite the need, small-molecule antiviral drug discovery is a relatively new area. One reason is that, for certain viruses, highly specialized skills, equipment, and laboratories are required to perform even the simplest cell-based assays. Furthermore, animal models that replicate human viral diseases are often inadequate. Therefore, early research in antiviral agents was limited to specialized organizations and laboratories.

Some of the first antiviral drugs were developed in the 1970s at Burroughs Wellcome to treat herpes simplex viruses (HSV). Drugs were characterized using a phenotypic assay (a plaque inhibition assay) that monitored HSV-1 replication in mammalian cells. The approach used was based on rational design to inhibit fundamental biochemical processes involved in herpesvirus replication and was related to the earlier work of Gertrude Elion and George Hitchings on antimetabolites to treat cancer. The success of this effort, which produced a highly safe, selective, and effective drug called acyclovir, spurred others to pursue antiviral drug discovery and validated the feasibility of targeting **viral-specific proteins**.

These discoveries were followed in the 1980s by massive efforts to identify treatments for AIDS. While the first AIDS drugs were known nucleosides, such as azidothymidine (AZT), dideoxycytidine (ddC), and dideoxyinosine (ddI), library screening later identified nonnucleoside inhibitors of the viral enzyme reverse transcriptase, such as nevirapine. The 1988 Nobel Prize in Medicine, awarded to Elion, Hitchings, and James Black, noted that the idea of designing drug molecules that bear structural similarity to natural substances was applied to the design of anti-AIDS drugs.

Investigation of the inhibition of viral proteases has led to many successful drugs, including the transition state mimetics saquinavir, indinavir, and ritonavir used to treat HIV infections. These compounds were some of the first drugs whose development was based on iterative structure-based, computer-assisted drug design using X-ray crystallography. More recent examples of viral protease inhibitors include the HCV-NS3 protease inhibitors grazoprevir and paritaprevir that are used to treat hepatitis C infections. Novel nucleosides, such as sofosbuvir, to treat HCV have also been developed and are curative when used in combination with other agents. Most recently, nirmeltravir, an inhibitor of SARS-CoV-2 protease, was introduced and is used in combination with ritonavir to treat COVID-19. The search for drugs to treat HIV/AIDS, HCV, and COVID-19 is an excellent example of how basic science research combined with drug discovery efforts allowed the rapid development of drugs to treat or cure a newly identified virus and the disease it causes.

Table 11.1 Common Infectious Viruses, Select Examples of Drugs, and Their Mechanisms of Action

Virus	Category	Family	Drug	Mechanism
Herpes simplex virus (HSV-1, HSV-2)	dsDNA	*Herpesviridae*	Acyclovir, valacyclovir	Polymerase inhibitors
Varicella zoster virus (VZV)	dsDNA	*Herpesviridae*	Acyclovir, valacyclovir	Polymerase inhibitors
Cytomegalovirus (CMV)	dsDNA	*Herpesviridae*	Ganciclovir	Polymerase inhibitor
			Fomivirsen (withdrawn)	Antisense, early gene products
Human immunodeficiency virus (HIV)	ssRNA	*Retroviridae*	Azidothymidine, nevirapine	Reverse transcriptase inhibitors
			Indinavir	Protease inhibitor
			Maraviroc	Entry/fusion inhibitor
			Raltegravir	Integrase inhibitor
Hepatitis C virus (HCV)	ssRNA	*Flaviviridae*	Ribavirin	Multiple viral and host targets
			Sofosbuvir	Polymerase inhibitor
			Telaprevir	Protease inhibitor
			Daclatasvir	NS5A inhibitor
Hepatitis B virus (HBV)	Partial dsDNA	*Hepadnaviridae*	Lamivudine, entecavir	Polymerase inhibitors
Influenza virus	ssRNA	*Orthomyxoviridae*	Amantadine	Uncoating inhibitor
			Oseltamivir, zanamivir	Neuraminidase inhibitor
SARS-CoV-2	ssRNA	*Coronavirus*	nirmeltravir	protease inhibitor
			remdesivir	polymerase inhibitor

While much of the notable progress in antiviral drugs has been made in the areas of HSV, HIV, and HCV, there are many other successes in treating viral infection. Drugs have been developed to treat influenza and infection by various members of the *Herpesviridae* family, including HSV-1 and HSV-2, varicella zoster virus (VZV), and cytomegalovirus (CMV). These herpesviruses can cause venereal disease and cold sores (HSV), chicken pox and shingles (VZV), and birth defects and retinitis (CMV). The idea of targeting viral-specific enzymes has also been applied to other viral infections, including Dengue virus, West Nile virus, and Zika virus. **Table 11.1** contains a representative list of viral diseases for which drugs are available, examples of drugs, and their mechanisms of action.

11.4 ANTIVIRAL DRUGS TARGETING VIRAL LIFE CYCLE

While the discovery of antiviral drugs was initially based on a combination of design and phenotypic assays, it has progressed to being based primarily on rational design targeting viral-specific proteins. Some drugs are thus specific for certain viruses, such as HIV protease inhibitors, while others may be effective against a number of different viruses, such as DNA polymerase inhibitors. Whether they target one virus, or several, all antiviral drugs affect a certain point in the viral replication cycle and can be categorized as such.

11.4.1 Drugs May Target Viral Attachment or Entry

There are a few drugs that target viral attachment or entry (in particular, fusion), and they are approved to treat HIV/AIDS. The HIV fusion inhibitors prevent HIV particles from infecting T cells by blocking protein–protein interactions (see Chapter 8). The initial steps of infection by HIV are shown in **Figure 11.3**. The HIV envelope glycoprotein, on the surface of the virion, consists of two subunits, gp120 and gp41. Binding between the virion and the target T cell is mediated by the interaction between CD4 proteins on the surface of the T cells and gp120. The HIV glycoprotein then undergoes a series of conformational changes and binds to a co-receptor, a chemokine receptor (usually CCR5, but sometimes CXCR4), also on the surface of the T cell. Ultimately, the membranes fuse, and the contents of the virion enter the T cell.

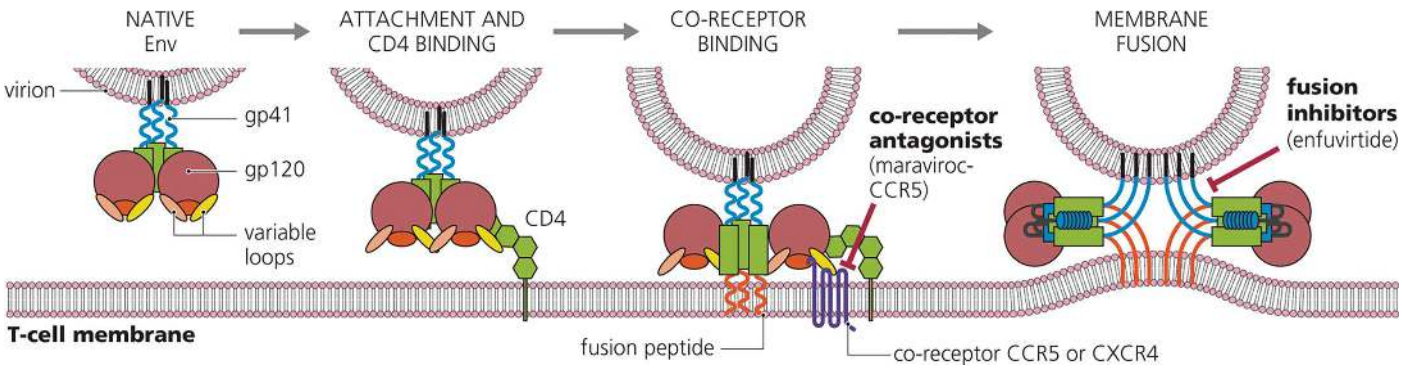


Figure 11.3 Attachment and entry of HIV viral particle into host T cell and points of intervention. The HIV surface glycoproteins gp120 (depicted in red, with variable loops in orange, pink, and yellow) and gp41 (blue lines/green rectangles) mediate attachment and membrane fusion with the target T cell. Gp120 binds to CD4 receptors (depicted as green hexagons) on the surface of the host T cell. Conformational changes of variable loops enable binding to a co-receptor (either chemokine receptor CCR5 or CXCR4, shown in purple), with activation of the fusion machinery (orange lines). Once bound, the membranes fuse and the contents of the viral particle enter the cell. Maraviroc, and other co-receptor antagonists, block the protein–protein interaction between gp120 and chemokine receptors required for entry. Enfuvirtide binds to a region of gp41 and prevents fusion.
Source: Adapted from Didigu, C. A., and Doms, R. W. (2012). *Viruses* 4:309–324. With permission from MDPI.

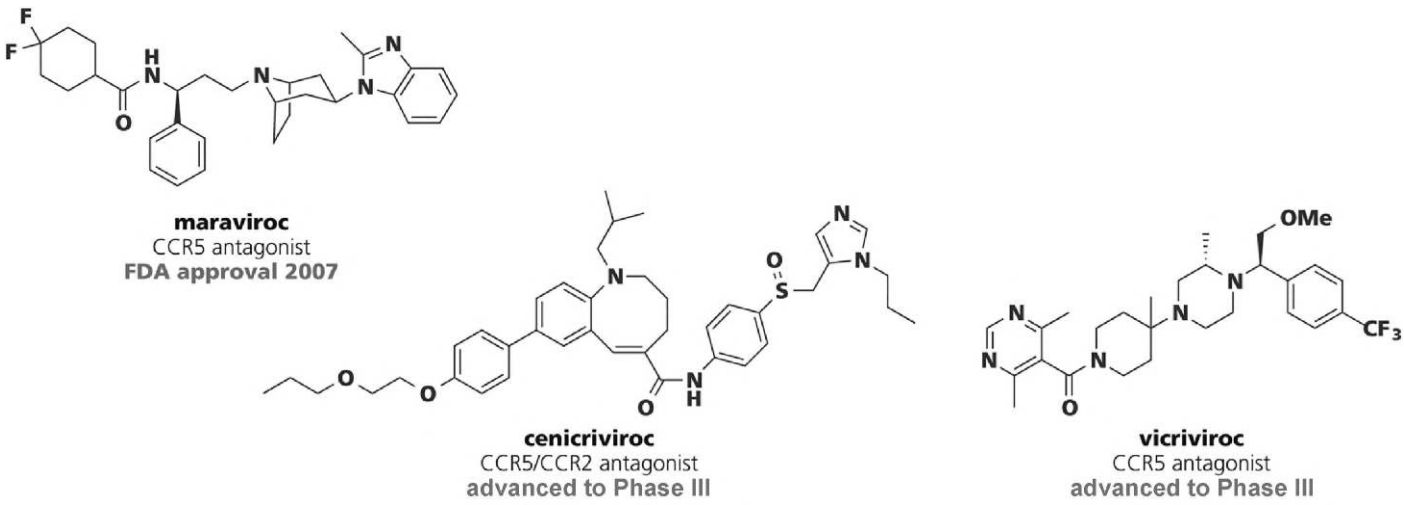


Figure 11.4 Structures of gp120–chemokine antagonists. Maraviroc was approved to treat AIDS in 2007; cenicriviroc and vicriviroc advanced to clinical trials but were discontinued for AIDS treatment. These compounds originated from hits identified from independent library screening campaigns.

One approach to blocking attachment or entry of HIV is to interrupt the interaction between the chemokine co-receptors and gp120. These chemokines are G protein–coupled receptors, and entry inhibitors block their binding to gp120. The first lead in the development of this type of inhibitor was a peptide based on the chemokine's natural ligand, the protein RANTES, although those leads did not result in viable drugs. Other small molecules were developed from screening programs and have resulted in marketed drugs (see Case Study 2 in Chapter 8 for a discussion of maraviroc). Currently, maraviroc is the only approved drug that works via this mechanism of blocking interaction between gp120 and co-receptors; however, several others, including cenicriviroc and vicriviroc (in combination), progressed to clinical trials for HIV infection (Figure 11.4).

Another approach to block HIV entry into cells led to enfuvirtide, which is a synthetic 36–amino acid peptide. The peptide binds to gp41, preventing it from adopting the specific conformation that is necessary for membrane fusion to occur.

11.4.2 Drugs May Target Uncoating of Viral Particles

The uncoating process varies among viruses and is complex, but inhibiting it is a viable therapeutic approach. To date, the only uncoating inhibitors approved are amantadine and rimantadine, and their effectiveness is poor. Amantadine (Figure 11.5) was found to have antiviral activity against the influenza A virus in 1963 and was approved for that use in the United States in 1966. The mechanism of action of amantadine and rimantadine involves binding to the viral M2 protein. This proton ion channel is specific to influenza A viruses and is involved in the early stage of uncoating.

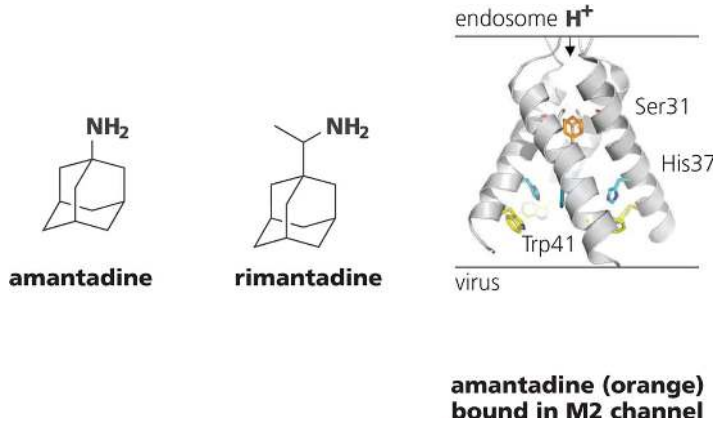


Figure 11.5 Structures of amantadine and rimantadine, and X-ray co-structure. *Left*, structures of amantadine and rimantadine. *Right*, X-ray crystal structure of amantadine bound in the M2 channel. Protons pass through the channel to allow uncoating, and amantadine (shown in orange) physically blocks the pore of the ion channel; mutations at Ser31, His37, and Trp41 confer resistance to amantadine.

Source: Left, from Das, K.; Aramini, J. M.; Ma. L.-C., et al. (2010) *Nat Struct Mol Biol* 17:530–538. With permission from Macmillan Publishers Ltd.

The viral M2 ion channel is a homotetrameric channel, and an influx of protons acidifies the interior of viruses trapped in endosomes, allowing the virus to dissociate and uncoat. Several X-ray and NMR structures of amantadine bound in the channel have been solved. Two of the structures place amantadine in the pore, essentially physically blocking the channel, as seen in the X-ray structure in **Figure 11.5**. Drug resistance is conferred by mutations to Ser31, His37, and Trp41. Side effects of amantadine and rimantadine, possibly caused by inhibition of host ion channels, as well as viral resistance, have significantly limited the use of these drugs, although there has been renewed interest in their effects against avian flu.

Another example of antiviral agents that block uncoating is based on a phenyl ether scaffold. A series of small molecules was investigated in the 1980s to treat rhinoviruses, the causative agent for the common cold. They were developed around the advanced analog disoxaril, which was, in turn, evolved from a lead compound called arildone (**Figure 11.6**). Screening at Sterling-Winthrop identified arildone as a compound with some activity against the poliovirus, and lead optimization led to disoxaril, which had activity against several rhinovirus serotypes. Its mechanism involves binding to a hydrophobic cleft on the viral capsid that prevents uncoating of the virus. The most advanced compound in the series, pleconaril, exhibited oral activity in mice infected with enteroviruses and rhinoviruses. Pleconaril entered clinical trials but was ultimately rejected as an oral drug due to safety issues and low efficacy.

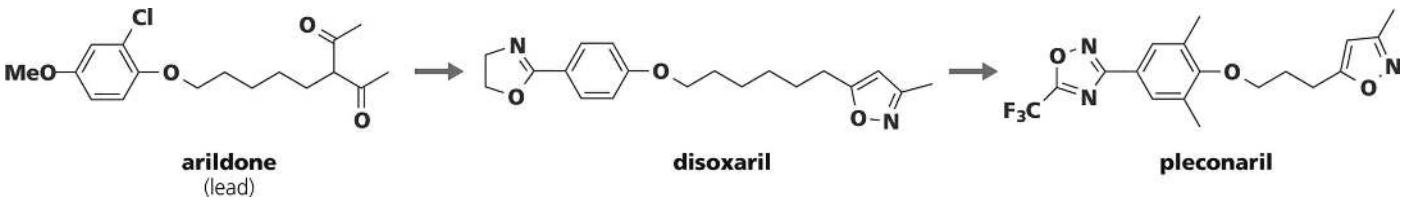
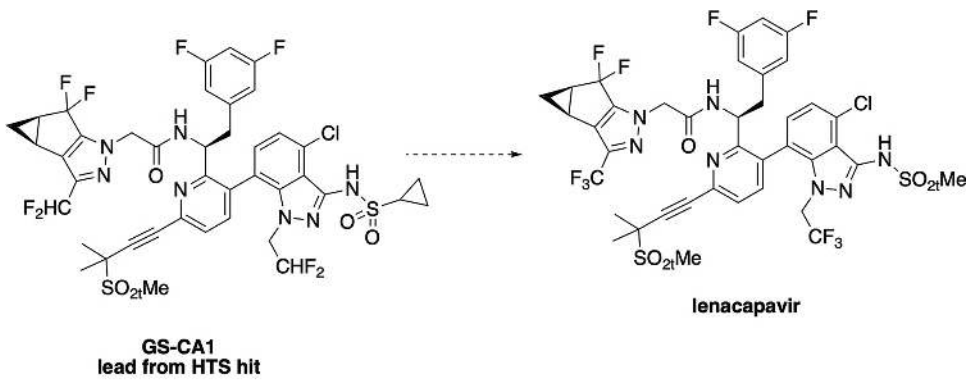


Figure 11.6 Structures of rhinovirus uncoating inhibitors. Arildone was identified as a compound with activity against poliovirus. Optimization led to disoxaril, which exhibited activity against rhinoviruses. Further optimization generated pleconaril, which advanced to clinical trials for the treatment of the common cold.



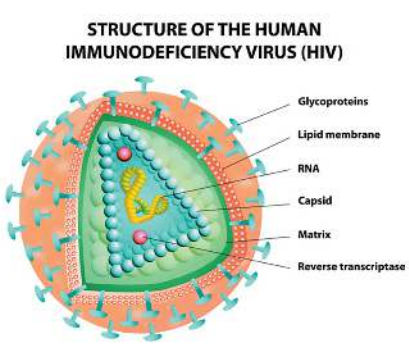


Figure 11.7 Structure of lead compound GS-CA1 and lenacapavir, with model of HIV showing capsid core.

A very different type of drug that binds to a capsid, lenacapavir, was approved in 2022 to treat people infected with resistant HIV. Lenacapavir is also remarkably effective in high-risk populations as a PrEP (pre-exposure prophylaxis) agent. This long-acting drug can be given just twice a year and is 90–100% effective in preventing transmission (Figure 11.7). Lenacapavir was discovered at Gilead Sciences, in collaboration with Wesley Sundquist at the University of Utah. Studies of the cone-shaped HIV capsid indicated its importance in viral replication. The HIV capsid (CA) is composed of ~1,500 individual CA monomers that assembly into hexamers and pentamers, and accurate assembly is required for infectivity. With the viability of the capsid as a valid drug target, high-throughput screening for molecules binding to CA was initiated and identified leads which were modified to give GS-CA1, with further modification to afford lenacapavir. This novel HIV capsid inhibitor protects people from HIV infection for six months after only one injection. It was approved in 2022, and *Science* magazine designated it a 2024 "Breakthrough of the Year," citing its "astonishing 100% efficacy in high-risk populations."

11.4.3 Many Antiviral Drugs Inhibit Transcription by Targeting the Active Site of Polymerases

As outlined in Chapter 7 in more detail, kinases can catalyze phosphorylation of the 5′-position of nucleosides, and polymerases or reverse transcriptases recognize these nucleotides and catalyze the formation of the 3′–5′ phosphodiester bond between the incoming nucleotide and the growing nucleic acid chain. Some of the first antiviral drugs were nucleosides or nucleoside mimics; they utilize this pathway to generate their corresponding triphosphates. These antimetabolites are recognized by the polymerase and then incorporated into the growing viral DNA chain. This results in chain termination and stops replication, as summarized in Figure 11.8. Points of structural modification of the nucleosides include the base and the ribose unit. Virtually all the early nucleoside antiviral agents maintain the 5′-hydroxyl group, and many are 2′- or 3′-deoxy analogs of naturally occurring nucleosides.

The earliest nucleoside antiviral drug, idoxuridine, containing a modified base (Figure 11.9a) was approved in 1962. It was synthesized by William Prusoff as a potential anti-cancer agent, as were many antiviral nucleosides. It is used topically to

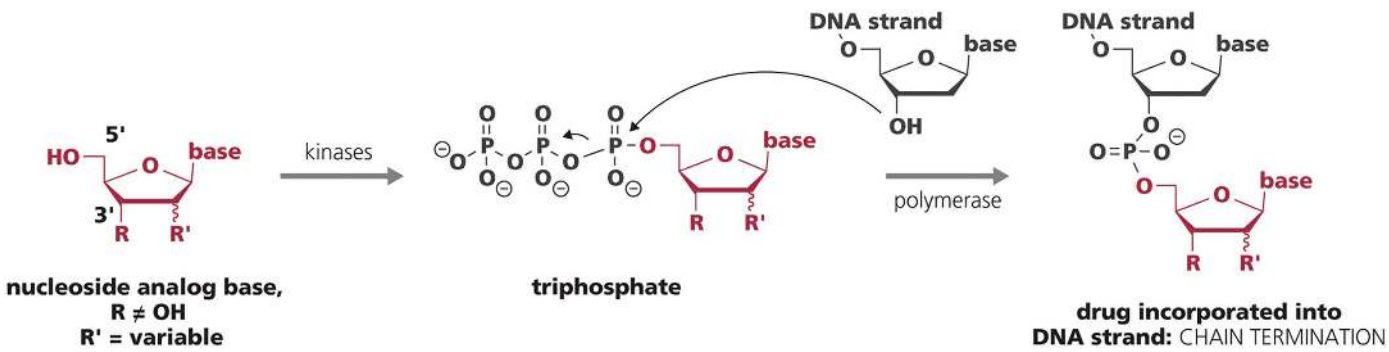
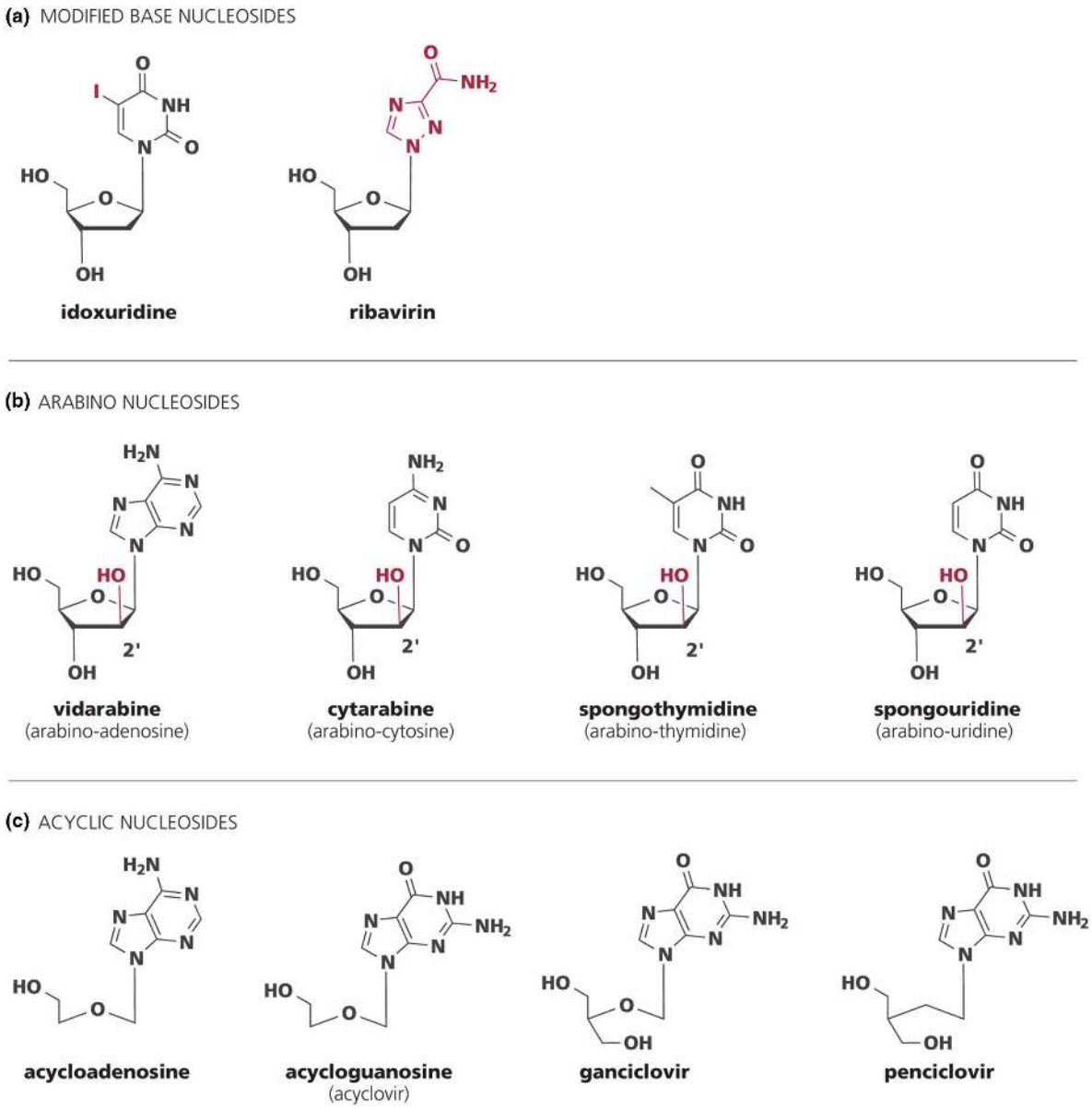


Figure 11.8 Incorporation of nucleoside antiviral drugs into DNA. The 5′-position of the nucleoside analog (red) is phosphorylated to the corresponding triphosphate by kinases. Attack by the 3′-hydroxyl of the DNA strand terminus incorporates the drug into the growing DNA strand. The R substituent cannot participate in subsequent polymerization reactions; therefore, chain termination results, and replication stops.



treat HSV eye infections, but its utility is limited by cardiotoxicity. Another nucleoside with a modified base is ribavirin, which was first synthesized in the 1970s. While it has broad-spectrum antiviral activity and was used in the 1980s to treat respiratory syncytial virus (RSV), since the 1990s, it has been used in combination with α-interferon to treat HCV infection. Recently, it has been supplanted by more effective drugs. However, ribavirin's mechanism is not well understood and appears to be more complex than solely polymerase inhibition.

The approval of idoxuridine was followed several years later by the approval of vidarabine (Figure 11.9b), an arabinose analog of adenosine that has activity against HSV and VSV. Its discovery can be traced to the 1950s, when spongothymidine and spongouridine were isolated from a Caribbean sponge by Werner Bergmann. Both

Figure 11.9 Early cyclic and acylic nucleoside antiviral drugs. (a) Idoxuridine, an iodo analog of 2′-deoxyuridine, and ribavirin, a triazole, contain modified bases (modifications are shown in red). (b) Vidarabine and cytarabine are arabino analogs based on the natural products spongothymidine and spongouridine (arabino OH configuration is shown in red). (c) Acycloadenosine, acycloguanosine (acyclovir), ganciclovir, and penciclovir are acyclic derivatives modified at the carbohydrate moiety.

natural products differ from thymidine and uridine at the 2′-stereogenic center by having the arabinose rather than the ribose configuration. Although modified bases had been prepared prior to this, activity of these compounds as polymerase inhibitors or antiviral agents gave the first indication that modification of the carbohydrate could also lead to active analogs. This idea was the basis for the synthetic analog vidarabine, as well as cytarabine (arabinocytosine), which entered clinical use in the 1960s.

While the earliest synthetic nucleosides either lacked the 2′-hydroxyl group or had inverted configurations, further modification of the carbohydrate by opening the ring generated acyclic nucleosides, an important structural motif. Compounds such as these were originally investigated not as antiviral agents but as substrates for adenosine deaminase, an enzyme in the purine biosynthesis pathway that converts adenosine to inosine. During these studies, it was found that, of the three hydroxyl groups present on ribonucleosides (2′, 3′, and 5′), only the one at the 5′-position was required for recognition by the enzyme. This observation led Howard Schaeffer at SUNY–Buffalo to synthesize the acyclic analogs acycloadenosine and acycloguanosine (acyclovir) in 1971, which were substrates for adenosine deaminase (**Figure 11.9c**).

Acyclovir was then investigated by Elion and Hitchings in 1977; it was found to have antiviral activity against HSV-1-infected cells and little toxicity toward uninfected cells. This selectivity is based on the compound's being selectively monophosphorylated by the virally encoded thymidine kinase, but not by mammalian kinases, and thus being converted to the active triphosphate polymerase substrate only in virally infected cells. As the triphosphate, it also selectively inhibits the viral DNA polymerase over host polymerases. Acyclovir, one of the most successful nucleoside antiviral agents, is used primarily to treat HSV and VZV infections. Other acyclic nucleosides include ganciclovir and penciclovir, which both contain a guanine base.

While the nucleoside drugs developed in the 1970s and early 1980s were successful antiviral agents, none was effective against the HIV infection that leads to AIDS and emerged as a pandemic in the mid-1980s. The first drugs to treat AIDS were not approved until the late 1980s; they were nucleoside analogs that targeted the enzyme reverse transcriptase and are known as nucleoside reverse transcriptase inhibitors, or NRTIs. HIV is an RNA-based retrovirus, which relies on its reverse transcriptase (RT) to catalyze the transcription of viral RNA to proviral DNA. Reverse transcriptase is a type of polymerase, and its mechanism was described in Chapter 7. In theory, treatment with an RT inhibitor immediately after exposure could prevent the initial formation of viral DNA and thereby prevent permanent infection. Indeed, this rationale is the basis for post-exposure prophylaxis (PEP), which is used to treat individuals who may have been exposed, for example, via occupational exposure to blood of an infected individual. However, RT inhibition is also effective once permanent infection is established, since it will prevent the formation of new viral particles. Unlike most host polymerases, however, the HIV RT does not possess a repair function, so that when an error in transcription is made, it persists. The lack of a repair mechanism allows NRTIs to be quite effective. These nucleoside analogs, like those described earlier, are considered competitive inhibitors/substrates and act at the active site. Despite the availability of X-ray structures of reverse transcriptase, the design of inhibitors is often based more on an empirical approach, since the molecule needs to be recognized and acted upon by multiple enzymes (kinases and RT) to be active. The considerable historical work in the area of nucleoside analogs that inhibited related polymerases or other DNA processing enzymes as anti-cancer agents was the basis of much of the earliest work on RT inhibitors.

In 1987, azidothymidine (AZT) became the first approved treatment for HIV/AIDS (**Figure 11.10**). It was originally synthesized in the 1960s as a potential anti-cancer drug and had been reported to be active against a retrovirus. It was one of many nucleosides screened for activity against HIV-infected cells. Its development relied on a collaboration between the pharmaceutical company Burroughs Wellcome and the US National Cancer Institute (NCI), based on the historical interest of Burroughs Wellcome in antiviral nucleoside analogs and the specialized capabilities of NCI for antiviral testing and access to patients. The approval of AZT occurred in record time and was based on data from only one clinical trial. The lack of HIV/AIDS animal models, as well as the urgent medical need, allowed a minimal development time. The

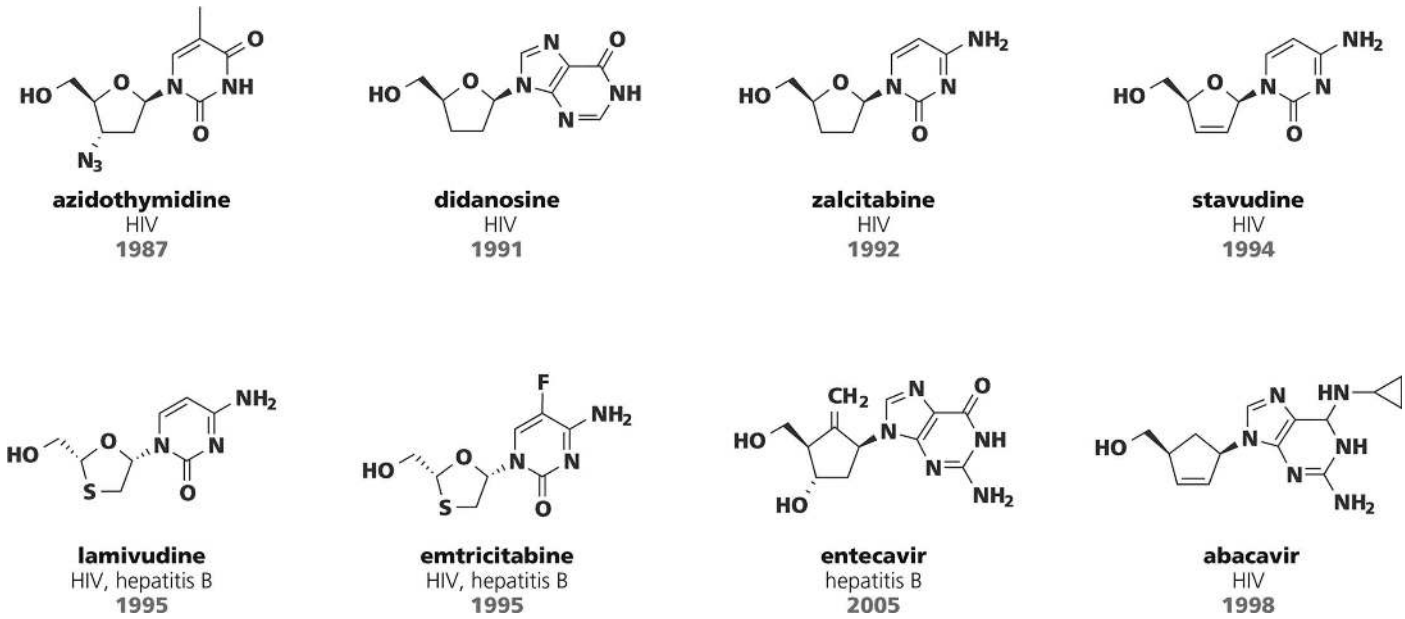


Figure 11.10 HIV nucleoside reverse transcriptase inhibitors and other nucleoside antiviral agents. Structures of NRTIs and the year of their approval are shown. Azidothymidine, didanosine, zalcitabine, and stavudine are structurally close to the natural nucleoside, with minor modifications on the ribose unit. Lamivudine, emtricitabine, and entecavir contain significant changes to the ribose; abacavir contains a modified base and unsaturated ribose. All compounds retain a 5′-OH group that is phosphorylated and incorporated into the growing nucleic acid chain, resulting in chain termination.

sole clinical trial was stopped after only 19 weeks, when data showed that the treatment group was much healthier than the placebo group, and it was deemed unethical to continue to administer a placebo to patients. Along with AZT, a number of other nucleoside analogs have been found to be effective against HIV: Didanosine (ddI), zalcitabine (ddC), and stavudine (d4T) had all been synthesized in the 1960s as potential anti-cancer agents and were among the earliest drugs approved for the treatment of HIV/AIDS. The enantiomer of the thiosugar analog of zalcitabine, lamivudine, and a base-modified derivative, emtricitabine, are also effective against HIV as well as against hepatitis B infection. Many of these drugs suffer from poor bioavailability and require frequent dosing.

Carbocyclic nucleoside analogs, where the ring oxygen is replaced with a carbon, have also been investigated. Two that have been successful are entecavir and abacavir, which are used to treat chronic hepatitis B and HIV infections, respectively. Lamivudine, emtricitabine, entecavir, and abacavir all contain significant structural modifications to the furanose ring. However, all NRTIs (see **Figure 11.9**) contain a 5′-hydroxyl group, are activated to the triphosphate, are recognized by RT, and act as chain terminators.

11.4.4 Prodrugs Improve the Properties and Effectiveness of Nucleosides

As shown in **Figure 11.8**, nucleoside drugs require a complex series of transformations to generate the active triphosphate. This requirement, coupled with some of their physical properties, requires high dosing; efforts to overcome the poor bioavailability have led to the application of prodrug strategies. One approach has focused on ester prodrugs of acyclic nucleosides. Valacyclovir, the valine ester of acyclovir, has approximately 4-fold better oral bioavailability than acyclovir (**Figure 11.11a**) and is used to treat various herpes infections, as is valganciclovir. Famciclovir is a diacetate ester of an all-carbon acyclic ribose unit and is also modified at the base. Hydrolysis of the esters and oxidation of the purine ring generate the acyclic guanosine nucleoside penciclovir. Due to its poor oral bioavailability, penciclovir is only used topically, but famciclovir, with 77% oral bioavailability, is used to treat both VZV and HSV infections. In these prodrug examples, the esters are cleaved by host enzymes to reveal the 5′-OH.

In addition to improving oral bioavailability, nucleoside prodrug strategies have also been applied to circumvent the first phosphorylation step on the pathway to the active triphosphate. This step, carried out by kinases, is considered to be the most

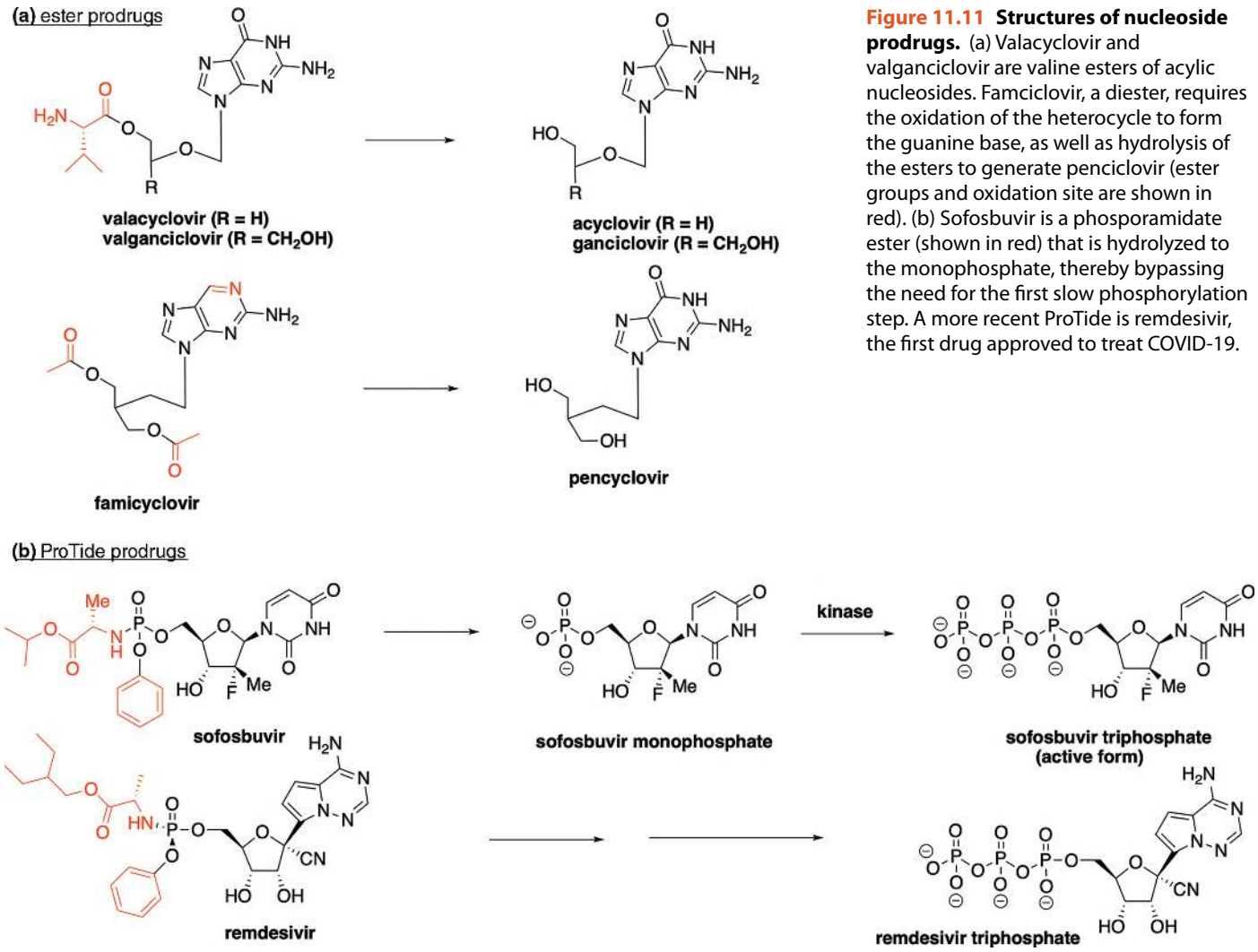


Figure 11.11 Structures of nucleoside prodrugs. (a) Valacyclovir and valganciclovir are valine esters of acylic nucleosides. Famciclovir, a diester, requires the oxidation of the heterocycle to form the guanine base, as well as hydrolysis of the esters to generate penciclovir (ester groups and oxidation site are shown in red). (b) Sofosbuvir is a phosporamidate ester (shown in red) that is hydrolyzed to the monophosphate, thereby bypassing the need for the first slow phosphorylation step. A more recent ProTide is remdesivir, the first drug approved to treat COVID-19.

discriminating, and therefore, efforts to design analogs that would bypass it are an active area of research. The monophosphate cannot be administered directly due to its highly charged nature and, therefore, lack of permeability, as well as its relative instability.

The prodrug sofosbuvir (see **Figure 11.11b**) is approved for the treatment of HCV infections, and its triphosphate inhibits the NS5B protein, which is HCV's viral RNA polymerase. Sofosbuvir is an example of the use of a pronucleotide (ProTide) strategy, developed by Christopher McGuigan. Hydrolysis *in vivo* of the aryloxy phosphoramidate triester generates the monophosphate, which is then converted by kinases to the active triphosphate. The enzymes responsible for its activation reside in the liver, the site of HCV infection, and this also contributes to its efficacy. Sofosbuvir was invented at Pharmasset in 2007 and proceeded rapidly to clinical trials in 2010. Its remarkable activity in combination with other HCV drugs (it is curative) led to rapid FDA approval in 2013 to treat hepatitis C infection. Remdesivir is another ProTide, originally investigated to treat Ebola virus infections but not approved. It was found to be effective against the SARS-CoV-2 virus and was the first drug approved to treat COVID-19 cases (see Case Study 2, Chapter 4).

A related strategy incorporates a phosphonate in place of the monophosphate; in addition to bypassing the first kinase step, this modification affords improved stability. The acyclic phosphonate cidofovir, due to its charged nature, is administered intravenously; it is used to treat CMV retinitis. It is converted to the triphosphate analog, which is a substrate for viral DNA polymerase (**Figure 11.12**). Phosphonate prodrugs have also

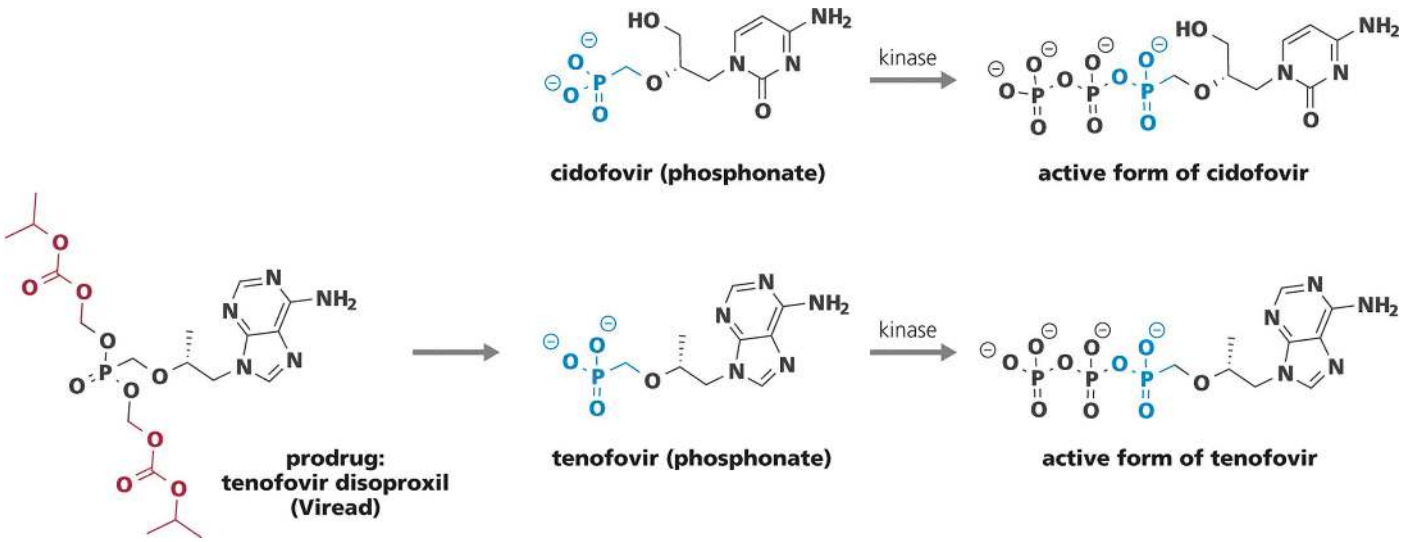


Figure 11.12 Examples of antiviral nucleoside phosphonates. Cidofovir and tenofovir are nucleoside phosphonates (phosphonate moiety is shown in blue). The phosphonates are more stable than the monophosphate but are still recognized by kinases to generate the active triphosphate analog. Tenofovir disoproxil is an ester prodrug of tenofovir; the prodrug moiety hydrolyzed is shown in red.

been developed to improve physical properties. Viread is a phosphonate ester prodrug of tenofovir. The esters are cleaved in the gut to generate the monophosphonate, which, unlike the monophosphate, is stable to esterases. This monophosphonate is a substrate for the next two kinases to afford the triphosphate analog, which is an inhibitor of HIV reverse transcriptase (see Figure 11.12). Although it was developed to treat HIV, Viread has also been approved for the treatment of hepatitis B infections.

11.4.5 Some HIV Drugs That Target Reverse Transcriptase Bind at Allosteric Sites

In addition to being inhibited by nucleosides, HIV reverse transcriptase is also sensitive to allosteric inhibition by nonnucleoside drugs. In the early search for drugs to treat AIDS, many compound libraries were screened for anti-RT activity, leading to hits from unexpected drug classes. Nevirapine is one such example. Screening for inhibitors of RT identified a hit originally synthesized as a potential muscarinic antagonist in the 1960s. The hit was optimized by chemists at Boehringer Ingelheim for potency and physical properties and approved to treat AIDS in 1996. Other nonnucleoside reverse transcriptase inhibitors (NNRTIs) were also developed from screening programs, leading to additional drugs in this class. At Upjohn, 1,500 compounds of distinct structural classes in their compound library were chosen for screening based on computational analysis for dissimilarity. Screening against HIV-1 RT led to 100 compounds that were active and, after optimization, to delavirdine. At DuPont, a similar screening of their compound library identified an unstable hit that was optimized, resulting in efavirenz. Rapid emergence of resistance to these NNRTIs was noted and was likely due to the molecules acting at an allosteric site that was not essential for enzymatic activity. Figure 11.13 shows the first-generation NNRTIs that are approved to treat HIV/AIDS.

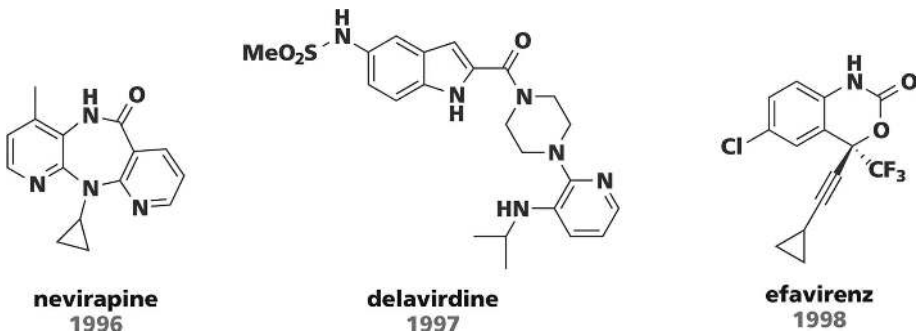


Figure 11.13 Structures of first-generation nonnucleoside reverse transcriptase inhibitors and year of approval. The NNRTIs nevirapine, delavirdine, and efavirenz were derived from hits identified from library screening for activity against HIV RT and act at an allosteric site.

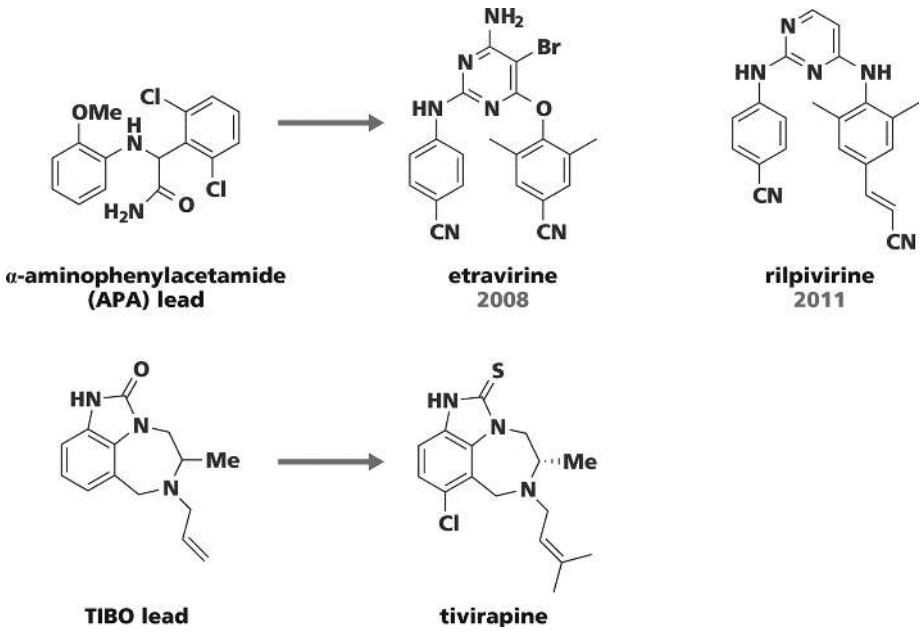


Figure 11.14 Second-generation nonnucleoside reverse transcriptase inhibitors. Etravirine and rilpivirine are approved to treat HIV and were derived from an α-aminophenylacetamide (APA) lead. Tivirapine was based on the tetrahydroimidazobenzodiazepinone (TIBO) lead and advanced to clinical trials.

Efforts to identify second-generation NNRTIs focused on screening and optimizing inhibitors that were active against multiple resistant strains of HIV. The structures of the newer NNRTIs etravirine and rilpivirine are considerably more flexible than the first-generation compounds (Figure 11.14). It is believed that the flexible nature of these compounds allows them to adapt to mutations in the active site while still maintaining their inhibitory effects. These compounds can be traced to early screening efforts at the Rega Institute for Medical Research. They screened 600 compounds from Janssen Pharmaceuticals and identified tetrahydroimidazobenzodiazepinone (TIBO) and related thione analogs as well as α-aminophenylacetamide (α-APA) scaffolds as leads that inhibited the replication of HIV-1. Mechanistic studies identified RT as the target of the compounds. One of the TIBO leads was investigated further in a search for compounds with activity against resistant strains, and this effort ultimately resulted in tivirapine, which progressed to clinical trials. The APA leads evolved into etravirine and rilpivirine, approved in 2008 and 2011, respectively (Figure 11.14). The development of rilpivirine from the APA lead is outlined in Case Study 2 in this chapter.

The NNRTIs bind at a common site on the enzyme ~10–15 Å away from the catalytic site, but precisely how they inhibit the enzyme is not well understood. The flexibility and large size of the enzyme (a heterodimer consisting of 66 and 51 kDa subunits), its multiple components, and its relative insolubility have made biophysical studies challenging. X-ray crystallography of HIV RT shows a structure like a right hand, with fingers, a thumb, and a palm region. One common theory is that NNRTIs act as a molecular wedge and lock the thumb region into an open conformation (characterized as an "arthritic thumb") (Figure 11.15), resulting in a change in orientation of the catalytic triad at the active site. It should be noted that this hypothesis is still

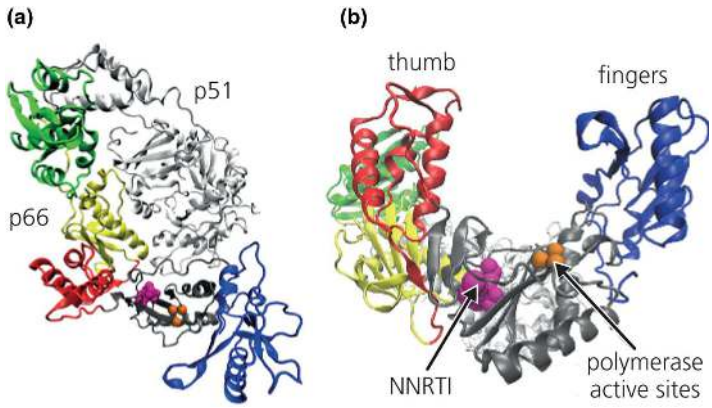


Figure 11.15 Structure of efavirenz bound to HIV RT. (a) HIV-1 RT is a heterodimer consisting of p66 (multicolored) and p51 (light gray) subunits with an RNase H domain (green, not labeled). (b) View along HIV-1 RT from the polymerase active site illustrates the resemblance of the p66 subunit to an open hand: The fingers are shown in blue, the palm in dark gray, the thumb in red, the linker in yellow, and the RNase H domain in green. In both figures, the location of the polymerase active site triad is depicted in orange, and the location of the bound NNRTI efavirenz is shown in magenta.

Source: From Wright, D. W.; Sadiq, S. K.; De Fabritiis, G.; and Coveney, P. V. (2012). *J Am Chem Soc* 134:12,885–12,888. With permission from the American Chemical Society.

controversial, and exactly how these allosteric modulators result in inhibition is not completely understood.

11.4.6 Viral Protease Inhibitors Have Been Developed by Using Structure-Based Drug Design

One of the most successful applications of structure-based drug design has been the development of HIV protease inhibitors. The role of the protease enzyme (PR) of HIV is to cleave the Gag-Pol fusion protein into smaller functional and structural viral proteins, such as reverse transcriptase and integrase, that are essential for the replication, function, and maturation of infectious viral particles. Sequencing of the HIV genome led to the conclusion that the virus encoded an aspartyl protease (see Section 7.4.1). Target validation was based on observations that mutating the putative active site aspartate residues resulted in immature, non-infectious viral particles. Drug discovery was accelerated by the considerable medicinal chemistry expertise accumulated in designing transition state mimetics (described in Chapter 4) for other aspartyl proteases, such as renin, which was a target for treating cardiovascular disease.

Many of the HIV protease inhibitors were designed on the basis of specific cleavage sites that are recognized by the enzyme. Phenylalanine-proline (Phe-Pro) and tyrosine-proline (Tyr-Pro) were two such recognition sites not commonly cleaved by human proteases. By the use of a substrate-like scaffold and the incorporation of transition state mimetics, a series of inhibitors of HIV protease was developed. An X-ray structure of the enzyme was solved in 1989, and it showed a homodimer with C_2 symmetry. This information, combined with knowledge acquired from the design of other aspartyl protease inhibitors, rapidly led to a number of transition state analogs being advanced as HIV protease inhibitors (**Figure 11.16**). Many of these (such as saquinavir, indinavir, and nelfinavir) retain vestiges of the original Phe-Pro peptide substrate on which the design was based. Some, such as ritonavir and lopinavir, were developed from C_2 symmetrical leads designed to inhibit the enzyme and still retain some of that symmetry, as evidenced by the presence of two benzyl groups flanking the transition state mimetic. One exception is tipranavir, which was based on hits from high-throughput screening.

Many of the early compounds suffered from very poor oral bioavailability and required frequently administered high doses. These poor pharmaceutical properties are likely due to the molecules' high molecular weights, the presence of multiple peptide bonds, and their highly polar nature. (One attribute of ritonavir is that it is also an inhibitor of cytochrome P450 metabolizing enzymes, resulting in higher drug levels of itself and other agents that are co-administered; see discussion on Paxlovid). Furthermore, like the reverse transcriptase inhibitors, resistance toward them soon developed. Newer compounds, such as tipranavir and darunavir, are less prone to the development of resistance.

The broad activity of darunavir against resistant strains is based on a design strategy to maximize interactions with backbone amides rather than side chain moieties. It was known that drug-resistant strains undergo significant reorganization of protein side chains compared to native strains, but minimal movement in the protein backbone. As predicted, X-ray co-crystal results of darunavir bound to HIV protease showed that the two oxygen atoms in the bicyclic ether group made H-bonds with Asp29 and Asp30 in the main chain (**Figure 11.17**). In addition, polar and ionic interactions between the aniline nitrogen of the inhibitor and the amide in the main chain and the carboxylate oxygen of Asp30′ were also observed.

The success of the HIV protease inhibitors led to the search for inhibitors of other virus-specific proteases. The hepatitis C virus was identified as a unique virus in 1989, and cloning of the HCV genome identified the proteolytic enzyme HCV NS3, a serine protease, as a potential drug target. Lessons learned from drug discovery efforts in the HIV arena have been applied to HCV, with a number of drugs approved. The observation that the NS3 protease was subject to feedback inhibition by two of the N-terminal peptide products led to the design of early peptide leads for enzyme inhibition. One of these is the N-acylated hexapeptide shown in **Figure 11.18a**, which is a cleavage product from proteolysis of the substrate at a cysteine residue. Several peptide

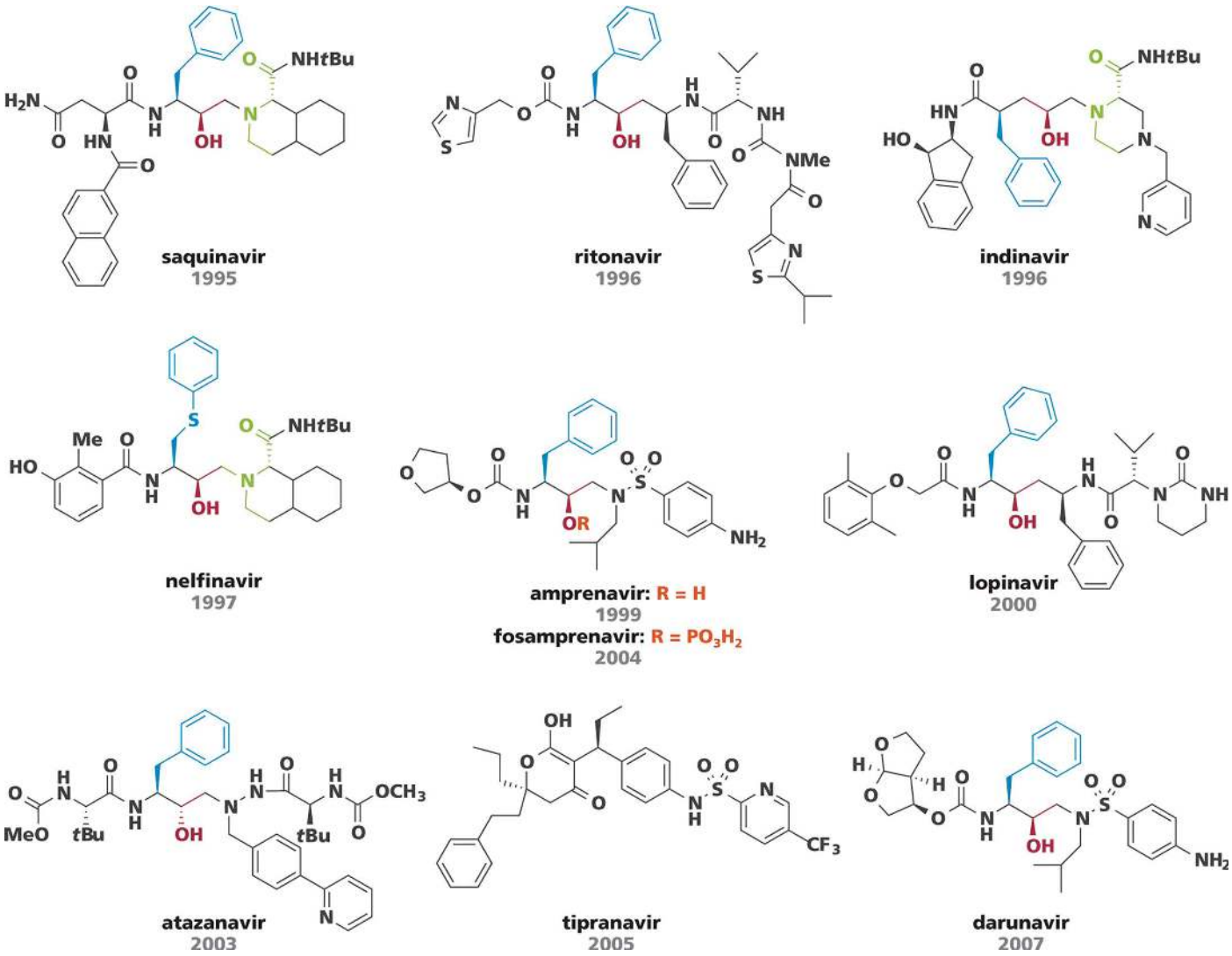
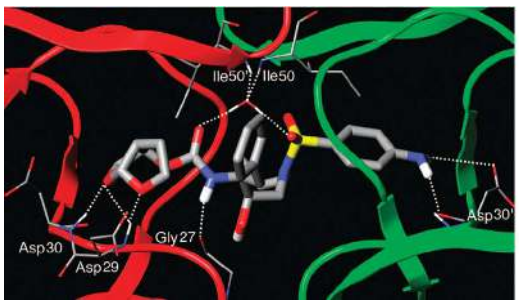


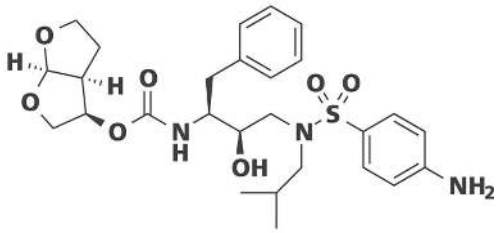
Figure 11.16 Structures of HIV protease inhibitors, in order of date of approval. All except tipranavir contain transition state mimetics (shown in red) and are based on mimicking the enzyme's natural substrate (fosamprenavir is a prodrug of amprenavir). In these cases, structural features of the Phe-Pro cleavage site are evident: Phe mimic in blue, and Pro mimic in green. The symmetry evident in ritonavir and lopinavir reflect the *C2* symmetry of the enzyme. The exception, tipranavir, was evolved from a high-throughput-screening hit.

analogs of this product were prepared as inhibitors, with two early leads shown in Figure 11.18a. Notably, the tetrapeptide lacking N-terminal residues and containing a vinylcyclopropane replacement for the P1 thiol was a potent enzyme inhibitor. Modification of these leads ultimately afforded the drugs boceprevir, telaprevir, and asunaprevir (Figure 11.18b). Common features based on the original peptide include the P3/P2 core based on the Val-Pro segment.

Boceprevir (Merck) and telaprevir (Vertex) were approved in 2011 to treat patients with the G1 genotype of HCV. Both contain a keto-amide warhead, with the active site serine forming a covalent bond to the ketone (see Section 7.4.1). Both drugs were given in combination with interferon and ribavirin and were effective in more than half of patients that were nonresponsive to the interferon/ribavirin combination alone. More recent efforts focused on drugs that can be given as a single oral dose without interferon, and these have supplanted telaprevir and boceprevir.

Newer drugs, approved since 2013, such as grazoprevir and the related glecaprevir and voxilaprevir, as well as simeprevir and paritaprevir (Figure 11.19), were developed by applying conformational constraint strategies to the linear peptides by tethering the binding elements to form a macrocycle (see Section 4.5.5). All these agents are given in combination with other drugs that target HCV-specific proteins. They are distinguished from telaprevir and boceprevir in that they do not require co-administration with interferon to be effective and have fewer side effects.



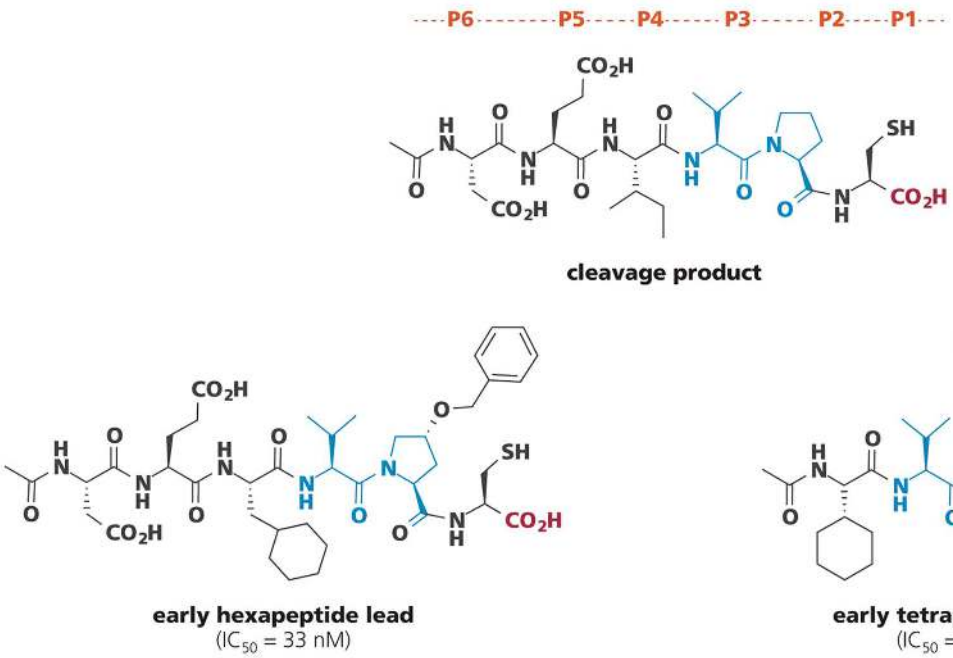


darunavir

Figure 11.17 Structure of darunavir bound to HIV protease. *Left*, X-ray crystal structure. Darunavir is shown in gray, with specific atoms colored: Oxygen is red, nitrogen is blue, sulfur is yellow, and specific hydrogens are white. Individual monomers of HIV protease indicated in red and green. Oxygen atoms of the bicyclic ether make interactions with backbone N-H groups of Asp29 and Asp30, and there is an ionic interaction between the protonated aniline nitrogen and the Asp30′ carboxylate, and as well as an H-bond to the backbone. As with nearly all HIV protease inhibitors, H-bonding between the inhibitor and the Ile50/Ile50′ residues of the flap is mediated by a water molecule. *Right*, structure of darunavir.

Source: Left, from Ghosh, A. K.; Chapsal, B. D.; Weber, I. T.; and Mitsuya, H. = (2008) *Acc Chem Res* 41:78–86. With permission from the American Chemical Society.

(a) EARLY PEPTIDE LEADS BASED ON CLEAVAGE PRODUCTS

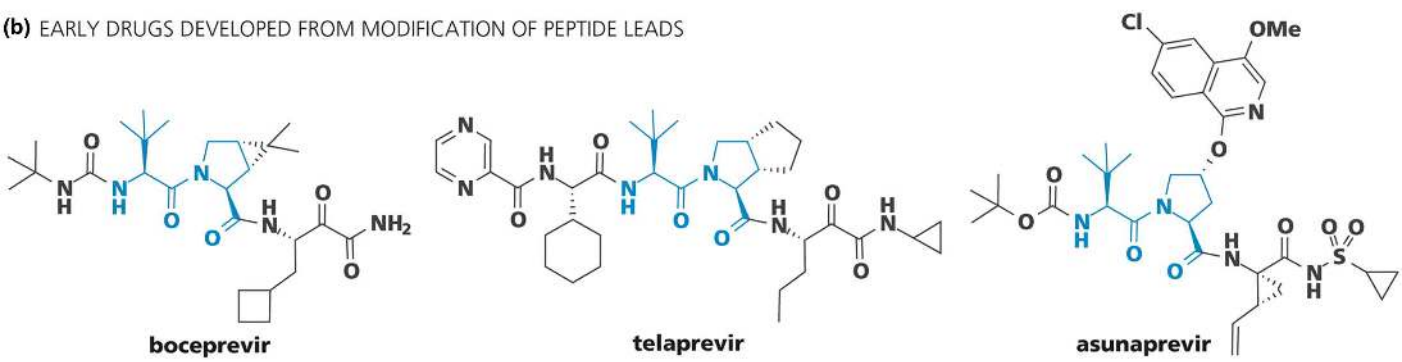


cleavage product

early hexapeptide lead
(IC_{50} = 33 nM)

early tetrapeptide lead
(IC_{50} = 13 nM)

(b) EARLY DRUGS DEVELOPED FROM MODIFICATION OF PEPTIDE LEADS



boceprevir telaprevir asunaprevir

Figure 11.18 Evolution of early HCV NS3 protease inhibitors. (a) The peptide was derived from cleavage products of the NS3 protease, which caused feedback inhibition of the enzyme. The hexapeptide and tetrapeptide leads were designed to mimic this cleavage product. (b) Modification of the tetrapeptide lead afforded the earliest HCV NS3 protease inhibitors boceprevir and telaprevir (both now withdrawn) and the later drug asunaprevir (approved in Japan). All maintain versions of the Val-Pro segment found in the original peptides (shown in blue), with the *tert*-butyl group replacing the isopropyl group.

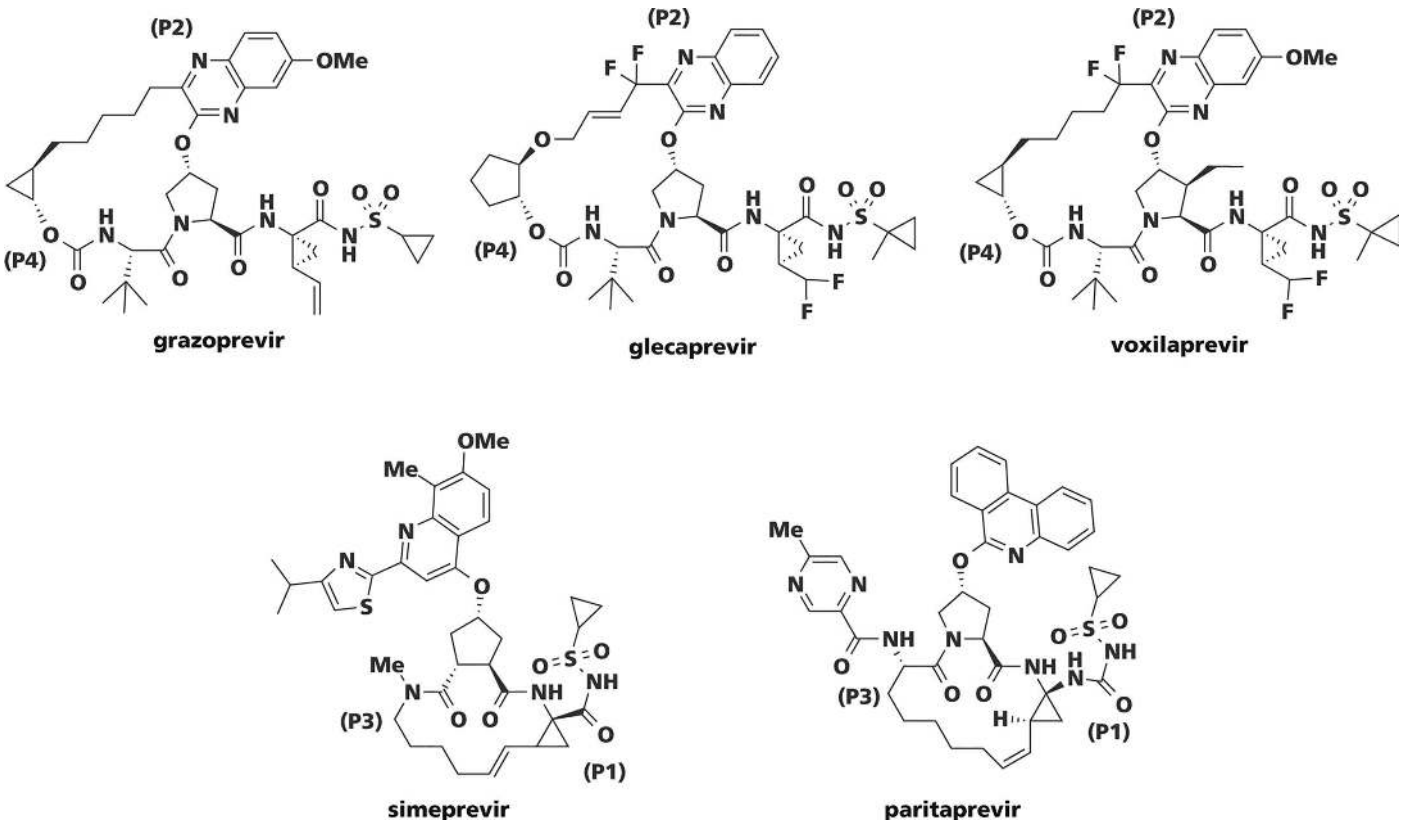


Figure 11.19 Structures of conformationally constrained macrocyclic HCV NS3 protease inhibitors. Grazoprevir, glecaprevir, and voxilaprevir resulted from linking between the P2 and P4 peptide positions; simeprevir and paritapevir are linked between P1 and P3 peptide positions. All contain the sulfonamide bioisostere for the carboxylic acid.

Two specific examples of application of the conformational constraint strategy are highlighted in Figure 11.20. BILN 2061, developed at Boehringer Ingelheim, was a more potent enzyme inhibitor than the linear tetrapeptide (IC_{50} of 3 nM vs. 13 nM for the tetrapeptide; Figure 11.20a). This compound was designed by linking the P1 and P3 regions of the lead and advanced to clinical trials. Docking of the linear tetrapeptide into the enzyme active site showed binding of the His528 residue in the protease to the carbonyl of the P4 amide and Gln526 binding to the P2 *O*-arylquinolone (Figure 11.20a). Because of these important interactions, chemists at Merck proposed to constrain those two binding elements by linking P2 to a truncated P4, as shown in Figure 11.20b. As with the P1–P3 constraint, the P2–P4 constraint developed into a compound with improved enzyme affinity compared to the acyclic peptide; incorporation of a sulfonamide bioisostere ("second series") for the P1 carboxylic acid provided a compound with sub-nanomolar affinity. Optimization occurred by introducing branching of the P3 side chain (to a tert-butyl moiety) and removal of the phenyl substituent on the quinolone. This advanced lead resulted in a compound with similar potency that achieved high concentration in the target organ (liver).

This conformational constraint strategy ultimately led to the development of MK-5172 (grazoprevir), which received FDA approval when used in combination with the NS5A inhibitor MK-8742 (elbasvir, described in the next section). This combination, like other combinations of HCV NS3 and NS5A inhibitors, was awarded breakthrough designation by the FDA, on the basis of Phase II results, which showed a 100% decrease in viral load. This co-therapy was also highly effective in the particularly challenging patient population co-infected with HIV and HCV, with 90% of treated patients showing a decrease in viral load. Several other combinations of HCV NS3 and NS5A inhibitors have shown similar results.

The SARS-CoV-2 3CL main protease represents another protease that has been successfully targeted. This is a cysteine protease that is highly conserved in coronaviruses and is absent in humans. With the proven success of HIV protease and HCV protease inhibitors in hand, chemists at Pfizer had investigated coronavirus protease inhibitors during the SARS-CoV outbreak of 2002 and identified PF-00835231, a

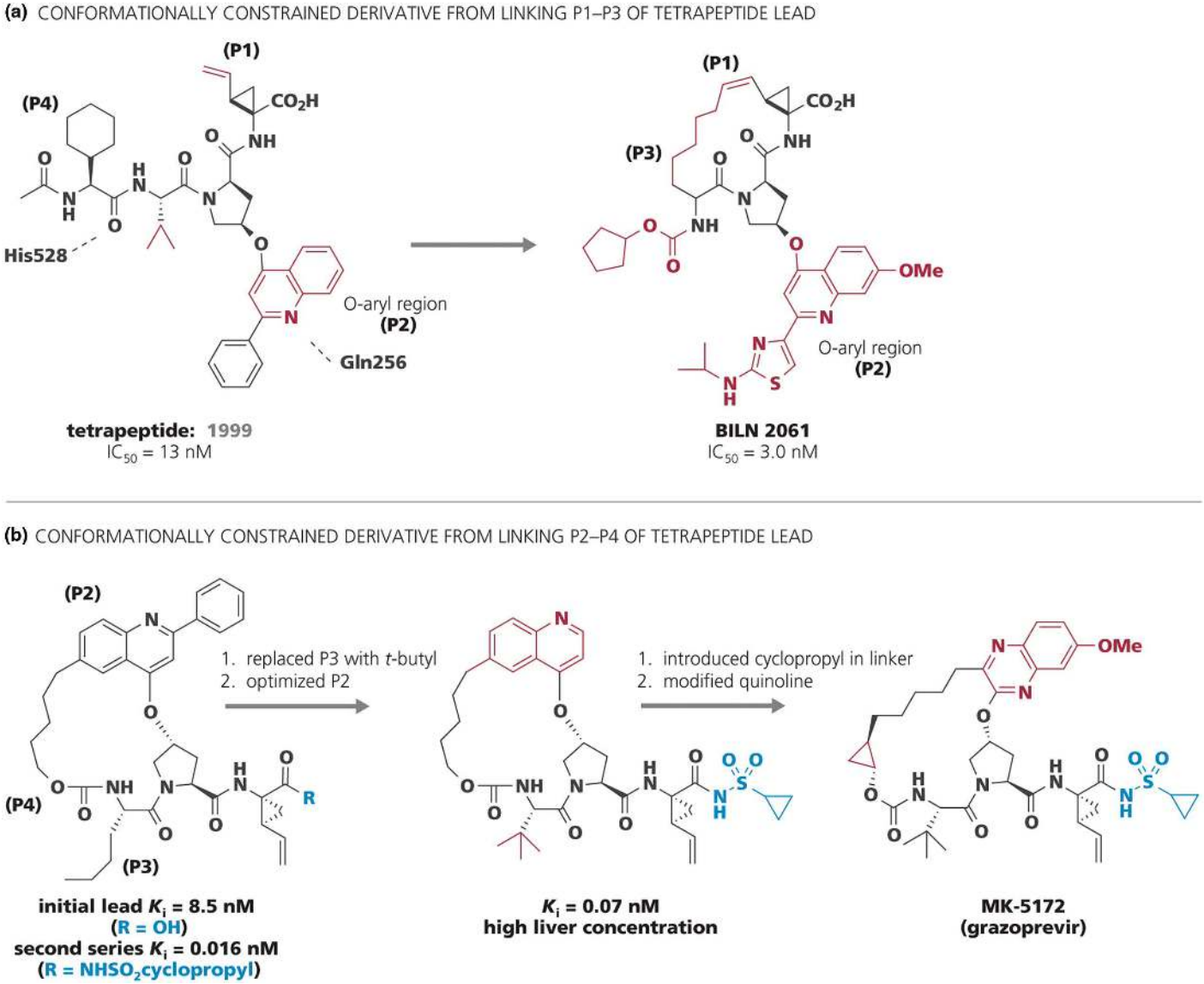
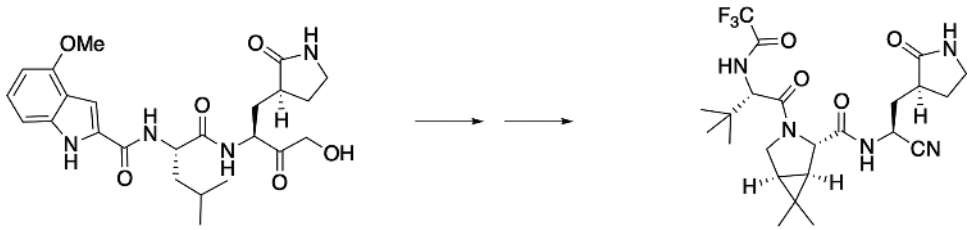


Figure 11.20 **Tethering strategy for design of HCV NS3 protease inhibitors.** (a) Linking the P1 and P3 components of the tetrapeptide lead, upon further optimization, led to BILN 2061, a compound that entered clinical trials (modifications shown in red). (b) Linking the P2 and P4 regions afforded an initial lead, and the incorporation of a sulfonamide bioisostere afforded a second series, both with significantly greater affinity compared to the acyclic peptide. In the sulfonamide series (in blue), optimization of the P2 group and changing the P3 group to a tert-butyl gave an optimized lead with high liver concentration (modifications in red). The sulfonamide series was further optimized by the introduction of a cyclopropyl group in the linker, and quinolone modification, ultimately leading to grazoprevir.

peptidomimetic based on the substrate for the coronavirus 3CL protease. Although PF-00835231 did not proceed, due to the containment of the 2002 SARS outbreak, it served as a lead for nirmatrelvir, which would become the active component of the Paxlovid combination used to treat COVID-19 (Figure 11.21). Ritonavir is the other component, whose purpose is to boost exposure levels of nirmatrelvir based on its inhibition of cytochrome P450 enzymes. Modifications to PF-008352 included reducing several hydrogen bonding sites and the introduction of rigidity, resulting in an orally active antiviral, which led to the approval of Paxlovid very rapidly, roughly one year after its nirmatrelvir's synthesis.



PF-00835231 nirmatrelvir

Figure 11.21 Structures of lead compound PF-00835231 and drug nirmatrelvir.

11.4.7 Inhibitors of the Hepatitis C Virus NS5A Replication Complex Exhibit Antiviral Activity

In in the early 2000s, a number of organizations initiated high-throughput screening campaigns using cell-based replicon assays to identify inhibitors of viral replication. Replicons consist of RNA from HCV that is non-infectious but capable of self-replicating in a cell. These cell-based assays allowed screening against the entire virus, rather than specific enzymes, and were an important advance for HCV drug discovery because HCV-infected cells are difficult to grow. This phenotypic screening, using the most prevalent genotypes of hepatitis C infections (GT1a and GT1b), identified compounds whose target was, at first, unknown but later was determined to be NS5A and led to starting points of an entire class of anti-HCV agents. The eventual identification of HCV NS5A as the target was based on assaying early leads in the development of one of the drugs, daclatasvir, in cells infected with drug-resistant mutant virus (see Case Study 1 in this chapter). The HCV NS5A protein complex is essential for viral replication; however, its specific function is unknown. Daclatasvir, developed at Bristol-Myers Squibb, was approved by the FDA in 2015 (**Figure 11.22**) in combination with sofosbuvir. Several other HCV NS5A inhibitors, including ledipasvir, elbasvir, ombitasvir, velpatasvir, and pibrantasvir, have also been approved, all in combination with other anti-HCV agents.

During the course of discovery of daclatasvir, the importance of symmetry of the molecules was noted. HCV NS5A exists as a symmetrical dimer, and therefore, symmetry of inhibitors is most likely related to their binding in the target. Several models of daclatasvir and other inhibitors binding to HCV NS5A have been developed that are based on X-ray structures of a portion of the protein; however, co-crystal structures have yet to be solved. Ledipasvir was optimized with a goal of high potency combined with a long half-life, and both symmetrical and unsymmetrical leads were investigated. An unsymmetrical imidazole/benzimidazole series ultimately led to ledipasvir, the first NS5A inhibitor approved. Elbasvir was elaborated from a hit derived from HTS of an in-house library. Ombitasvir was derived from compounds containing chiral pyrrolidines to match the symmetry of the target, and velpatasvir was developed to be a pan-genotype inhibitor, improving on the spectrum of the first-generation NS5A inhibitors. Pibrantasvir is highly potent (picomolar) against all the common HCV genotypes as well as against common resistant mutants. These drugs are used in combination with HCV NS3 protease inhibitors, such as grazoprevir, or the NS5B polymerase inhibitor sofosbuvir. These combinations have revolutionized the treatment for hepatitis C infection, which previously relied on a combination of interferon and ribavirin, which was much less effective and had severe side effects.

11.4.8 Drugs to Treat Retroviruses May Target the Viral Integrase

Retroviruses, such as HIV, utilize a viral integrase (IN), which catalyzes the incorporation of viral DNA into the host genome through a complex multi-step process. The newly synthesized viral DNA (formed by the action of reverse transcriptase) is

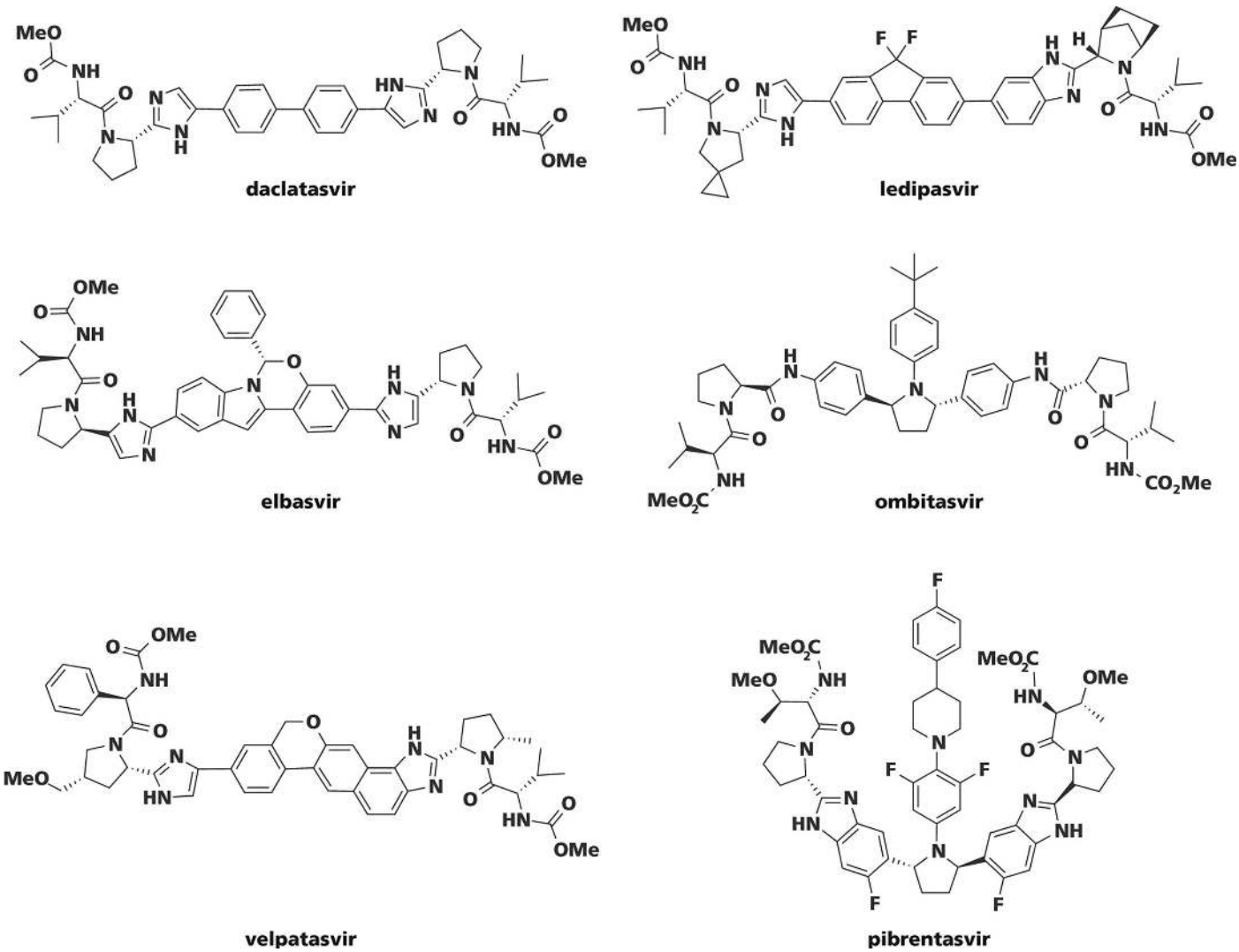


Figure 11.22 Examples of HCV NS5A inhibitors. These drugs inhibit the viral replication complex, and all have elements of symmetry in common. All are administered with other anti-HCV agents. The most recently approved agents, velpatasvir and pibrantasvir, are effective against multiple resistant strains of HCV.

processed at the 3′-end and forms a pre-integration complex. The complex migrates to the nucleus and is inserted into the host DNA by a metal-mediated transesterification reaction through a mechanism like the one shown in **Figure 11.23**. Two metal cations (M_A and M_B) at the enzyme active site are essential for the reaction and serve to activate and properly orient the nucleophile (3′-hydroxyl group of the viral nucleic acid) as well as to stabilize the pentacoordinate transition state, once formed. Following initial transesterification, the gaps between the viral and host DNA are closed by host polymerases (not shown).

Understanding the role and mechanism of IN suggested that screening for inhibitors of the enzyme might yield useful starting points for the development of antiviral agents. For example, at Merck, 250,000 compounds were screened, and the hits all contained a diketo function as well as an aromatic moiety. They were characterized as active site inhibitors that were competitive with the host DNA substrate. Other groups identified compounds with similar structural motifs. Optimization of hits initially led to two drugs: the *N*-methylpyrimidone raltegravir from Merck and the quinolone-derived elvitegravir from Gilead (**Figure 11.24**). Structural studies indicate that the coplanar oxygen atoms present in raltegravir and elvitegravir coordinate to the two metal (Mg^{2+}) ions in the enzyme active site and also bind to the viral DNA substrate. Second-generation integrase inhibitors, now called INSTIs (integrase strand transfer inhibitors), include dolutegravir, bictegravir, and cabotegravir. These are now used as front-line therapy to treat AIDS, with excellent efficacy and good tolerability. Like all

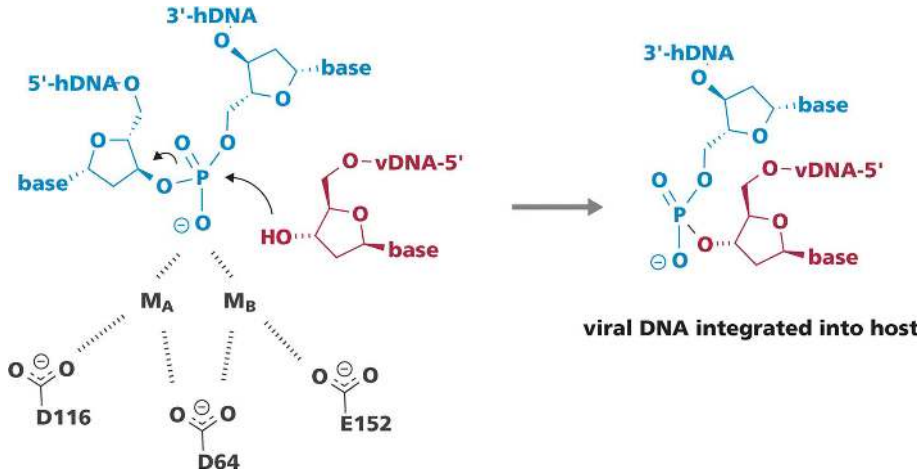


Figure 11.23 Representation of the reaction catalyzed by HIV integrase. The 3′-hydroxyl of viral DNA (vDNA, shown in red) adds to the phosphodiester of the host DNA (hDNA, shown in blue), which is coordinated through two metal ions (M_A and M_B) to acidic side chains of HIV integrase (D116, D64, and E152). The viral DNA becomes integrated into the host DNA after a further round of integration mediated by host polymerases (not shown), and a permanent infection results.

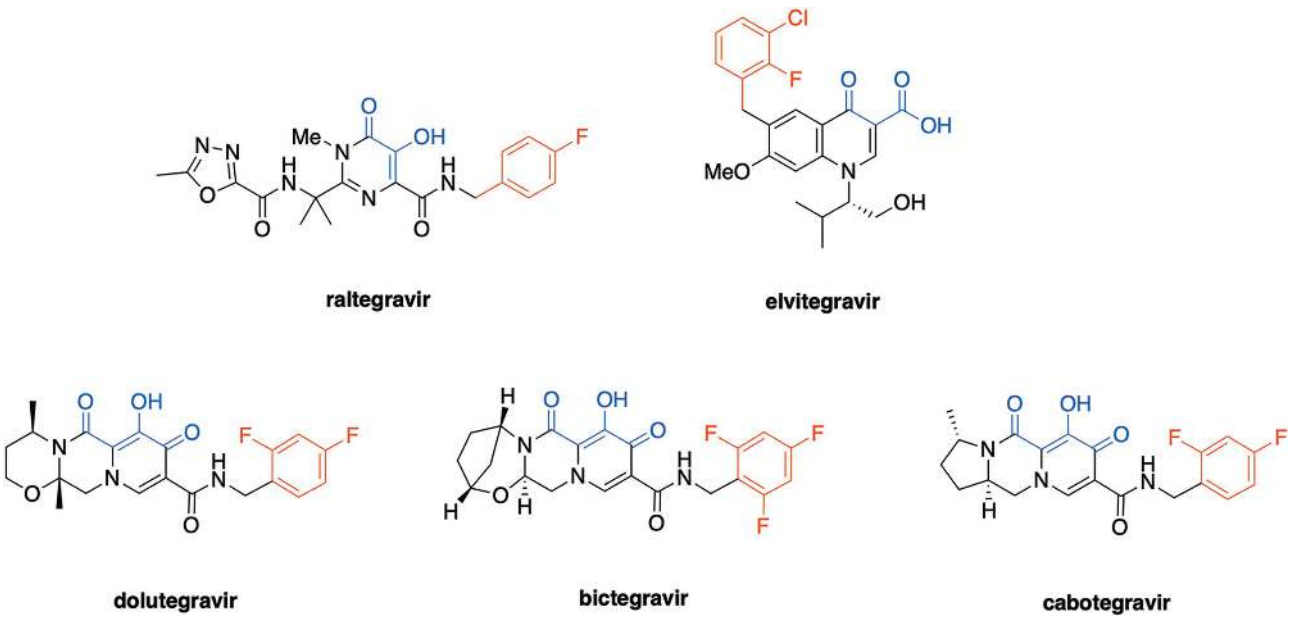
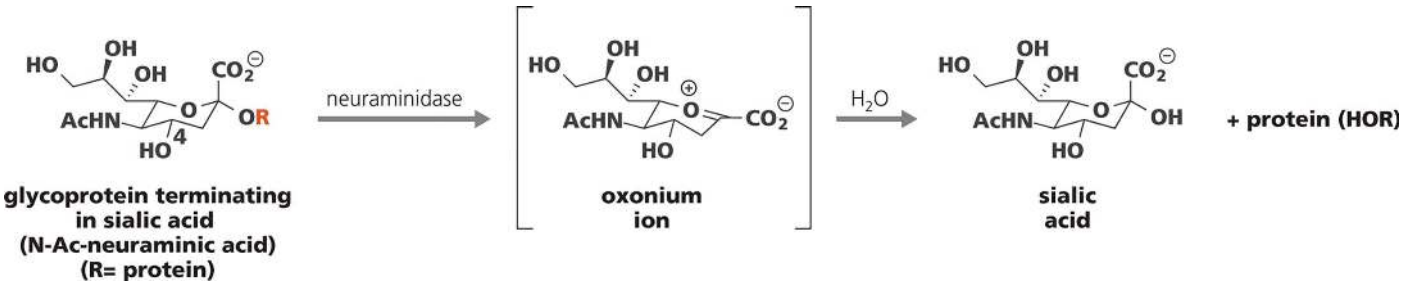


Figure 11.24 Structures of viral integrase inhibitors. Common pharmacophore elements are shown in blue (metal binding) and red (aromatic and hydrophobic).

other modern drug regimens, these compounds are used in combination with other anti-HIV drugs.

11.4.9 Influenza Drugs' Targets Include Viral Budding and an Endonuclease

Viral budding is an essential process for a virus to spread from one cell to another, and therefore, targeting this step is a viable therapeutic approach. To date, the only drugs that inhibit this process target influenza virus, an enveloped single-stranded RNA virus. Its viral surface glycoproteins are made up of hemagglutinin (H) and neuraminidase (N), and serotypes are characterized by antibody responses to H and N. For instance, the H5N1 strain causes avian flu, while the seasonal flu is primarily from H1N1, H1N2, and H3N2 variants. While the primary treatment is prevention of infection by vaccination, the most widely used small-molecule drugs are neuraminidase inhibitors. As mentioned earlier, the uncoating inhibitors amantadine and rimantadine (see **Figure 11.5**) are no longer recommended.



The viral enzyme neuraminidase is a glycoside hydrolase that cleaves terminal sialic acid (also called *N*-acetylneuraminic acid) residues from glycoproteins (**Figure 11.25**). The mechanism of the cleavage process involves an intermediate oxonium ion where the pyran oxygen holds a positive charge; hydrolysis generates the free sialic acid and the new truncated glycoprotein. Sialic acid–containing glycoproteins are widely distributed on cell surfaces and are recognized by the protein hemagglutinin for attachment of influenza virions. When the new virions are ready to bud from the cell, cleavage of sialic acid residues from cellular glycoproteins facilitates their release. When neuraminidase is inhibited, influenza virions do not bud from cells, and their further spread to new cells is prevented.

Figure 11.25 Neuraminidase-catalyzed cleavage of sialic acid from glycoprotein. Glycoproteins containing terminal sialic acid (also known as N-Ac-neuraminic acid) are substrates for neuraminidase, which cleaves off the carbohydrate. This abbreviated mechanism shows the intermediate where the pyranose ring develops a positive charge. Hydrolysis of the oxonium ion generates free sialic acid and the processed glycoprotein. Viral neuraminidase activity is required for budding by influenza virions.

The first X-ray crystal structure of a neuraminidase was solved in the early 1980s and led to structure-based drug discovery programs. On the basis of the structure of the active site, it was proposed that an amine or equivalent replacing the C-4 hydroxyl would afford important ionic binding interactions with a glutamate at the active site. Another important interaction is the ionic bond between the carboxylate of the substrate and two arginines of the enzyme. One example of a drug developed on the basis of the enzyme's X-ray structure is oseltamivir (**Figure 11.26**). An early neuraminidase inhibitor, Neu5Ac2en, which is identical to *N*-acetylneuraminic acid except for unsaturation at C2, mimics the intermediate oxonium ion. This compound had some activity in a mouse model of influenza and was a basis for later drugs. Oseltamivir replaces the unsaturated pyran ring with a cyclohexene, the C4 hydroxyl group with an amine, and the glycerol side chain with a 3-pentyl ether. While incorporation of the ether eliminated the possibility of hydrogen bonding to the glutamate residue, reorientation of it exposes a hydrophobic surface that interacts with the alkyl chain of the ether. These interactions can be seen in **Figure 11.26**.

The drugs oseltamivir and zanamivir are transition state mimetics (see Section 4.5.4) of the pyranosyl cation (oxonium ion) formed during the neuraminidase-catalyzed hydrolysis reaction. Both contain an sp^2 carbon that resembles both

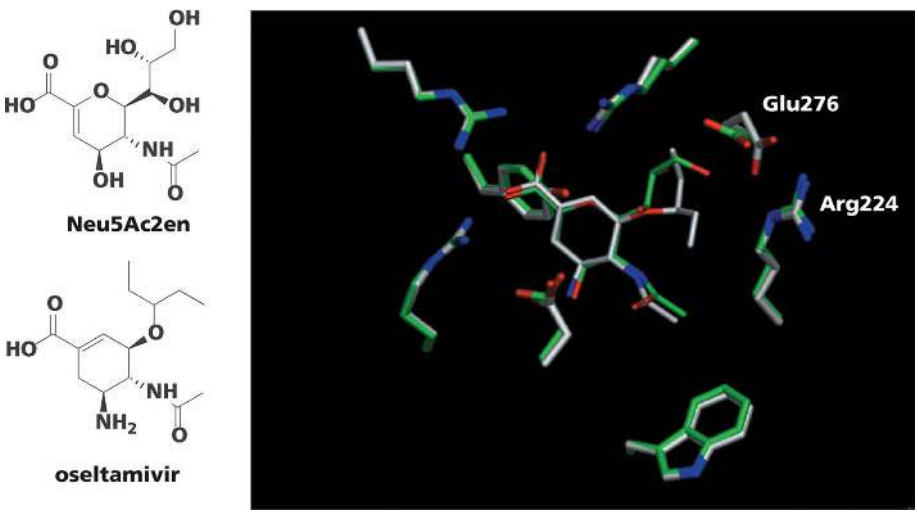


Figure 11.26 Overlay of X-ray structures of oseltamivir and Neu5Ac2en bound to neuraminidase. Oseltamivir (white) and Neu5Ac2en (green) occupy the same binding site. While Neu5Ac2en forms hydrogen bonds to Glu276, oseltamivir does not. In the oseltamivir co-structure, Glu276 reorients to form an ionic interaction with Arg224, exposing a hydrophobic surface that interacts with alkyl ether side chains.

Source: From von Itzstein M. (2007). *Nat Rev Drug Discov* 6:967–974. With permission from Macmillan Publishers Ltd.

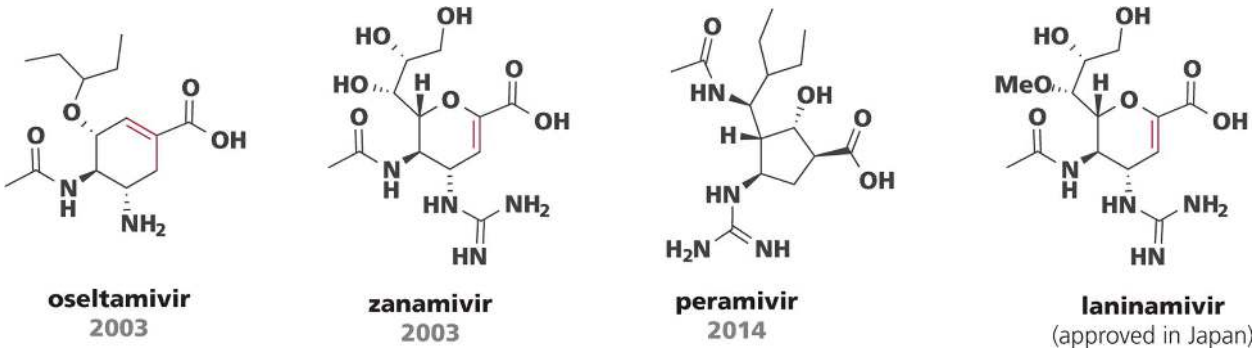


Figure 11.27 Structures of the neuraminidase inhibitors used to treat influenza. Approved neuraminidase inhibitors, which are unsaturated at C2 (unsaturation is shown in red) and mimic the oxonium ion (Figure 11.25), except for peramivir, which has a contracted cyclopentane ring. Oseltamivir and zanamivir were approved in the early 2000s, peramavir in 2014, and laninamivir is approved in Japan.

the oxonium ion and the twist boat conformation generated in the transition state and seen in X-ray co-structures of sialic acid bound to neuraminidase. Additional neuraminidase inhibitors include the cyclopentane peramivir and laninamivir, which is a methoxy analog of zanamivir and is approved in Japan (Figure 11.27). Each structure contains a mimic of the flattened pyranose ring that forms in the transition state; in some cases, it is achieved through the incorporation of unsaturation, and in others through the contraction of the ring.

With resistance to neuraminidase inhibitors increasing, a new approach to influenza is targeting of a viral-specific endonuclease that is required for the transcription of viral RNA. Baloxavir marboxil is a prodrug, converting *in vivo* to baloxavir acid, developed by scientists at Shionogi Pharmaceuticals. It targets the cap-dependent endonuclease of the viral polymerase complex. This enzyme is part of a viral replication complex that cleaves the bound mRNA, generating primers needed for viral transcription in the "cap-snatching" process. Reminiscent of HIV integrase, two metal ions are found in the active site, and baloxavir (approved in 2018) was designed to bind those ions (Figure 11.28) through a diketo functional group. An X-ray model, seen in Figure 11.28, shows the binding of the metal ions to the diketo group.

As seen in preceding sections, the life cycle of a virus contains a number of points of intervention for therapeutic agents. While some antiviral agents were developed on the basis of phenotypic screening, a majority were based on rational design. In contrast to anti-cancer drug discovery, the history of small-molecule antiviral drug discovery and medicinal chemistry is much more recent. In part, this is due to the nature of viruses, which can evolve and cause new diseases that spread rapidly. In fact, HIV/AIDS and HCV were virtually unknown prior to the mid-1980s, and SARS CoV-2 was first detected in late 2019. The ability of viruses to mutate has encouraged the development of agents that are effective against multiple variants. More commonly, however, multidrug regimens are used that rely on different drugs affecting different points in the viral life cycle in an effort to minimize the development of resistance. Current practice for treating HIV/AIDS, called highly active anti-retroviral therapy (HAART), and HCV uses multiple drugs targeting multiple steps in the viral life cycle. In fact, many of the recently approved therapies are combinations of drugs formulated in one pill in an effort to ease the burden on patients and improve compliance. Examples for AIDS include Stribild, which combines an integrase inhibitor, two RT inhibitors, and a pharmacokinetic enhancer that improves the bioavailability of the active agents, and Complera, which includes two NRTIs and one NNRTI. Harvoni, a combination of sofosbuvir and ledipasvir, has a cure rate of >90% in certain HCV patient populations. Approvals in 2017 for HCV include Vosevi, a combination of sofosbuvir, velpatasvir, and the new drug voxilaprevir, and Mavyret, a combination of glecaprevir and pibrentasvir that requires only 8 weeks of treatment. Recent successes in antiviral drug discovery, for example, the rapid development polymerase/RT and protease inhibitors soon after the emergence of HIV/AIDS and SARS-CoV-2, highlight how strategies that were developed for specific target classes (e.g., polymerases, proteases) for older diseases can be applied to new and emerging pathogens

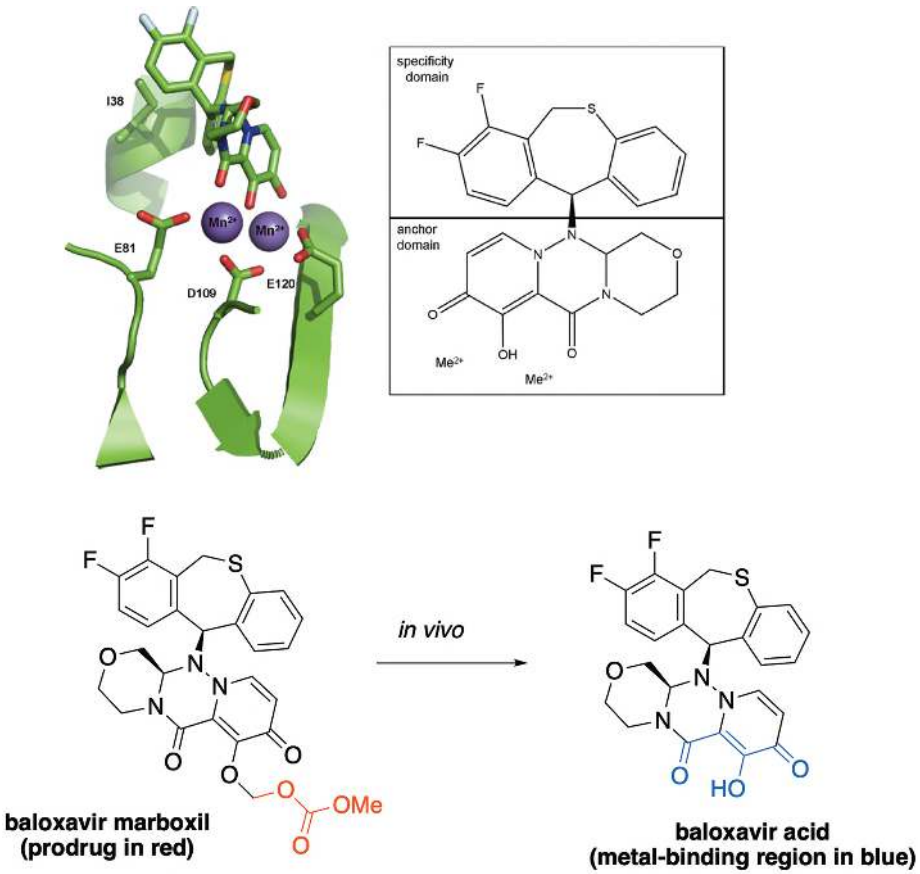


Figure 11.28 Structures of baloxavir marboxil and acid (active form), with model of drug bound in active site with two metal ions as blue spheres. Prodrug moiety in red, metal binding moiety in blue.

Source: Todd, B.; Tchesnokov, E. P.; and Gotte, M. (2021). *J Biol Chem* 296:10046. With permission from Elsevier.

11.5 ANTIFUNGAL AGENTS: COMMON FUNGAL STRUCTURES, REPLICATION PROCESS, AND IMPACT

Unlike viruses, fungi are eukaryotic organisms; they include molds and yeasts as well as mushrooms. Some of the more common pathogenic fungi include *Aspergillus fumigatus*, *Candida albicans*, *Cryptococcus neoformans*, and *Pneumocystis* and *Histoplasma* species. Local or systemic infections can result from exposure to fungi in the environment (air, surfaces, soil) or from overgrowth of organisms that are present in normal human flora. Exposure alone does not necessarily lead to disease; however, symptomatic infection is usually the result of fungal organisms evading host defenses in their target organs, which may be the skin, lung, or other deep tissue sites. Pathology is a result of numerous factors, such as release of toxic or destructive proteins and metabolites by the fungi, competition between fungi and host for nutrients, and host immune response. The specific symptoms associated with a particular fungal infection depend on the location of the infection and the immune status of the patient.

Over the past 50 years, there has been an increase in the number of fungal infections, both topical and systemic. While topical infections, such as athlete's foot and thrush, are somewhat minor, systemic infections may affect internal organs and are deadly. For example, mortality rates for patients infected with the three most common invasive species are 20–40% for *C. albicans*, 50–90% for *A. fumigatus*, and 20–70% for *C. neoformans*. The increase in invasive fungal infections is due, in part, to an increase in the number of patients who are immunocompromised from chemotherapy, organ transplantation, and certain disorders, such as AIDS. In addition, the increased use of indwelling catheters and medical

devices provides surfaces on which fungi can grow. Fungal biofilms on hospital surfaces are also a growing problem, mainly due to *C. albicans* and the emerging threat *C. auris*.

The recorded history of fungal infections precedes that of bacterial and viral infection. In 1665, thrush (caused by *C. albicans*) was recorded as a fatal disease. In 1835, silkworm disease was attributed to the fungus *Beauveria bassiana*; this report is considered the first time an organism was identified as the cause of a disease. Until the 1940s, fungal infections were treated with nonselective toxic agents, such as phenols. The first selective drugs were discovered in the 1940s, with the introduction of the natural product griseofulvin.

Unlike prokaryotic bacteria or simpler viral particles, fungi are eukaryotes. They range from single-celled organisms, like yeasts, to multicellular organisms, like mushrooms. Their structure is characterized by an outer cell wall composed of chitin, glucose polymers (primarily β-(1,3) glucan polymers, but other linkages as well), and mannoproteins (proteins linked to mannose residues). The inner cell membrane is composed of phospholipids containing the steroid ergosterol, rather than cholesterol, the constituent of animal cell membranes. The internal features of part of a fungal cell are very similar to mammalian cells, with a nucleus, mitochondria, and ribosomes and other cellular structures (**Figure 11.29**).

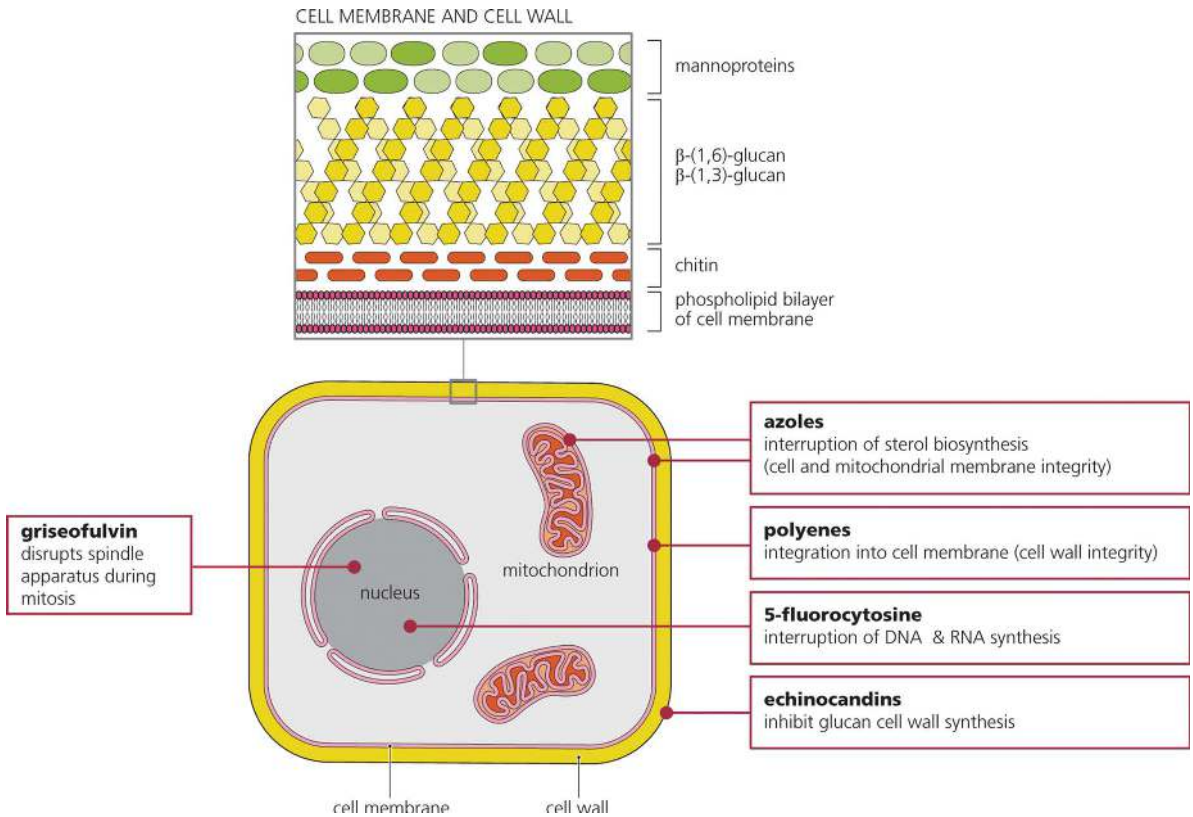


Figure 11.29 Model of fungal cell structure, with points of drug interaction. The fungal cell wall is composed of an outer layer of mannoproteins covering β-(1,3) and (1,6)-glucan polymers and a layer of chitin over the phospholipid bilayer of the cell membrane, which contains ergosterol. Inside the cell, there are mitochondria, ribosomes, and a nucleus. Points of action of the various drug classes are highlighted.

Table 11.2 Examples of Small-Molecule Antifungal Drugs and Their Mechanisms of Action

Antifungal Drug	Mechanism	Representative Indication
Amphotericin B, nystatin	Binds ergosterol, disrupts membrane integrity	Amphotericin: serious systemic fungal infection; nystatin: (topical) skin and mucosal infections
Ketoconazole, terbinafine, itraconazole	Inhibits ergosterol biosynthesis, disrupts membrane integrity	Ketoconazole and terbinafine (topical): skin infections (athlete's foot; dandruff); itraconazole: blastomycosis, histoplasmosis, aspergillosis (in immunocompromised patients), fungal nail infections
Griseofulvin, flucytosine	Inhibits cell replication	Flucytosine: serious fungal infection caused by *candida* or *cryptococcus*; griseofulvin: skin, hair, nail infections
Echinocandins	Inhibits cell wall synthesis	*Aspergillus* and *candida* systemic infections

11.6 ANTIFUNGAL DRUG DISCOVERY

While fungal cell structure is similar to that of mammalian cells, there are some enzymes and structures, unique to fungi, that are targeted by antifungal drugs. Most drugs to treat fungal infections have, for the most part, been discovered by phenotypic assays for overall antifungal activity, and elucidation of their mechanisms has identified why they are selective. There are only a few classes of antifungal drugs: polyenes and azoles affect fungal cell membrane integrity, 5-fluorocytosine affects DNA and RNA synthesis, echinocandins affect cell wall synthesis, and griseofulvin disrupts the spindle apparatus assembled during mitosis (**Figure 11.26**). **Table 11.2** highlights a few examples of antifungal agents that work via these different mechanisms.

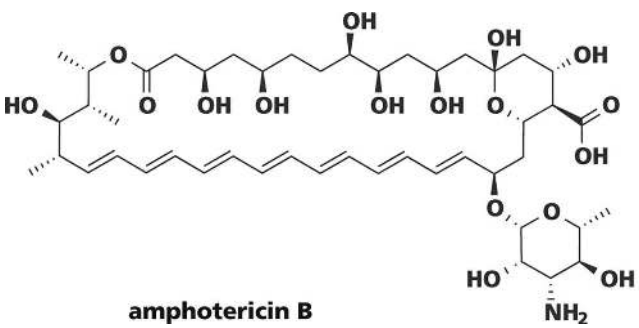
11.7 DRUGS TARGETING FUNGI

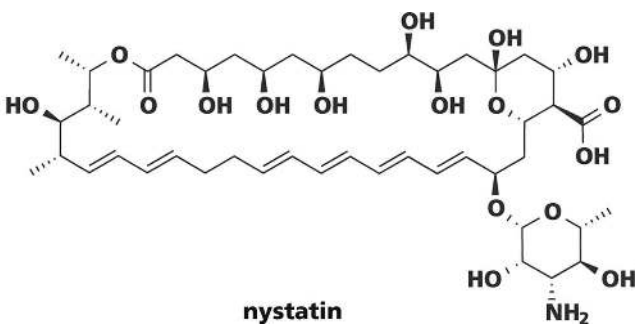
The discovery of antifungal agents has been based on screening of compounds, often natural products, for their ability to inhibit the growth of a fungus. Many drugs target the fungal cell wall (its integrity and biosynthesis) since this structure is distinct from mammalian cell membranes, and this difference accounts for their selectivity. While the specific targets of many older drugs were identified after their antifungal effects were realized, the current focus is on identifying small molecules that specifically exploit these targets and others that are unique to fungi. Some drugs are more active against certain fungi, but recent efforts are concentrated on identifying targets and associated pharmacophores that lead to broad-spectrum activity. In addition, phenotypic assays that are biased toward specific mechanisms (such as cell wall biosynthesis, efflux, or resistance) are also being adopted.

11.7.1 Polyene Drugs Target the Fungal Cell Membrane

The polyene antifungal drugs include the natural products amphotericin B and nystatin (**Figure 11.30**). Their mechanism involves a lipid–lipid interaction with ergosterol in the cell membrane, leading to sequestration of ergosterol (see Section 8.4). Nystatin, used as an oral preparation to treat an oral *Candida* infection

Figure 11.30 Structures of polyene antifungal drugs. Amphotericin B and nystatin, which work by interaction with ergosterol, differ only in their hydroxylation patterns and the saturation of one alkene in nystatin.





(thrush), is well-tolerated. However, amphotericin B, used intravenously to treat systemic *Candida* and other fungal infections, often has serious acute and chronic side effects.

Amphotericin B was discovered at Squibb as the result of a natural product screening effort that examined fungal *Streptomyces* strains with the goal of identifying novel compounds that could be used to treat systemic infections. In 1953, they identified an isolate from a Venezuelan soil culture that provided amphotericin B, which was potent and could be prepared in a formulation that could be administered intravenously. Nystatin was also discovered in the early 1950s, by Elizabeth Hazen and Rachel Brown, at the New York State Department of Health (hence the name nystatin). The compound was isolated from *Streptomyces noursei* in a soil sample and was patented and licensed to Squibb for development.

11.7.2 Azoles and Allylamines Inhibit Ergosterol Biosynthesis

Like the polyene antifungal compounds, the azole and allylamine class also disrupt the fungal cell membrane, but their mechanism involves inhibiting the synthesis of ergosterol. The biosynthesis of all steroids (human and fungal) begins with squalene, which is oxidized to 2,3-oxidosqualene, in a reaction catalyzed by squalene-2,3-epoxidase. Cyclization to lanosterol is catalyzed by lanosterol synthase. In humans, lanosterol is then converted to steroids, such as cholesterol; in fungi, it is converted to ergosterol. One step in the conversion of lanosterol to ergosterol (a steroid unique to fungi) is catalyzed by the enzyme sterol 14α-demethylase. The key steps in the transformation from squalene to ergosterol are shown in **Figure 11.31**.

The discovery of azoles in the 1970s was spurred by both an increased incidence of fungal infections and a need for drugs that were less toxic than the available treatments. The earliest imidazole approved, clotrimazole, was identified as having antifungal activity against *Candida* and *Aspergillus* in the late 1960s. It was synthesized at Bayer in Germany by Karl Buchel and marketed in the early 1970s. The drug was much less toxic than amphotericin B and nystatin and could be administered either orally or intravenously. Variations in the aryl substituents led to a number of analogs, including miconazole, ketoconazole, fluconazole, and itraconazole (**Figure 11.32**). These azoles are used to treat *Candida* infections. While ketoconazole is no longer widely used to treat systemic infections, it is present in over-the-counter drugs used to treat mild infections, such as athlete's foot, dandruff, and jock itch.

More recent antifungal drugs include posaconazole and voriconazole. Posaconazole, approved in 2006, was designed on the basis of tetrahydrofuran-based azoles, which were identified as showing improved activity over the dioxolane itraconazole. Optimization of initial compounds focused on the introduction of hydroxyl groups

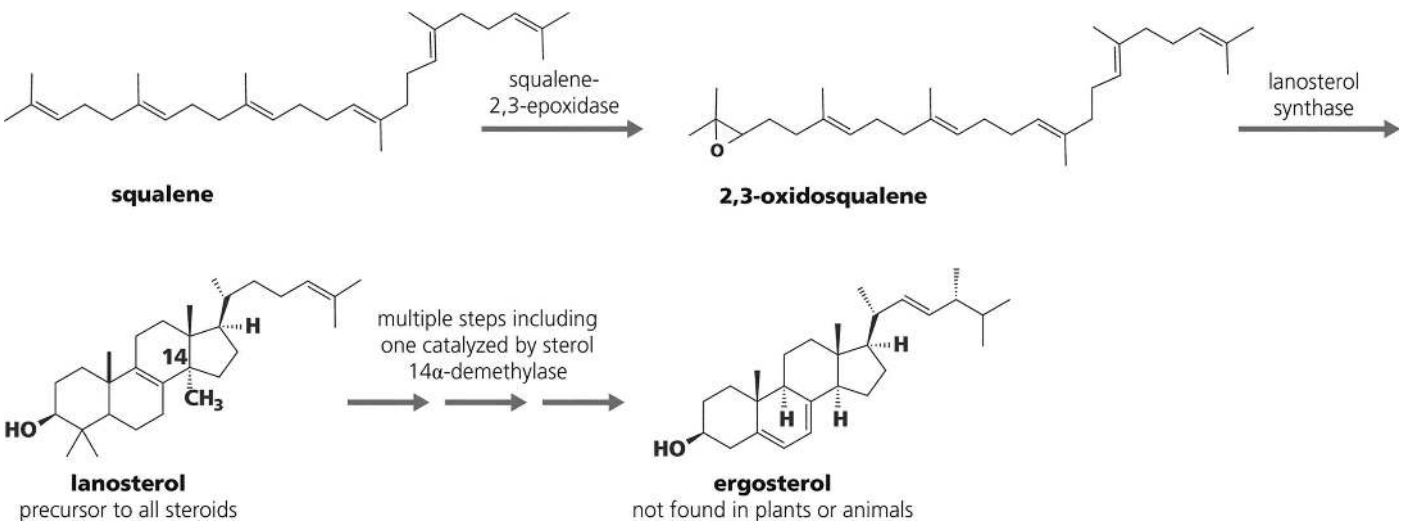


Figure 11.31 Key steps in the biosynthesis of ergosterol. The precursor squalene is epoxidized to 2,3-oxido squalene, in a reaction catalyzed by squalene-2,3-epoxidase. Lanosterol synthase catalyzes the cyclization of the epoxide to lanosterol. Multiple steps, one of which is catalyzed by sterol 14α-demethylase, convert lanosterol to ergosterol.

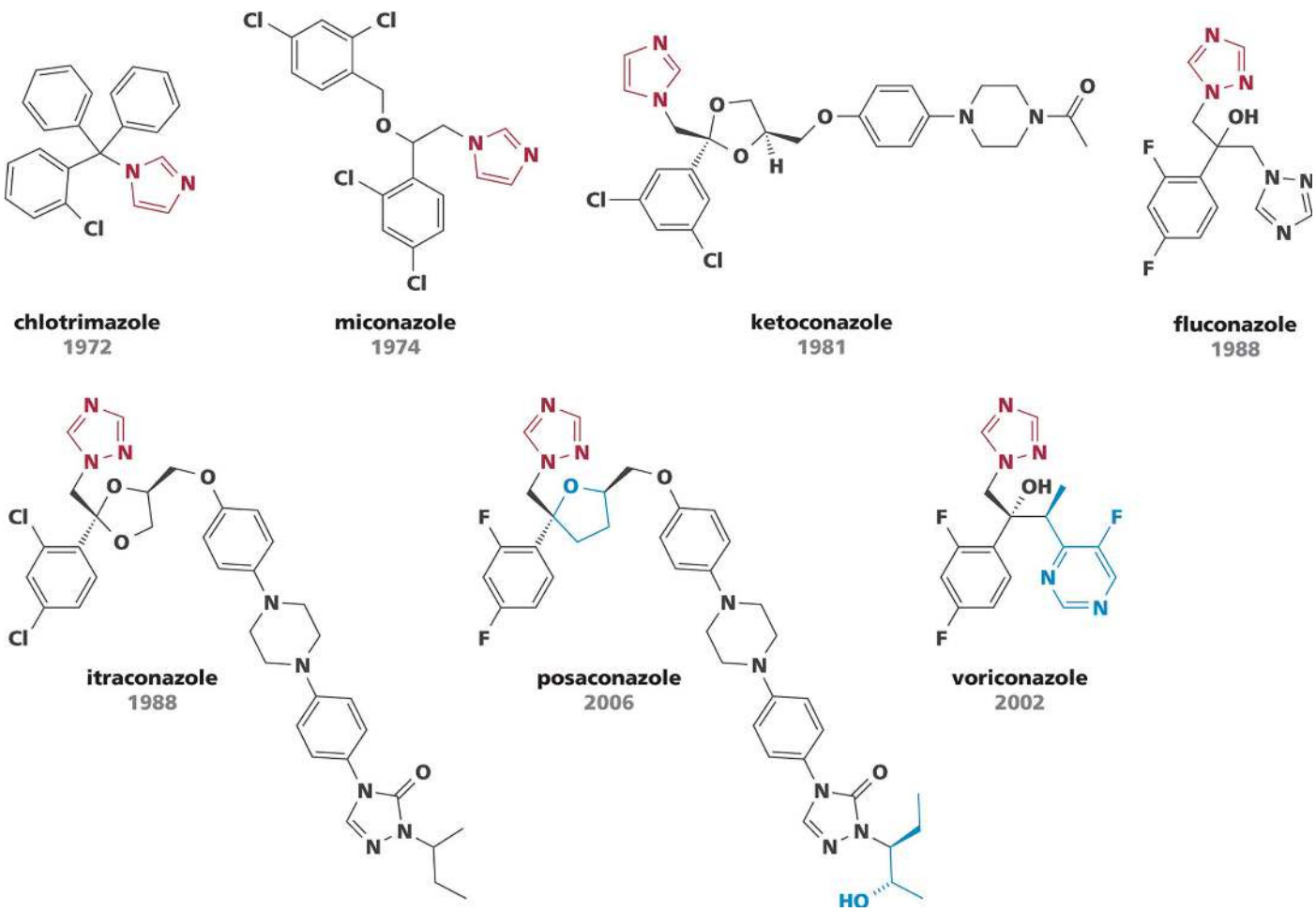


Figure 11.32 Structures of azole antifungal drugs and year of approval. The defining common structural feature, imidazole or triazole, is shown in red. Recent antifungal drugs include posaconazole (modifications compared to itraconazole are shown in blue) and voriconazole (modifications compared to fluconazole are shown in blue).

on the periphery to improve the physical properties of the lipophilic parent compounds. This lipophilicity, characteristic of all azoles, makes formulating them into an intravenous delivery system challenging, so design focused on incorporating polarity. Voriconazole was designed to have a broader spectrum of activity compared to fluconazole, based on the observation that substitution alpha to one of the triazoles resulted in an increase in potency toward *A. fumigatus*. A series of analogs were prepared that contained α-substitution, as well as replacement of one of the triazoles with a six-membered ring heterocycle. Voriconazole, approved in 2002, has a broader antifungal spectrum than fluconazole.

The azoles are all inhibitors of the enzyme sterol 14α-demethylase. This fungal enzyme, also called CYP51, is an oxidoreductase (see Chapter 7) and a member of the class of P450 enzymes. The same class is found in humans and is responsible for metabolism of xenobiotics, among other activities. Some of the side effects of azoles are due to their activity toward these related human homologs. These enzymes contain a heme group at their active site, where the imidazole or triazole nitrogen binds. The relatively lipophilic side chains occupy hydrophobic pockets that are normally occupied by the substrate, as seen in **Figure 11.33**. The availability of X-ray structures of this enzyme has aided drug design.

Another class of drugs, the allylamines, also inhibit the synthesis of ergosterol, but through a different mechanism. These compounds inhibit the enzyme squalene-2,3-epoxidase, which catalyzes the first step in the conversion of squalene to lanosterol (**Figure 11.28**). These are broad-spectrum topical antifungal agents, which are widely used to treat dermatomycoses (skin and toenail infections). The parent compound, naftifine, was discovered at Sandoz through random screening of their compound library for antifungal activity and introduced in 1988. This compound was originally isolated as a by-product in a synthetic program directed at synthesizing

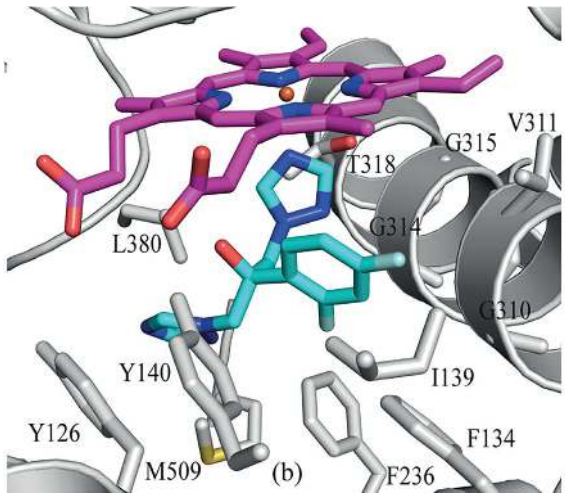
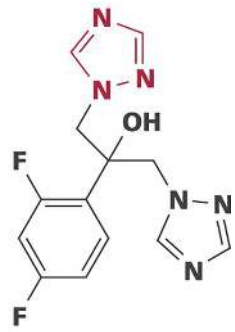
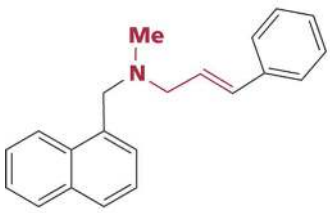


Figure 11.33 X-ray co-structure of fluconazole bound to sterol 14α-demethylase. In this X-ray co-structure, the protein is shown in gray, with specific residues in the active site labeled; the heme group is in magenta, and fluconazole is in cyan. One triazole ring of fluconazole (indicated in red in the structure on the right) interacts with the heme group. The difluorophenyl and second triazole sit in hydrophobic pockets of the enzyme.

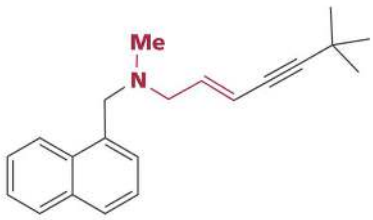
Source: From Sagatova, A. A.; Keniya, M. V.; Wilson, R. K.; et al. (2015). *Antimicrob Agents Chemother* 59:4,982–4,989. With permission from the American Society for Microbiology.



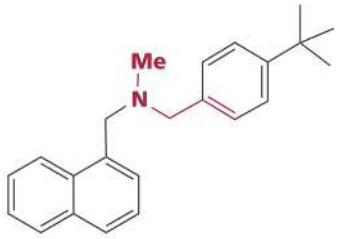
fluconazole



naftifine



terbinafine



butenafine

Figure 11.34 Structures of allylamine antifungal drugs. Naftifine was developed from a screening campaign; the incorporation of a *tert*-butyl group, as seen in terbinafine and butenafine, provided increased potency. The required allylamine common feature is shown in red; these drugs act by the inhibition of squalene-2,3-epoxidase.

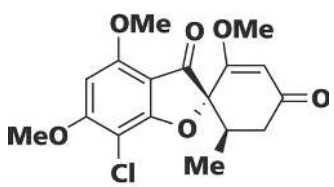
spironaphthalenes (Figure 11.34). The allylamine is a requirement for activity, and the incorporation of *tert*-butyl moieties provided terbinafine and butenafine, which are reversible, noncompetitive inhibitors of the enzyme and show improved potency compared to naftifine.

11.7.3 Some Antifungal Drugs Inhibit Cell Replication

Griseofulvin, a natural product isolated in 1939 from *Penicillium griseofulvum*, was the first antifungal drug marketed (1959) (Figure 11.35). It is used as a treatment for ringworm, and although many analogs have been prepared, it is the only drug in its class.

The mechanism of action of griseofulvin is still not completely understood. It has been shown to bind to tubulin and interfere with mitosis. However, the binding affinity is low, and it is possible that its antifungal effect is due to binding to a microtubule-associated protein instead.

Flucytosine, a derivative of cytosine, was originally synthesized as an anti-tumor agent but was inactive. It is also inactive against fungi but is taken up into fungal cells by the action of cytosine permease. Once in the fungal cells, it is converted to 5-FU by cytosine deaminase, an enzyme not found in mammals. Like other nucleosides, 5-FU is converted to the monophosphate (FdUMP) and then to the corresponding triphosphate, which inhibits fungal DNA replication (Figure 11.36). Flucytosine is generally used in combination with other antifungal drugs and is limited by toxicity.



griseofulvin

Figure 11.35 Structure of griseofulvin. Griseofulvin was the first antifungal agent approved. It inhibits mitosis, but its exact mechanism of action is not well understood.

11.7.4 Echinocandins Inhibit Fungal Cell Wall Synthesis

The echinocandins are a relatively new class of semisynthetic antifungal drugs, introduced in the early 2000s. They are derived from naturally occurring cyclic lipopeptides that inhibit fungal cell wall synthesis. Specifically, they are noncompetitive inhibitors of 1,3-β-glucan synthase, an enzyme involved in the synthesis of

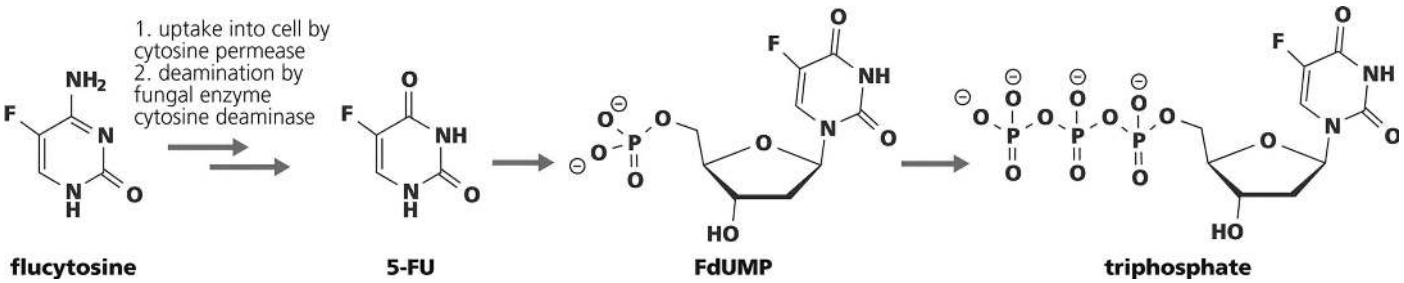


Figure 11.36 Mechanism of action of flucytosine. Flucytosine, a modified nucleobase, is taken up into fungal cells by cytosine permease. It is converted to 5-FU in fungal cells by the fungal-specific enzyme cytosine deaminase, and then eventually to the triphosphate, which inhibits DNA replication.

the glucan polymers that are found in fungal cell walls. Since this enzyme has no mammalian counterpart, targeting it is expected to be relatively safe. Echinocandin B was isolated in 1974 from an *Aspergillus* strain; a related class, the pneumocandins, were isolated in the 1980s (**Figure 11.37**).

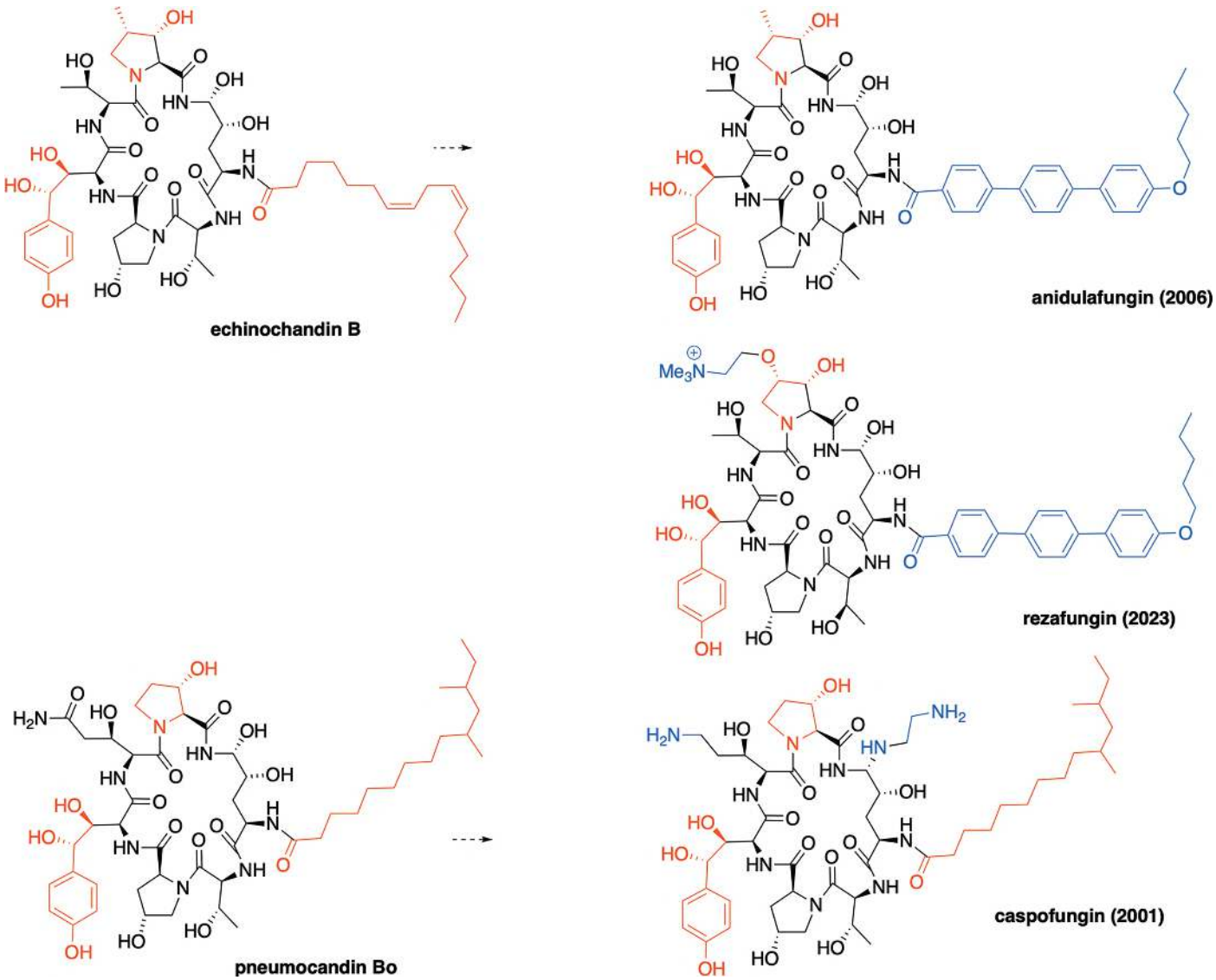


Figure 11.37 Structures of 1,3-β-glucan synthase inhibitors. Natural products echinocandin B and pneumocandin B_0 are on the left, and approved drugs anidulafungin, rezafungin, and caspofungin are on the right. Moieties required for activity are shown in red, and semisynthetic changes shown in blue.

While echinocandin B was found to be too toxic for use, several semisynthetic analogs are now used clinically, including anidulafungin, rezafungin, and caspofungin. SAR studies on echinocandin B analogs found that the lipophilic side chain as well as the hydroxlyated l-homotyrosine residue were required for activity. While not crucial, the 3-hydroxy-4-methylproline found in echinocandin B enhanced antifungal activity and is maintained in anidulafungin. The approved drugs maintained the essential features, incorporating modifications in the lipophilic side chain in anidulafungin and additional of amino groups in rezafungin and caspofungin. The most recent in this family, rezafungin, was approved in 2023 and has a longer half-life, allowing for less frequent dosing.

A novel class of glucan synthase inhibitor, the triterpene ibrexafungerp, was approved in 2021, the first novel structural class of antifungals since the early 2000s (**Figure 11.38**). This is a semisynthetic drug derived from the natural product enfumafungin and is a non-competitive inhibitor of 1,3-β-glucan synthase. Enfumafungin was isolated from *Hormonema carpetanum ATCC 74360*, a black yeast found in decaying leaves in Spain.

Chitin synthase is another target involved in cell wall biosynthesis. Chitin is a polymer of N-Ac-glucosamine and not found in humans. Polymerization occurs by reaction of the 3′-hydroxyl group of the growing polymer chain with uridine diphosphate *N*-acetylglucosamine (UDP-GlcNAc) in a reaction catalyzed by chitin synthase (**Figure 11.39**). The natural product nikkomycin Z is an inhibitor of this enzyme.

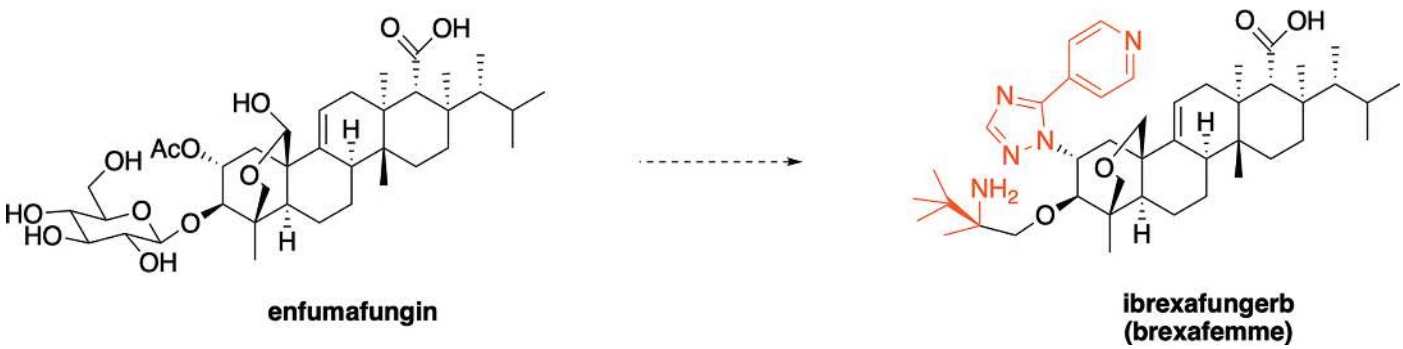


Figure 11.38 Structures of the natural product enfumafungin and the drug ibrexafungerp (changes in red).

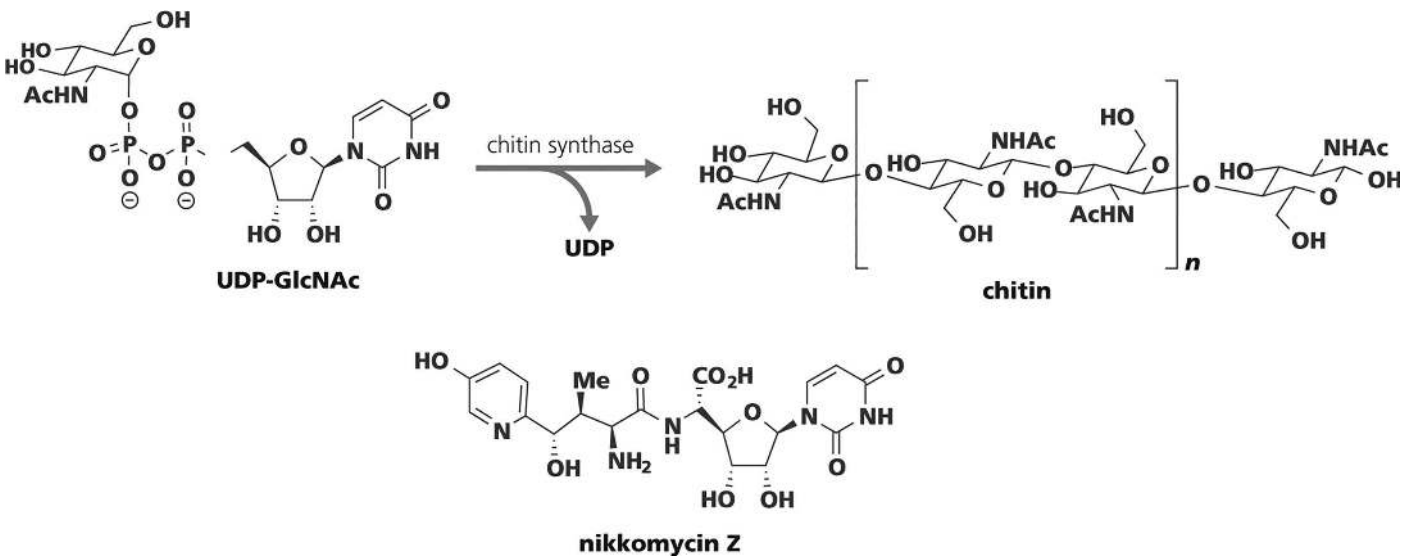


Figure 11.39 Formation of chitin and structure of nikkomycin Z. Chitin synthase is a transferase that catalyzes the addition of an *N*-acetylglucosamine unit (from uridine diphosphate *N*-acetylglucosamine, UDP-GlcNAc) to the chitin polymer. The inhibitor, nikkomycin Z, is structurally similar to the natural substrate.

Nikkomycin Z resembles the enzyme's substrate and progressed to Phase II clinical trials as a treatment for valley fever (coccidioidomycosis). This relatively rare disease is localized in the Southwestern United States and Central and South America and causes respiratory symptoms that can be fatal in immunocompromised patients.

While the number of antifungal drugs lags far behind the number of antiviral drugs, recent efforts in searching for novel antifungal compounds should add to those on the market. Many of these drugs were identified via phenotypic screening. However, in some cases, extensive medicinal chemistry optimization afforded improved analogs. For example, the extensive SAR developed around the azole scaffold, as well as an understanding of its binding mode to its enzymatic target, has allowed a more rational approach to drug design that has culminated in new analogs with broad-spectrum activity and improved physical properties. In addition, new screening strategies that specifically target one pathway are leading to new structural leads. These approaches are starting to address the difficulties in developing drugs selective for fungal targets.

11.8 SUMMARY

In the past 30 years, much progress has been made in the development of both antiviral and antifungal drugs. The significant advancement in the development of antiviral drugs was directly related to the initial success of drugs to treat herpes in the 1970s and then the massive efforts devoted to the development of drugs to address the AIDS pandemic in the late 1980s. Both library screening and structure-based drug design were used to identify RT inhibitors, HIV protease inhibitors, and fusion inhibitors. Structure-based drug design was also applied to the development of neuraminidase inhibitors that are currently approved to treat influenza. The current therapies for hepatitis C, many of which are curative, also rely on similar strategies: inhibition of viral polymerase and protease. These studies led to a deep understanding of how to target specific viral proteins and paved the way for the rapid development of drugs to treat emerging infections, such as SARS CoV–2.

Antifungal compounds represent a much smaller family of drugs, and initial leads have primarily been identified through screening rather than drug design. With an increase in fungal infections, including dangerous systemic infections, and the identification of specific mechanisms, medicinal chemistry has led to the introduction of improved compounds.

CASE STUDY 1 DISCOVERY OF DACLATASVIR

In a search for compounds to inhibit replication of hepatitis C, scientists at Bristol-Myers Squibb identified BMS-858 from a library screening program. Over 1 million compounds were screened in a high-throughput assay that was specifically designed to identify compounds with a mechanism distinct from NS3 protease and NS5B polymerase inhibition. A screening strategy was used that simultaneously evaluated antiviral activity and selectivity vs. human cells. The hit pursued, BMS-858, contains a thiazolidine core and inhibited HCV replication at an EC_{50} of 0.57–1 μM with no cytotoxicity. Exploration of this promising result by synthesis of analogs afforded the lead compound BMS-824, which replaced the carbamate of BMS-858 with an amide and had ~100-fold improvement in potency (Figure 11.40).

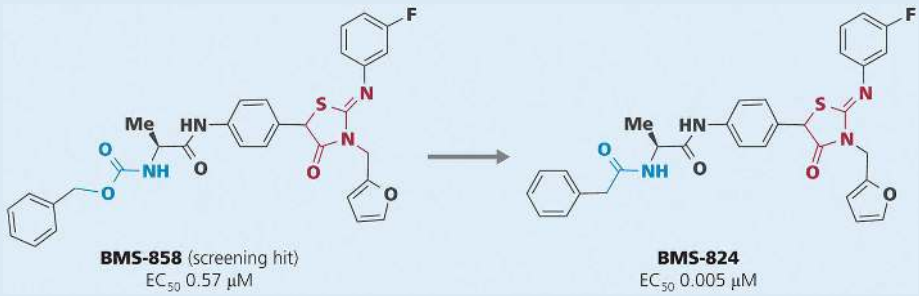


Figure 11.40 Structures of hit and early lead compound in the development of daclatasvir. BMS-858 was identified from a high-throughput screen for inhibitors of HCV replication that was biased away from known targets. Further optimization afforded the more potent BMS-824. The thiazolidine core is shown in red, and the change from carbamate (BMS-858) to amide (BMS-824) is shown in blue.

Investigation of the mechanism of BMS-858 indicated that it was inactive against both the HCV NS3 protease and NS5B polymerase. The target was ultimately identified by generating cells that exhibited BMS-858-resistant phenotypes; all had mutations in the N-terminal region of the HCV NS5A replication protein. The antiviral activity of the compound confirmed that this was a new, valid target for HCV drugs.

Despite its potent activity, BMS-824 was found to be unstable in the assay medium. Furthermore, the scaffold is related to rhodanines, which are characterized as PAINS compounds (see Section 2.7.1). To determine the source of its instability, BMS-824 was incubated in the assay medium. HPLC and bioassay of individual components showed that oxidation/rearrangement product **1** and a thiourea fragment **2** (Figure 11.41) were major decomposition products. Neither was active, but further

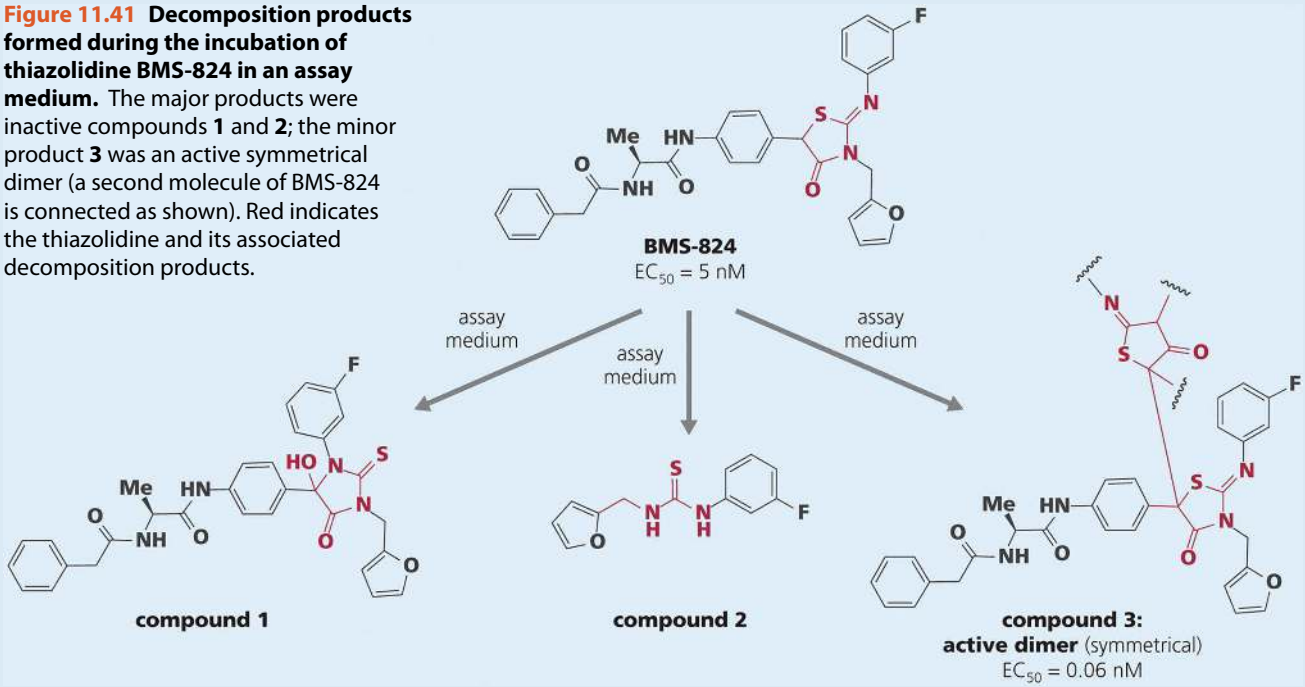


Figure 11.41 Decomposition products formed during the incubation of thiazolidine BMS-824 in an assay medium. The major products were inactive compounds **1** and **2**; the minor product **3** was an active symmetrical dimer (a second molecule of BMS-824 is connected as shown). Red indicates the thiazolidine and its associated decomposition products.

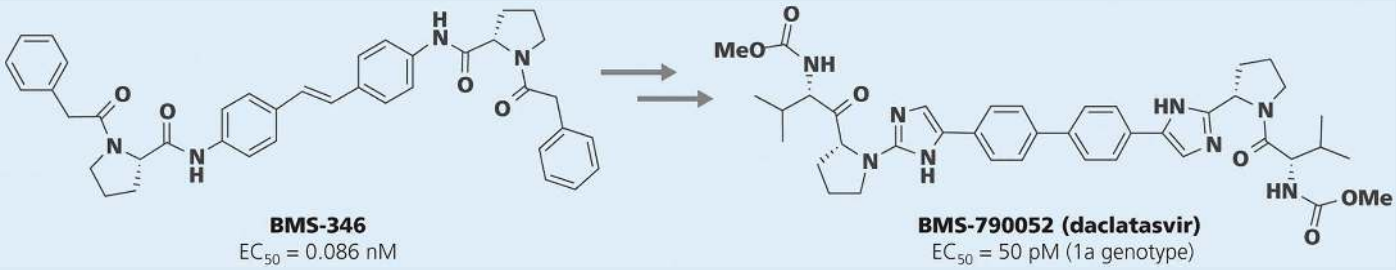


Figure 11.42 Symmetrical analog of active dimer and structure of daclatasvir. Replacement of the unstable thiazolidine with a styrene and further optimization, including the incorporation of L-proline caps, afforded BMS-346. Daclatasvir retains the proline groups but replaces the styrene with a biphenyl group. It is a highly potent inhibitor of viral replication.

isolation of minor products identified the dimeric structure **3**, which was a very potent inhibitor of HCV replication (EC_{50} = 0.06 nM); it was nearly 100-fold more potent than BMS-824.

On the basis of the structure of the active dimeric compound, stable symmetrical analogs were designed, and the unstable thiazolidine was replaced. The stilbene analog BMS-346, with L-proline caps, was the most active, with the corresponding D-proline isomer being completely inactive (**Figure 11.42**). The inhibitor binds in the amino terminus of the HCV NS5A protein, which crystallizes as a dimer. This dimeric structure complements the symmetry of the inhibitors.

The stilbene BMS-346 maintained the potency seen in the thiazolidine dimer, with an EC_{50} of 0.086 nM. Further modifications of the core led to the biphenyl analog BMS-790052 (daclatasvir), with EC_{50} values of 50 pM against the HCV 1a genotype and 9 pM against the 1b genotype. Daclatasvir is orally available and was found to be well-tolerated in phase I clinical trials at doses up to 200 mg. In infected patients, a combination of daclatasvir and sofosbuvir, given as an oral dose once a day, led to a sustained response after 12 weeks of treatment. This combination was approved by the European Union in 2014 and by the US FDA in 2015.

CASE STUDY 2 DISCOVERY OF RILPIVIRINE

In the late 1980s, Janssen Pharmaceuticals initiated a search for compounds to inhibit HIV-1 replication in collaboration with the Rega Institute for Medical Research. A library of 90,000 compounds was screened for their ability to inhibit HIV-1 replication in a cell-based assay, and 2,000 hits were chosen applying the additional criteria of low toxicity and lack of off-target effects.

One of the hits identified from the library was the α-aminophenylacetamide (α-APA) R15345 (**Figure 11.43**), which had an EC_{50} of 610 nM in the HIV-1 replication assay, although it was cytotoxic at higher doses. A series of aryl substitutions were

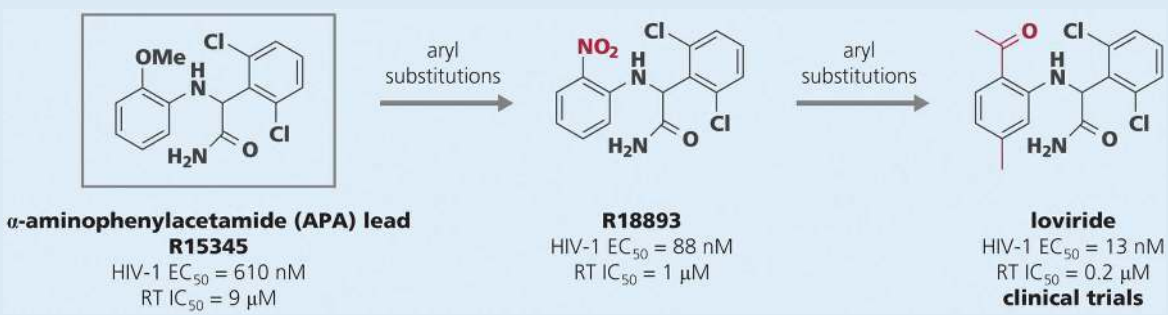


Figure 11.43 Early leads in the development of rilpivirine: from APA lead to loviride. The APA lead was identified by screening for inhibition of HIV-1 replication. Aryl substitutions led to the nitro analog R18893, with improved potency and reduced cytotoxicity. Further substitutions led to the acetyl derivative loviride, which progressed to clinical trials (changes to structure shown in red).

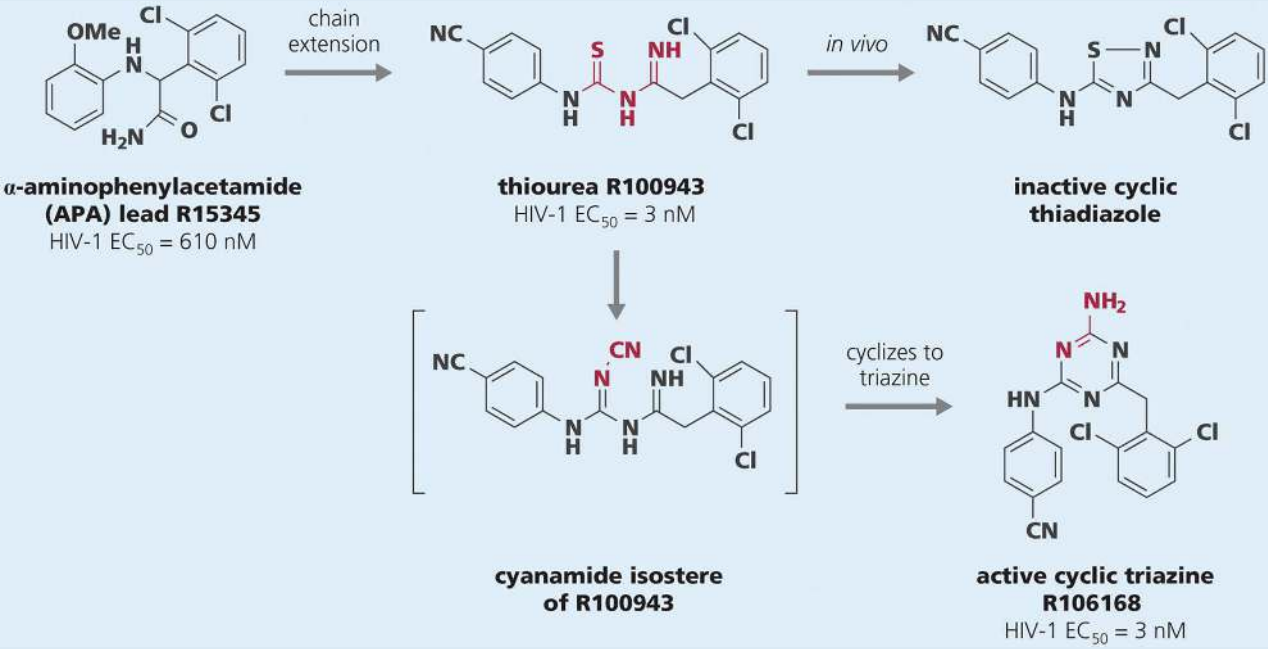


Figure 11.44 Early leads in the development of rilpivirine: from APA lead to cyclic triazine. Chain extension of the APA lead afforded a potent thiourea; however, it cyclized *in vivo* to an inactive thiadiazole metabolite. The incorporation of the *N*-cyano bioisostere gave the cyanamide, which spontaneously cyclized to an active triazine. This compound was highly potent. Changes in structure shown in red.

investigated, and the nitro analog R18893 had improved potency and lower cytotoxicity. The replacement of the nitro group with an acetyl group and the addition of a methyl substituent on the aryl ring afforded loviride, which had further improved potency and low cytotoxicity. Although the initial discovery was based on the inhibition of viral replication, the specific mechanism was investigated and found to be inhibition of RT, with IC_{50} values of 9 μM for the APA lead and 0.2 μM for loviride.

On the basis of potency as well as pharmacokinetic profile, loviride progressed to clinical trials. Although active, loviride had no advantage over other NNRTIs, such as delavirdine, that were already on the market.

In addition to the aryl modifications that afforded loviride, lead modification also included chain extension of the aryl linker of R15345, and although most analogs lost activity, a thiourea analog showed increased potency. This finding led to the synthesis of thiourea R100943, a second clinical candidate (**Figure 11.44**). Unfortunately, its activity was limited by rapid metabolism to an inactive metabolite, a thiadiazole, which resulted from cyclization of the thiourea moiety. In an attempt to avoid the formation of the metabolite, the cyanamide isostere was synthesized. However, this compound was not isolated, as it also cyclized as soon as it was formed to the triazine R106168. Unlike the cyclic thiadiazole, the triazine had excellent potency.

Replacement of one of the triazine ring nitrogens by a carbon to generate a diarylpyrimidine was based on molecular modeling studies and, after further optimization, led to the diarylpyrimidine R147681 (dapivirine) (**Figure 11.45**), as well as the bis(nitrile) R165335 (etravirine).

One of the most promising aspects of dapivirine was its activity against strains of the virus that were resistant to other NNRTIs. In a further medicinal chemistry optimization effort, extensive SAR studies were carried out, based on further improving potency against resistant strains. Modeling into the binding pocket of RT shows adoption of a butterfly-like conformation where one wing interacts with tyrosines Y181 and Y188, as well as tryptophan W229, and one of the NH bonds to the backbone amide of lysine K101, as seen in NNRTIs, such as efavirenz. The allosteric NNRTI binding pocket is flexible and lipophilic. The pocket is small in the absence of a ligand but opens upon binding of an inhibitor. This conformational change results in the inhibition of

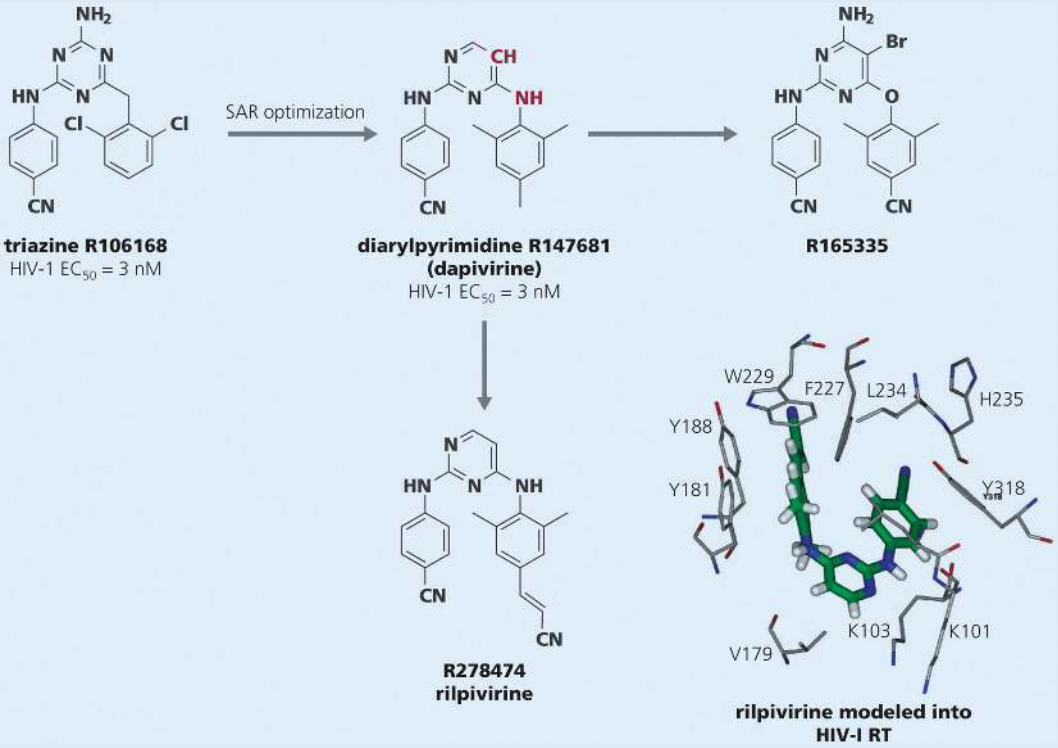


Figure 11.45 Development of rilpivirine: from triazines to diaminopyrazines. SAR optimization of the triazine R106168 led to the diarylpyrimidines. Both R147681 (dapivirine) and R165335 (etravirine) advanced to clinical trials. Molecular modeling based on R165335 indicated that inserting a spacer between one aryl ring and the nitrile would improve binding, leading to rilpivirine. Rilpivirine is seen modeled into HIV-1 RT, showing binding between the nitrile and W229, as well as hydrophobic interactions with Y181 and Y188.

Source: From Janssen, P. A. J.; Lewi, P. J.; Arnold, E; et al. (2005). *J Med Chem* 48:1,901–1,909. With permission from the American Chemical Society.

polymerase activity. Molecular modeling based on etravirine indicated that inserting a spacer between one aryl ring and the *p*-nitrile group would improve binding, leading to rilpivirine.

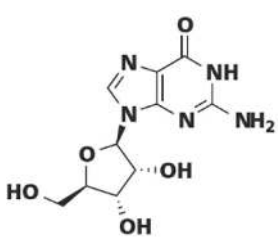
Goals in the development of these novel NNRTIs included high potency against resistant strains. Both the diarylpyrimidine analogs and the binding pocket are flexible, which may contribute to activity against resistant strains. Other goals included a long half-life and high oral bioavailability to allow once-daily dosing, minimal side effects, and ease of synthesis and formulation. Rilpivirine fulfilled all these criteria, with high potency, a half-life of up to 39 hours (in dogs), and oral availability of approximately 35%. Rilpivirine also exhibited no side effects in multiple animal models and is synthetically accessible in a six-step sequence.

11.9 REVIEW QUESTIONS

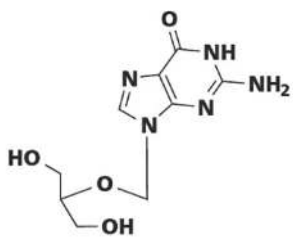
1. Identify at least two antiviral drugs described in this chapter that are prodrugs. Explain why the prodrug strategy was applied to these compounds and what problem it solved. Draw each of their structures and a reaction scheme that shows how each is converted to the active moiety.
2. Explain the difference between NRTIs and NNRTIs.
3. Identify three drugs that act at different points in the viral replication cycle. What are the drugs, what are their structures, and at what point in the viral life cycle do they interfere?
4. How do polyene and azole antifungal drugs differ in their mechanisms of action?
5. The HCV NS5A inhibitors (see **Figure 11.22**) are all either symmetrical or nearly symmetrical. Explain why this might be (refer to Case Study 1 in this chapter).
6. Explain why there are fewer small-molecule drugs to treat viral and fungal infections than bacterial infections.
7. Why are nucleoside drugs often toxic? Acyclovir is surprisingly nontoxic, despite being a nucleoside. Explain its lack of toxicity.
8. Identify one drug described in this chapter that works by inhibiting a channel. What is the drug, what channel does it inhibit, and what does it treat?
9. Identify one drug described in this chapter that works by inhibiting a protein–protein interaction. What is the drug, what are its targets, and what does it treat?
10. Identify one drug described in this chapter that works by interfering with lipids. What is the drug, what are its targets, and what does it treat?
11. Identify three different compounds that exhibit antiviral or antifungal activity that are natural products.

11.10 APPLICATION QUESTIONS

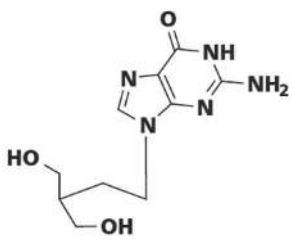
1. HCV NS3 is a serine protease described in Section 11.6, and its inhibitors boceprevir and telaprevir (see **Figure 11.18**) contain a warhead like those described in Chapter 7.
   a. Draw the general structure of the tetrahedral intermediate formed by reaction of one of these inhibitors with HCV NS3.
   b. Design two new inhibitors based on the structure of boceprevir or telaprevir by incorporating different warheads into the inhibitor scaffold.
2. Using the structures of the neuraminidase inhibitors in **Figure 11.27**, apply Lipinski's rules to each of them. What are the parameters for each compound? Do these drugs comply with Lipinski rules?
3. Viruses are examples of a type of agent causing a disease that can be eradicated or cured. Name one example of a virus that has been eradicated or nearly eradicated. Research the history of its eradication.
4. Compare the structures of penciclovir and ganciclovir to the natural substrate they mimic. How does each differ from guanosine? What is the mechanism of action of ganciclovir? Draw the structures of the intermediate products.



guanosine
(natural ligand)



ganciclovir



penciclovir

5. For three different enzymes, identify one representative antiviral or antifungal drug that works through its inhibition. Applying the concepts in Chapter 7, describe the enzyme, what reaction the enzyme catalyzes, and how the inhibitor works.
6. From the examples in this chapter, find five different examples of molecules containing a privileged structure. Draw the structures and highlight the privileged structures.
7. Many of the HIV protease inhibitors in **Figure 11.16** contain a hydroxyethylene transition state mimetic. Based on information from Chapter 4, design three other possible inhibitors based on the structure of saquinavir by incorporating alternative transition state mimetics.
8. AZT was the first compound approved to treat HIV. Research its history. Who synthesized it, when, and why? Draw the initial synthesis. Research a more recent synthesis and compare the two.
9. A number of individuals were awarded Nobel prizes for their work on HIV over several years. Name four of them, with the years of their awards, and describe their contributions.
10. Ebola is a life-threatening virus. Research the small-molecule approaches that are being taken to identify drugs to fight it. Draw the structures of two, and describe their mechanisms of action.

FURTHER READING

HIV/HCV

De Clercq E (2007) The design of drugs for HIV and HCV. *Nat Rev Drug Discov* 6:1001–1018.

De Corte BL (2005) From 4,5,6,7-Tetrahydro-5-methylimidazo[4,5,1-*jk*] (1,4)benzodiazepin-2(1*H*)-one (TIBO) to etravirine (TMC125): fifteen years of research on non-nucleoside inhibitors of HIV-1 reverse transcriptase. *J Med Chem* 48:1689–1696.

Ghosh AK, Chapsal BD, Weber IT & Mitsuya H (2008) Design of HIV protease inhibitors targeting protein backbone: an effective strategy for combating drug resistance. *Acc Chem Res* 41:78–86.

Gordon CP & Keller PA (2005) Control of hepatitis C: a medicinal chemistry perspective. *J Med Chem* 48:1–20.

Huryn DM & Okabe M (1992) AIDS-driven nucleoside chemistry. *Chem Rev* 92:1745–1768.

Pommier Y, Johnson AA & Marchand C (2005) Integrase inhibitors to treat HIV/AIDS. *Nat Rev Drug Discov* 4:236–248.

Thematic issue on HCV therapy. (2014) *J Med Chem* 57:1625–2166.

Acyclic Nucleosides and Phosphonates

De Clerq E & Holy A (2005) Acyclic nucleoside phosphonates: a key class of antiviral drugs. *Nat Rev Drug Discov* 4:928–940.

Freeman S & Gardiner JM (1996) Acyclic nucleosides as antiviral compounds. *Mol Biotechnol* 5:125–137.

Fyfe JA, Keller PM, Furman PA et al. (1978) Thymidine kinase from herpes simplex virus phosphorylates the new antiviral compound, 9–(2-hydroxyethoxymethyl)guanine. *J Biol Chem* 253:8721–8727.

Neuraminidase Inhibitors

Varghese JN, McKimm-Breschkin JL, Caldwell JB et al. (1992) The structure of the complex between influenza virus neuraminidase and sialic acid, the viral receptor. *Proteins: Struct Funct Genet* 14:327–332.

von Itzstein M (2007) The war against influenza: discovery and development of sialidase inhibitors. *Nat Rev Drug Discov* 6:967–974.

von Itzstein M, Wu W-Y, Kik GB et al. (1993) Rational design of potent sialidase-based inhibitors of influenza virus replication. *Nature* 363:418–423.

Antifungal Agents

Denning DW (1997) Echinocandins and pneumocandins: a new antifungal class with a novel mode of action. *Antimicrob Agents Chemother* 40:611–614.

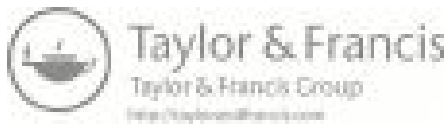
Georgopapadakou N (1998) Antifungals: mechanism of action and resistance, novel and established drugs. *Curr Opin Microbiol* 1:547–557.

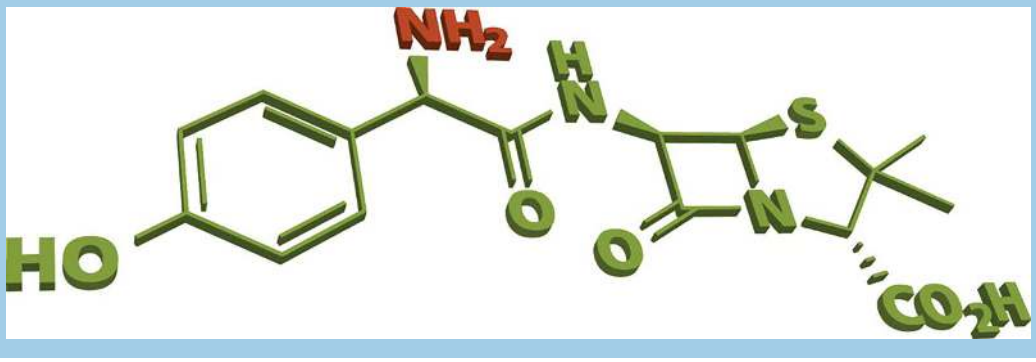
Nussbaumer P, Petranyi G & Stutz A (1991) Synthesis and structure–activity relationships of benzo[*b*]thienylallylamine antimycotics. *J Med Chem* 34:65–73.

Odds FC, Brown AJP & Gow NAR (2003) Antifungal agents: mechanisms of action. *Trends Microbiol* 11:272–279.

Petranyi G, Meingassner JG & Mieth H (1987) Activity of terbinafine in experimental fungal infections of laboratory animals. *Antimicrob Agents Chemother* 31:1558–1561.

Richardson K, Cooper K, Marriott MS et al. (1990) Discovery of fluconazole, a novel antifungal agent. *Clin Infect Dis* 12(Suppl 3): S267–S271.


Taylor & Francis
Taylor & Francis Group



Antibacterial and Antiparasitic Drugs

12

LEARNING OBJECTIVES

- Learn the historical background of the discovery and development of antibacterial and antiparasitic drugs.
- Describe different mechanisms of action of antibacterial agents.
- Identify different classes of antibacterial drugs acting by each mechanism.
- Explain antibacterial drug resistance mechanisms.
- Understand the structure–activity relationships of several classes of antibacterial drugs.
- Become familiar with drugs used to treat parasitic infections and their mechanism of action.

12.1 INTRODUCTION

Infections caused by bacteria and parasites have historically been, and continue to be, a significant worldwide cause of mortality. It has been estimated that almost half of the world's population is at risk of malaria, a disease that has been known for over 4,000 years and is the result of infection by *Plasmodium* parasites. Prior to the early 1900s, bacterial infections, such as tuberculosis, typhoid fever, syphilis, bubonic plague, leprosy, and sepsis from infected wounds, were a leading cause of death. The antibiotic age brought about new treatment for bacterial infections, and the increase in life expectancy in the 20th century has been attributed, in part, to the introduction of antibiotics, such as penicillin. Similar advances have been made in the area of antiparasitic drugs.

Today, these infections continue to be a source of morbidity and mortality, but with less of an impact than they once had. Efforts to purify water sources, minimize exposure to infectious agents, and eliminate vectors of parasites have significantly reduced the incidence of certain bacterial and parasitic infections. Several bacterial infections are also preventable by vaccination, with a few examples being diphtheria, pertussis, meningococcal disease, and typhoid, as well as *Haemophilus influenzae* and *Streptococcus pneumoniae* infections.

Despite the large variety of bacteria and parasites, the same drug is often effective against multiple species, thereby broadening the utility and number of drugs available to treat these infections. In some cases, however, the complex life cycle of certain parasites makes drug discovery particularly challenging, and there are fewer drugs to treat infections caused by parasites than by bacteria. For both bacteria and parasites, resistance to current drugs is a serious problem, and drug discovery efforts are focusing on addressing this issue.

DOI: 10.1201/ 9781003465607-15

12.2 ANTIBACTERIAL AGENTS: BACTERIAL STRUCTURE, POINTS OF DRUG INTERACTION, AND IMPACT

Bacteria are prokaryotic microorganisms, and unlike eukaryotes, they do not have a nucleus and typically do not have organelles. Their DNA and ribosomes are surrounded by a cell wall in addition to a cell membrane. Each of these are drug targets. Bacteria display pili and flagella on their surface that are responsible for adhesion and locomotion, respectively. They are broadly divided into two classes, Gram-positive and Gram-negative, on the basis of their ability to take up the crystal violet of the Gram stain. Gram-positive bacteria have a high percentage (50–70%) of peptidoglycan content in the bacterial cell wall, whereas the cell walls of Gram-negative bacteria are only about 10% peptidoglycan but contain an additional outer membrane that contains pore proteins (porins) and displays lipopolysaccharide (LPS) on the cell surface. LPS is a macromolecule composed of a hydrophobic lipid portion (lipid A), linked to a polysaccharide core, which is linked to O-antigenic oligosaccharides. Gram-positive bacteria contain different molecules on their surface, including lipoteichoic acids and wall teichoic acids. Both contain integral membrane and peripheral membrane proteins, as well as wall-associated proteins, such as those involved in peptidoglycan synthesis. A representation of a bacterial cell is shown in **Figure 12.1a**, with key structural

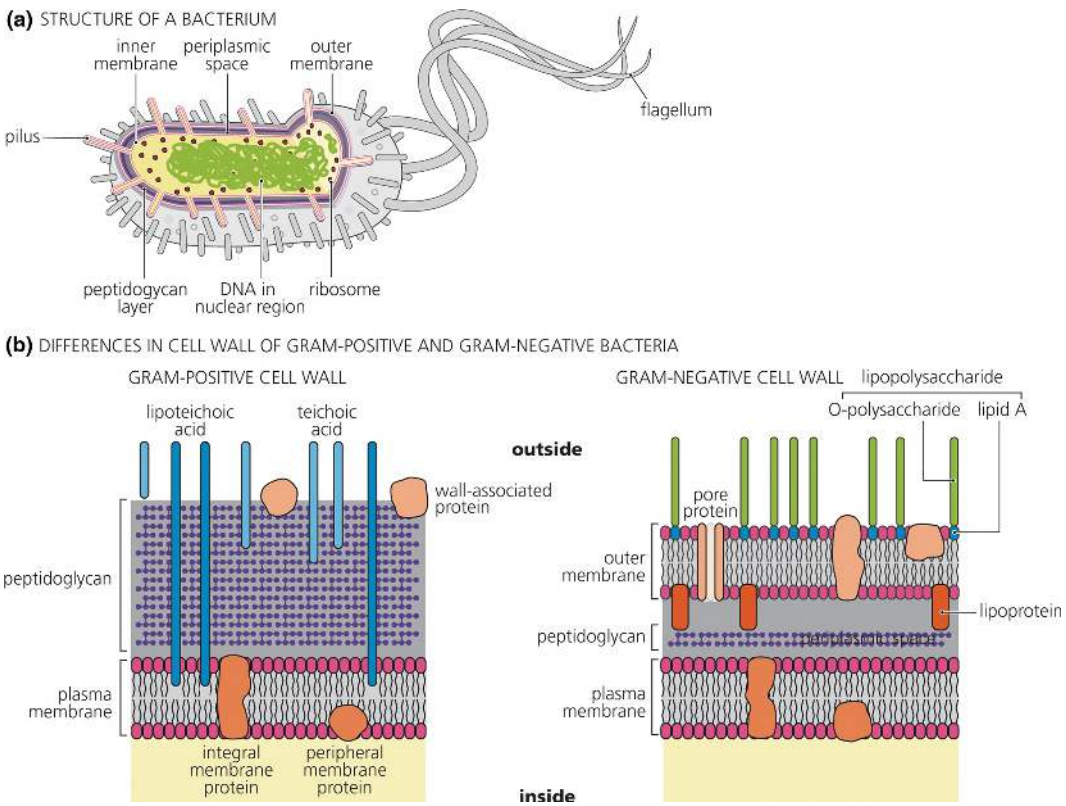


Figure 12.1 Bacterial structure and cell wall details. (a) General structure of a bacterium consisting of pilus (for adhesion) and flagella (for motility), as well as drug targets, such as the cell wall (peptidoglycan), cell membrane, ribosomes, and DNA. (b) Differences in the bacterial cell wall between Gram-positive and Gram-negative bacteria. Gram-positive bacteria have a thick peptidoglycan layer in the cell wall with teichoic and lipoteichoic acid displayed on the surface. Gram-negative bacteria have a thin peptidoglycan layer but have an outer membrane that contains pore proteins (porins) and displays lipopolysaccharides (made up of a lipid A linked to a polysaccharide core). Various proteins (wall-associated, peripheral membrane, and integral membrane) are found in each.

Source: (a) Adapted from Strelkauskas, A.; Edwards, E.; Fahnert, B.; et al. (2015). *Microbiology: A Clinical Approach*, 2nd ed. Garland Science.
(b) Adapted from Alberts, B.; Johnson, A.; Lewis, J.; et al. (2014). *Molecular Biology of the Cell*, 6th ed. Garland Science.

elements highlighted as well as points for drug intervention. Figure 12.1b illustrates the differences in the cell wall between Gram-positive and Gram-negative bacteria.

The process of infection by bacteria involves entering the host, adhering to cells, and then multiplying and spreading. Bacteria tend to colonize tissues that are in contact with the external environment, like wounds or the respiratory tract. Adhesion takes place either by nonspecific interactions between the host cell and the bacterium or by specific interaction between the pili on the bacterial cell surface and host cell-surface receptors.

Unlike viruses, most bacteria multiply outside of cells, forming microcolonies that can spread either to multiple tissues or into the lymph and bloodstream (bacteremia). Bacteria may or may not be pathogenic and can be classified as primary pathogens, which always cause disease; opportunistic pathogens, which usually cause disease only in immunocompromised patients; and nonpathogens, which are part of the normal flora in a human body and do not cause disease.

The pathogenicity of a specific organism, or **virulence**, is the result of numerous factors, including the effectiveness and specific nature of the host's own immune response, the route of entry into the body, and specific virulence factors of the bacterium. Symptoms of a bacterial infection are the result of not only the bacteria themselves and often their toxic secretions but also the host defense mechanisms. In the latter case, immune cells release toxic factors in an effort to eliminate the infection, but at the same time, these factors can destroy the surrounding tissue (sepsis) and may lead to widespread organ failure and death (septic shock).

Since bacteria are rapidly evolving systems, the emergence of resistant organisms is a constant problem, and its impact on health is growing. The inappropriate use of antibiotics contributes to the development of resistant organisms. Bacteria can pass on their resistance genes either in a linear fashion, to their progeny, or horizontally through plasmids to other nonresistant bacteria, exacerbating the problem. The most common causes of resistance are mutations in the bacteria that result in the following:

- Production of drug-inactivating enzymes
- Evolution of altered drug targets with reduced drug affinity
- Decreased permeability of the bacterium to drugs
- Production of drug efflux systems

In 2025, the US Centers for Disease Control (CDC) estimated that, each year in the United States, 2.8 million people are infected with drug-resistant organisms and ~35,000 die from those infections. Even more alarming are estimates that, if resistance continues unchecked, the number of deaths per year worldwide will grow to 10 million. While there are currently many antibacterial drugs, there is clearly a need for new agents due to this escalation in resistant organisms.

12.3 ANTIBACTERIAL DRUG DISCOVERY

In this day and age, it is difficult to appreciate the importance of the discovery of antibacterial drugs. Treatment of bacterial infections with highly toxic metals, such as mercuric chloride and arsenic, was common from the Middle Ages until the early 1900s. At that time, Ehrlich began a systematic investigation of agents to treat syphilis, a significant public health problem. On the basis of the recent identification of the causative agent *Treponema pallidum* and his interest in the use of organic dyes to stain bacteria, he and his colleagues embarked on a study of organoarsenic compounds. Number 606, Salvarsan, or arsphenamine (Figure 12.2), was introduced in 1910. It was effective against *T. pallidum*, as well as the parasite that causes sleeping sickness, but did not target a wide range of bacteria. For many years, the structure of Salvarsan was believed to be analogous to that of azo dyes, containing an As = As double bond; however, the correct structure was determined in 2005 by Brian Nicholson to be a mixture of cyclic species (see Figure 12.2).

Other early drugs used as antiseptics to inhibit the growth of microorganisms were aminoflavins, such as proflavine (see Figure 12.2), introduced in the early 1900s

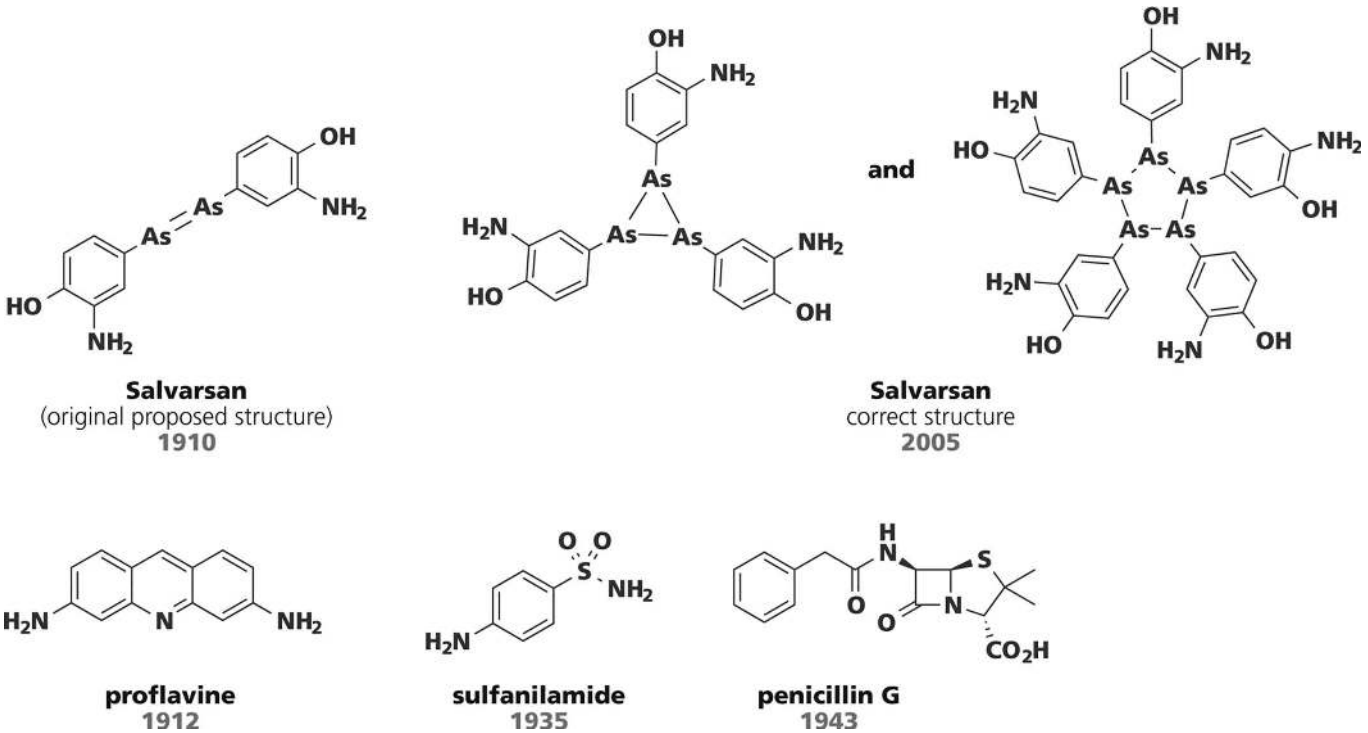


Figure 12.2 Early antibacterial drugs and years introduced. Structures of Salvarsan (purported and actual), proflavine (an aminoflavin), sulfanilamide (from the sulfonamide, or sulfa, drug class), and penicillin G are shown.

to treat topical infections. The first broadly effective and safe antibacterial agents were the sulfonamides (commonly called sulfa drugs), such as the prototype sulfanilamide that was launched in 1935. Soon after, in the early 1940s, the penicillin fungal metabolites represented by penicillin G became widely used. They were bactericidal rather than bacteriostatic like the sulfa drugs, and they had a lower incidence of side effects. Since then, there have been hundreds of antibacterial drugs introduced that led to the successful treatment of most bacterial infections. Despite these advances, there is a continual need for new antibacterial agents, primarily due to the development of drug resistance.

A particular problem is that of nosocomial or hospital-acquired infections, which result in approximately 100,000 deaths per year. Many of these infections are due to drug-resistant organisms. The increased severity of this problem is evidenced by the fact that, in between 1990 and 2021, the number of deaths attributed to *Staphylococcus aureus* (SA) that was resistant to the drug methicillin doubled. Furthermore, a significant portion of methicillin-resistant *Staphylococcus aureus* (MRSA) strains (78%) have developed cross-resistance to other classes of antibacterials.

In 1975, only 2% of strains of *Staphylococcus aureus* (SA) were penicillin-resistant, but by 1988, that number had increased to 80%. Of those, 40% were also resistant to another drug, methicillin, and a significant portion of methicillin-resistant *Staphylococcus aureus* (MRSA) strains (78%) have developed cross-resistance to other classes of antibacterial drugs. Even organisms that were nearly eradicated have made a comeback: *Mycobacterium tuberculosis* infection was the number one cause of infectious disease–related deaths in 2023. Some of the most problematic organisms have been labeled as ESKAPE pathogens, since they cause most hospital infections and escape the action of antibacterial drugs. These include *S. aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* species. Further concerning is that despite the increase in resistant organisms and the lack of efficacy of current drug classes toward them, since between 2017 and 2025, only 12 new antibacterial drugs have been introduced, and only two novel structural classes (cyclic lipopeptides and oxazolidinones) have been introduced between 2000 and 2025.

Over the last ~10 years, increasing attention has been focused on the very real possibility of a return to the pre-antibacterial era. In the United States, promising new developments include passage in 2012 of the Generating Antibiotics Incentive Now Act (GAIN), which gives an extra 5–7 years of patent life for antibacterial drugs. In the

European Union, the Innovative Medicines Initiative New Drugs for Bad Bugs (IMI ND4BB) is a $280 million fund that supports clinical development of new antibacterial drugs, as well as basic research. These initiatives have spurred renewed interest in developing new antibacterial drugs, and some compounds that languished for years due to regulatory issues have recently been approved. One example is dalbavancin, which was discovered in 1996 but did not receive FDA approval until 2014.

Although the terms are often used interchangeably, the term **antibiotic** refers to products such as penicillin that are derived from microorganisms and inhibit bacterial growth. An **antibacterial agent** is any compound, natural or synthetic, that inhibits bacterial growth. Some drugs are **bactericidal**, killing the bacteria, while others are **bacteriostatic** and inhibit replication but do not kill existing bacteria. **Antiseptics** include a number of common compounds, such as phenols and alcohols, that do not have a defined mechanism but are used topically on skin and tissue to inhibit the growth of microorganisms. Some common pathogenic bacteria include the Gram-positive bacteria *S. aureus*, *Streptococcus pyogenes*, and *S. pneumoniae*. One important group of pathogenic Gram-positive bacteria is mycobacteria, with *M. tuberculosis* causing about 2 million deaths per year worldwide. Mycobacteria have an outer waxy cell wall that is difficult for drugs to penetrate. Examples of Gram-negative bacteria that cause disease include *Escherichia coli*, *Neisseria gonorrhoeae*, *Salmonella* species, and *P. aeruginosa*. **Table 12.1** highlights some disease-causing bacteria and classes of drugs that are effective against them.

Some of the components of bacteria targeted by drugs include the cell wall, plasma membrane, ribosomes, and DNA function (see **Figure 12.1a**). The four major categories of antibacterial drugs, based on mechanism of action, include cell wall synthesis inhibitors, protein synthesis inhibitors, nucleic acid synthesis inhibitors, and

Table 12.1 Examples of Common Pathogenic Bacteria, Their Associated Diseases and Conditions, and the Drugs or Drug Classes Used to Treat Each

Bacterium	Disease or Condition	Drugs
Gram-positive bacteria		
Streptococcus pneumoniae	Pneumonia, meningitis	β-lactams, tetracyclines, macrolides
Streptococcus pyogenes	Strep throat, skin infections, rheumatic fever	β-lactams, macrolides
Staphylococcus aureus	Pneumonia, blood infections, skin infections	β-lactams, tetracyclines (resistant: Vancomycin, daptomycin, sulfa),
Clostridium difficile	Diarrhea	Vancomycin, metronidazole
Mycobacterium tuberculosis	Tuberculosis	Isoniazid, rifampicin, ethambutol, pyrazinamide, quinolones, aminoglycosides
Gram-negative bacteria		
Escherichia coli	Food poisoning, gastroenteritis	β-lactams, aminoglycosides, quinolones, sulfonamides
Salmonella species	Food poisoning, typhoid	β-lactams, quinolones, sulfonamides
Neisseria gonorrhoeae	Gonorrhea	β-lactams, macrolides
Shigella species	Food poisoning, dysentery	β-lactams, quinolones
Pseudomonas aeruginosa	Pneumonia, sepsis	β-lactams, aminoglycosides, quinolones
Helicobacter pylori	Gastritis, ulcers	β-lactams, metronidazole, macrolides
Legionella pneumophila	Legionnaire's disease	Macrolides, quinolones, tetracyclines
Haemophilus influenzae	Pneumonia, meningitis	β-lactams, quinolones, macrolides
Borrelia burgdorferi	Lyme disease	Tetracyclines, β-lactams
Treponema pallidum	Syphilis	β-lactams, tetracyclines
Klebsiella pneumoniae	Pneumonia, sepsis	β-lactams, aminoglycosides, quinolones

drugs that affect membrane permeability. Ideal antibacterial drugs work against multiple strains of both Gram-positive and Gram-negative bacteria and are orally active with low toxicity.

12.4 BACTERIAL CELL WALL SYNTHESIS INHIBITORS

The bacterial cell wall structure depends on the cross-linking of **peptidoglycan** strands. The strands are composed of *N*-acetylmuramic acid (NAM) and *N*-acetylglucosamine (NAG) polysaccharides, with peptide strands linked to the NAM strand, containing diaminopimelic acid (DAP) and terminating in D-Ala-D-Ala residues. Cross-linkage is catalyzed by peptidoglycan transpeptidase, a serine protease, also known as **penicillin-binding protein** (PBP), of which there are multiple isoforms. These enzymes react at the amide between the D-Ala-D-Ala residues to generate an acyl-enzyme intermediate, and then the amine of DAP on a second peptidoglycan strand displaces the enzyme. The basic process for the cross-linking step is outlined in Figure 12.3 (see Chapter 7 for the specific mechanism of serine proteases).

Cell wall synthesis inhibitors comprise about 44% of all antibacterial drugs on the market and include the β-lactam antibacterial compounds, as well as vancomycin and its analogs. The discovery of penicillin led the way to the discovery and development of numerous analogs of β-lactams, including the penicillins, cephalosporins, carbapenems, penems, and monobactams.

β-lactam antibacterial compounds are irreversible inhibitors of the bacterial serine protease, with the active site serine adding to the β-lactam's reactive carbonyl, acylating the enzyme (mechanism outlined in Figure 7.23). Vancomycin and its analogs also inhibit cell wall synthesis, but instead of acting as enzyme inhibitors, they bind to the D-Ala-D-Ala portion of the enzyme substrate.

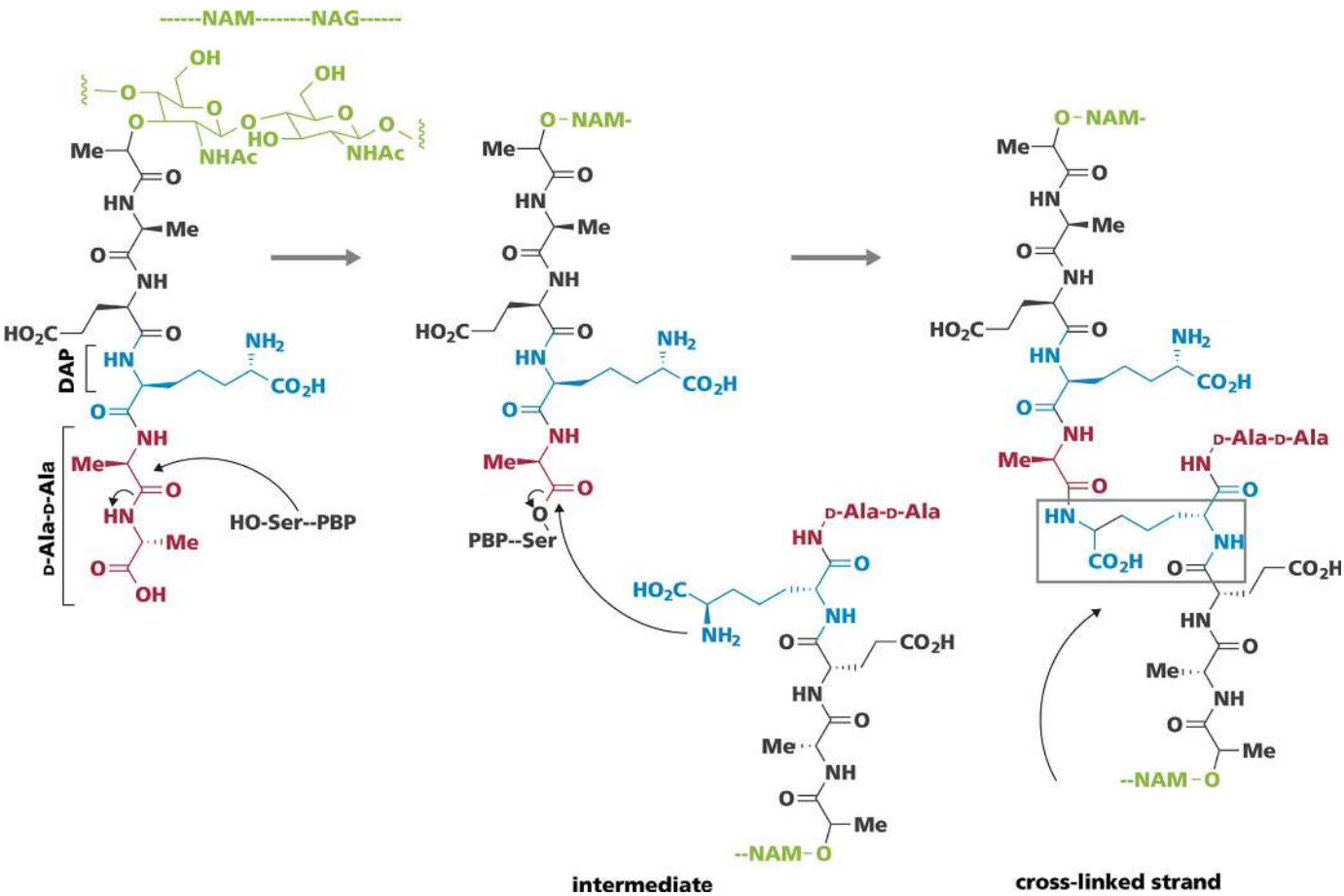
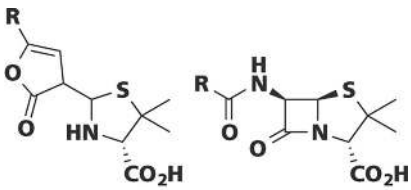


Figure 12.3 Formation of bacterial cell wall cross-links catalyzed by penicillin-binding protein. Cell walls are composed of *N*-acetylmuramic acid (NAM) and *N*-acetylglucosamine (NAG) polysaccharides (green), with peptide strands linked to the NAM strand, containing diaminopimelic acid (DAP, blue) and terminating in D-Ala-D-Ala residues (red). The serine hydroxyl group in PBP attacks the amide bond between the two terminal D-alanine residues of a peptidoglycan strand to form an acyl enzyme intermediate; addition of the amino group of DAP from a second peptidoglycan strand to the acyl-enzyme intermediate forms a cross-link (boxed) and releases the enzyme.

12.4.1 β-Lactam Antibacterial Compounds Are the Largest Class of Cell Wall Synthesis Inhibitors

The discovery of penicillin in 1928 by Alexander Fleming is one of the most important chapters in medicinal chemistry. A bacteriologist, Fleming had previously isolated the enzyme lysozyme and was studying the lysis, or cell breakdown, of bacteria. Upon returning from several weeks of vacation, he observed that an uncovered plate of the bacterium *S. aureus* had been contaminated with a mold, and lysis had occurred where the bacterial colonies were in contact with the mold. He identified the mold as *Penicillium rubrum* (later corrected to *Penicillium notatum*), cultured the broth, and found it to be an effective antibacterial agent against a number of Gram-positive organisms, as well as the Gram-negative *N. gonorrhoeae*. He also injected the mold into rabbits and mice and noted its lack of toxicity. However, he was unable to isolate the active component, and his isolation attempts were abandoned in 1932. Three years later saw the introduction of the synthetic sulfonamide antibacterial drugs, and then in 1938, Ernst Chain and Howard Florey carried out a systematic study of all reported antibacterial compounds. Believing the active ingredient in Fleming's mold to be an enzyme, Chain was able to isolate penicillin in 1938 by using milder techniques than conventional isolations that used acid and base extractions (see Chapter 1.2.1). With the advent of World War II in Europe, the need for antibacterial agents became urgent, and the scale-up and production of penicillin through isolation was primarily done in the United States, at the Merck, Squibb, and Pfizer pharmaceutical companies. Until 1944, penicillin was in short supply, and availability was restricted mainly to the military. The structure of penicillin was not fully clarified until Dorothy Hodgkin elucidated it by use of X-ray crystallography. Until then, the two most common structures proposed were a thiazolidine oxazolone, supported by Sir Robert Robinson, and a β-lactam, proposed by Robert B. Woodward and supported by Ernst Chain (**Figure 12.4**). Hodgkin's X-ray structure unequivocally confirmed the bicyclic β-lactam structure.



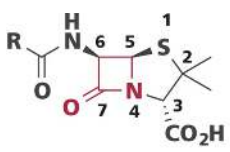
thiazolidine oxazolone **β-lactam structure**

Figure 12.4 Proposed penicillin structures. The incorrect thiazolidine oxazolone and the correct β-lactam structure, confirmed by X-ray crystallography (R = benzyl), are shown.

Penicillins are bicyclic β-lactam antibacterial compounds, with the lactam ring fused to a five-membered thiazolidine ring (**Figure 12.5**). As found by Fleming and later confirmed by Chain and Florey, penicillins exhibit good activity against Gram-positive bacteria.

The original fermentation products were penicillin G (benzylpenicillin) and penicillin V (phenoxymethylpenicillin), which are still in clinical use. Only one enantiomer is active, and requirements for activity include the β-lactam thiazolidine bicyclic core, a free carboxylic acid at C3, and an amide side chain at C6 (see **Figure 12.5**). The amide substituent can be varied, as long as it contains an aryl group. Many active amide substituents also contain an amine or carboxylic acid at the position α to the carbonyl. The specific amide substituent can have a profound influence on the activity of the compound. In some cases, its steric bulk protects the lactam ring from deactivation by β-lactamases (a mechanism of resistance), and in others, it can afford improved permeability. The substitution of either oxygen or carbon for the sulfur atom in the thiazolidine ring maintains activity, but no compounds based on these scaffolds are used clinically. Substitution on one of the C2 methyls with a heterocycle maintains potency. Due to the mechanism of action (reaction with the serine protease), reactivity of the β-lactam carbonyl is essential.

There are approximately 20 penicillin derivatives marketed, with the majority being semisynthetic C6 amide derivatives. Variations in the side chain amide are prepared by selective amide hydrolysis of the readily available natural products penicillin



penicillin structure (β-lactam in red)

STRUCTURE–ACTIVITY RELATIONSHIPS
position 1 = usually sulfur; oxygen or carbon also active
postion 2 = usually dimethyl; substitution on methyl with heterocycle tolerated
position 3 = carboxylic acid required
position 4, 7 = β-lactam moiety required
position 6 = amide; R group variable aryl
stereochemistry: as shown

Figure 12.5 Structure–activity relationships of penicillins. All active penicillins contain the bicyclic core, with the stereochemistry shown. An amide at C6 is required, as is the C3 acid. The sulfur may be replaced with either oxygen or carbon, and one of the C2 methyl groups may be substituted with a heterocycle.

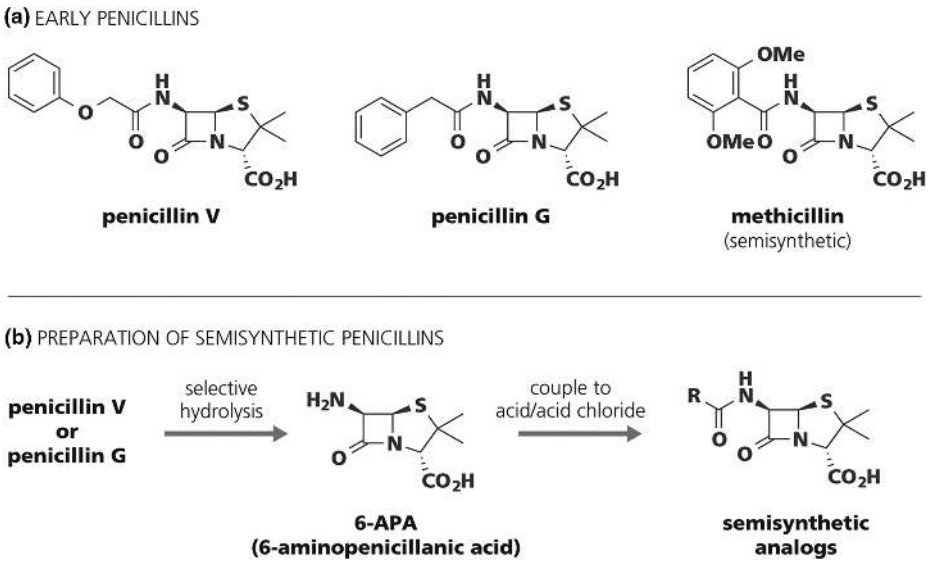


Figure 12.6 Structures of early penicillins and preparation of semisynthetic analogs. (a) Early penicillins include natural products penicillin V and G and semisynthetic analog methicillin. (b) Synthesis of semisynthetic analogs. The hydrolysis of the natural products penicillin V or G affords 6-APA, which is used to prepare various amide semisynthetic derivatives.

V or G (Figure 12.6a) to give 6-aminopenicillanic acid (6-APA), followed by coupling to a wide variety of acids or acid chlorides, as seen in Figure 12.6b.

Numerous semisynthetic analogs have been prepared, and penicillins are still some of the most widely used antibacterial drugs. Methicillin (1959) is an example of an early analog that has some activity against resistant organisms (see Figure 12.6a). All penicillins are susceptible to acid degradation in the stomach, with some having better oral availability than others. Resistance to penicillins has been a problem almost since their introduction, in part due to bacterial production of the penicillin-inactivating enzyme β-lactamase. Some penicillins that have activity against β-lactamase-producing strains include dicloxacillin, oxacillin, and nafcillin. These are active against Gram-positive bacteria exclusively (Figure 12.7).

Amine analogs of penicillin G (aminopenicillins), such as amoxicillin and ampicillin, show improved acid stability as well as activity against some Gram-negative bacteria, such as *E. coli* and *H. influenzae*. Of the two, amoxicillin has better oral activity than ampicillin, which is used as its prodrug, pivampicillin. The prodrug contains a pivaloyloxymethyl ester at C3 that is hydrolyzed *in vivo* to give ampicillin. The ester, although inactive until hydrolyzed, provides greater oral bioavailability. The carboxypenicillins and ureidopenicillins have a broader spectrum of activity than the earlier penicillins and can be used to treat Gram-negative infections. The carboxypenicillins, examples of which are ticarcillin and carbenicillin (used for veterinary medicine), contain a carboxylic acid group as part of the C6 amide side chain. The ureidopenicillins, such as mezlocillin, azlocillin, and piperacillin, were introduced in the 1980s and can be considered analogs of the aminopenicillins. They have improved activity against *P. aeruginosa* as well as some anaerobic bacteria. Of these, piperacillin is still used clinically.

As mentioned earlier, a major antibiotic resistance mechanism is the production of β-lactamase, an enzyme that evolved from PBP and hydrolyzes the β-lactam ring in penicillins and other β-lactam antibacterial compounds. It has no function in cell wall biosynthesis and serves only to inactivate the β-lactams. Successful strategies to mitigate this resistance mechanism include identifying drugs with low susceptibility to β-lactamase and identifying inhibitors of it that can be co-administered with β-lactam antibacterial drugs. The earliest β-lactamase inhibitor used is clavulanic acid, a natural product isolated from *Streptomyces clavuligerus* in the 1970s (Figure 12.8). Although this oxapenem possesses a bicyclic β-lactam, it has no amide at C6 and, therefore, has only minimal antibacterial activity. However, it is a potent inhibitor of the β-lactamase enzyme and is an example of a suicide inhibitor, or mechanism-based inactivator (see Figure 7.24).

To increase its effectiveness, amoxicillin is widely prescribed in combination with the β-lactamase inhibitor clavulanate, and the carboxypenicillin ticarcillin

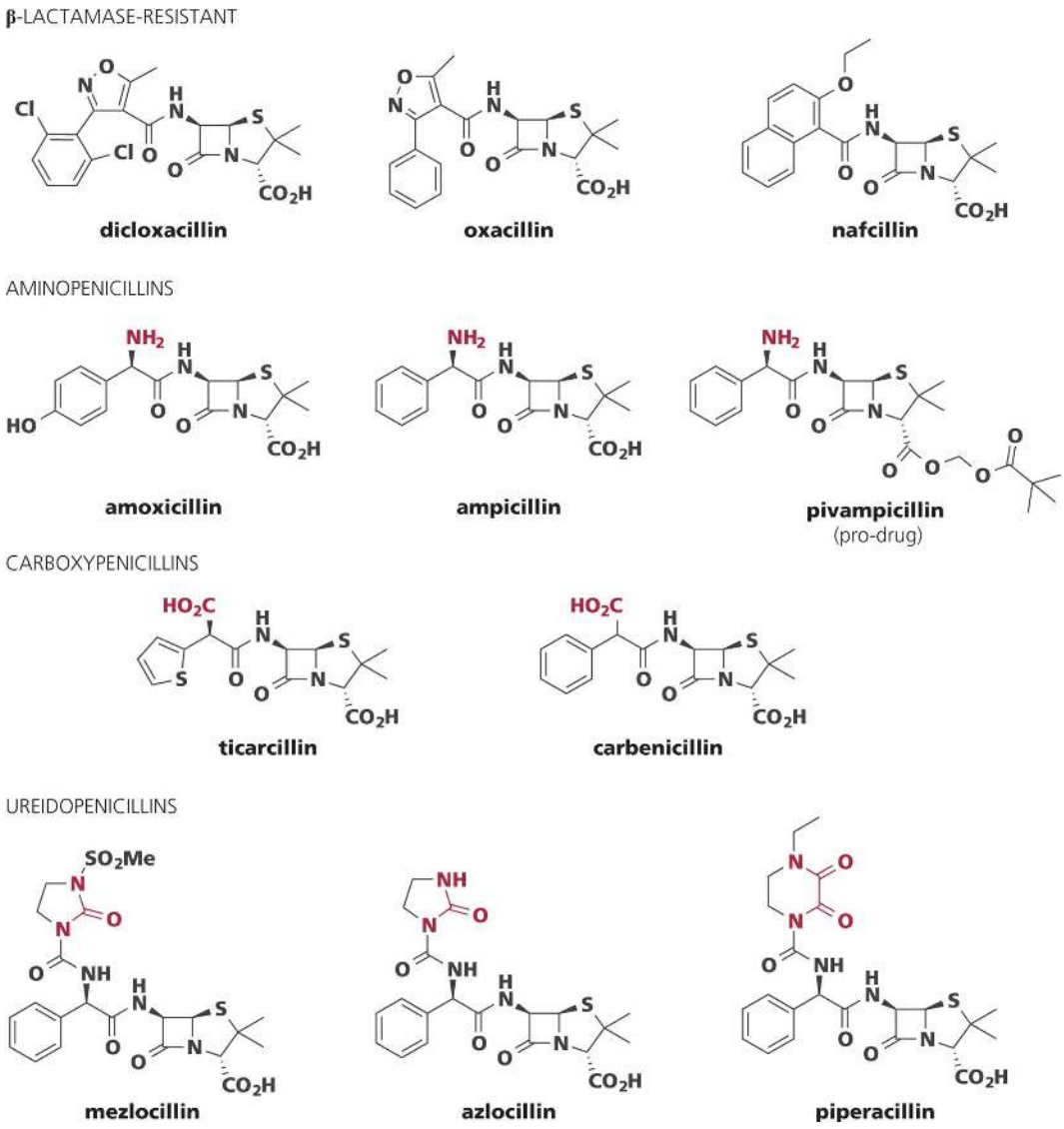


Figure 12.7 Structures of semisynthetic penicillins. Families include β-lactamase-resistant penicillins; aminopenicillins, characterized by the presence of a primary amine; carboxypenicillins, which all contain a carboxylic acid in the amide side chain; and ureidopenicillins, typified by a cyclic urea in the amide side chain. Common structural features within each family are shown in red.

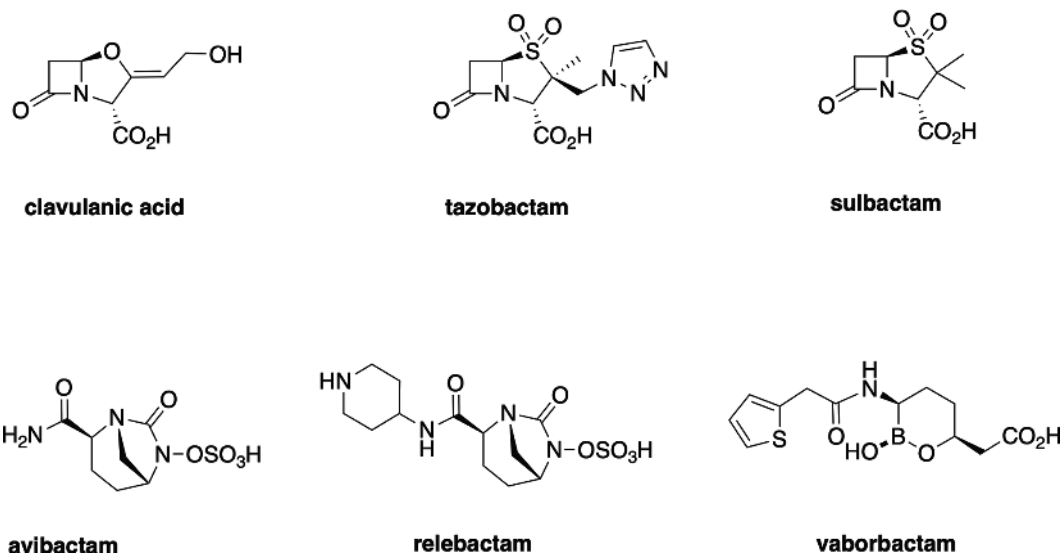


Figure 12.8 Structures of β-lactamase inhibitors. The natural product clavulanic acid and five other inhibitors are shown, which are used in combination with penicillins, carbapenems, penems, and cephalosporins to minimize resistance mechanism via β-lactamase inhibition.

is used almost exclusively in combination with clavulanate. Other β-lactamase inhibitors used in combination with penicillins include the penicillin sulfones tazobactam and sulbactam (see **Figure 12.8**). Avibactam and relebactam are novel diazabicyclooctanes that bind reversibly to β-lactamases and are effective against a wide range of β-lactamases. Vaborbactam is a cyclic boronic acid that was identified through a specific search for compounds that would inhibit *Klebsiella pneumonia* carbapenemases, which are the most prevalent carbapenemases. Recent combinations of these β-lactamase inhibitors with β-lactam antibiotics include ceftazidime/avibactam, meropenem/vaborbactam, and imipenem/relebactam.

Another resistance mechanism used by bacteria is the production of PBPs that can still catalyze cell wall biosynthesis but have reduced affinity for penicillins. This type of resistance is seen in MRSA, an organism that carries the *mecA* gene, which encodes a mutant PBP. This modified enzyme does not bind the drug methicillin but still catalyzes cell wall biosynthesis. Other resistant bacteria with altered PBPs include *S. pneumoniae* and *N. gonorrhoeae*.

Cephalosporins were the second family of β-lactam antibiotics to be isolated. Cephalosporin C was found in a fungus by Giuseppe Brotzu in Sardinia in the early 1940s. This discovery was prompted by efforts to reduce the incidence of typhoid fever (caused by *Salmonella typhi*) in Sardinia. The *S. typhi* organism was found in the sewers, but not in nearby seawater, so Brotzu cultured the seawater and identified a fungus (*Cephalosporium acremonium*) responsible for the antibacterial activity. He isolated an active component that he named mycetin and used it to treat patients with typhoid as well as other bacterial infections. During World War II, he was unable to pursue further development, but after the war, a sample of the fungus was sent to Florey in England. Ultimately, in 1962, Florey and Edward Abraham produced and identified cephalosporin C from the fungus. The structure of cephalosporins differs from penicillins by replacement of the thiazolidine five-membered ring with a six-membered dihydrothiazine ring (**Figure 12.9**), but both contain the strained β-lactam ring that is essential for activity. In addition, cephalosporins may contain a leaving group at the 3′-position, although this is not necessary for activity. A simplified mechanism shows attack of a PBP serine hydroxyl group on the β-lactam carbonyl, followed by ring opening and expulsion of the leaving group (when present), resulting in a covalent bond to the inactivated enzyme (see **Figure 12.9**).

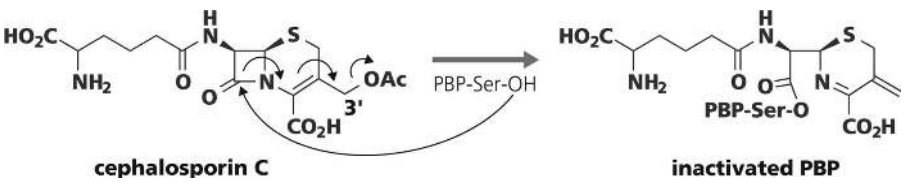
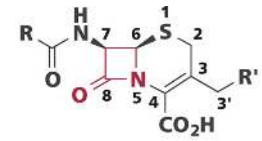


Figure 12.9 Abbreviated mechanism for inactivation of penicillin-binding protein by cephalosporin C. A serine hydroxyl of PBP attacks the β-lactam carbonyl, and a leaving group at C3′ is expelled when the β-lactam ring opens, resulting in a covalently modified, inactivated PBP.

Cephalosporin C itself did not possess very potent antibacterial activity, and the most frequently used cephalosporins are semisynthetic derivatives with variations at both the side chain amide and the 3- or 3′-position (**Figure 12.10**). Like the penicillins, requirements for activity include the bicyclic core structure with the embedded β-lactam with the stereochemistry as drawn, a carboxylic acid at C4, and



cephalosporin structure (β-lactam in red)

STRUCTURE–ACTIVITY RELATIONSHIPS
position 1 = usually sulfur; oxygen or carbon also active
position 3 = Cl, vinyl, propenyl, methyl, CH_2R (R = OAc, heterocycle)
postion 3' (R') = acetate (Gram-positive), heterocycle (improved Gram-negative)
position 4 = carboxylic acid
position 5, 8 = β-lactam moiety required
position 7 = amide; many R groups tolerated
stereochemistry: as shown

Figure 12.10 Structure–activity relationships of cephalosporins. All active cephalosporins contain a β-lactam embedded within a bicyclic core, with stereochemistry as drawn. An amide at C7 is required, as is the C4 acid. The sulfur atom may be replaced with either oxygen or carbon. The C3 position may be substituted with a small lipophilic group or methylene group containing a leaving group or heterocycle.

an amide side chain at C7. Various substitutions at position 3 are tolerated. As in cephalosporin C, C3 sometimes has a methylene substituted with a leaving group, such as acetoxy. However, this is not a requirement, and other groups at C3, such as chloro, vinyl, methyl, or other substituted methyl groups, are also tolerated. As with the penicillins, the sulfur may be replaced with either oxygen or carbon without loss of activity. While the β-lactam is required for activity, the specific amide and C3 substituents influence the spectrum of antibacterial activity, as well as physical and pharmaceutical properties. For example, when the C3 substituent is CH_2OAc, analogs exhibit Gram-positive activity, but when it is a CH_2-heterocycle, the derivatives have an improved Gram-negative profile.

The first approved cephalosporin was introduced in 1964. Cephalosporins are classified by generation, corresponding roughly to the era of introduction, as well as by spectrum of activity (**Figure 12.11**). The first-generation cephalosporins have penicillin-like amide side chains, and they exhibit an activity profile similar to the early penicillins: good activity against Gram-positive bacteria but limited activity against Gram-negative bacteria. Examples include cefalotin (also known as cephalothin), which requires parenteral administration, due in part to acetoxy hydrolysis, as well as cefazolin and cefalexin. Most have limited oral availability, with the exception of cefalexin. In the late 1970s and 1980s, medicinal chemistry efforts focused on variation of the C7 amide and the 3-position, leading to second-generation cephalosporins, such as cefuroxime, cefoxitin, and cefaclor. These are all characterized by improved Gram-negative potency but reduced Gram-positive potency. The oxime in the amide side chain of cefuroxime provides improved stability toward β-lactamase. Cefoxitin is an example of a cephamycin, a semisynthetic derivative of a naturally occurring cephem that possesses an α-methoxy group at C7.

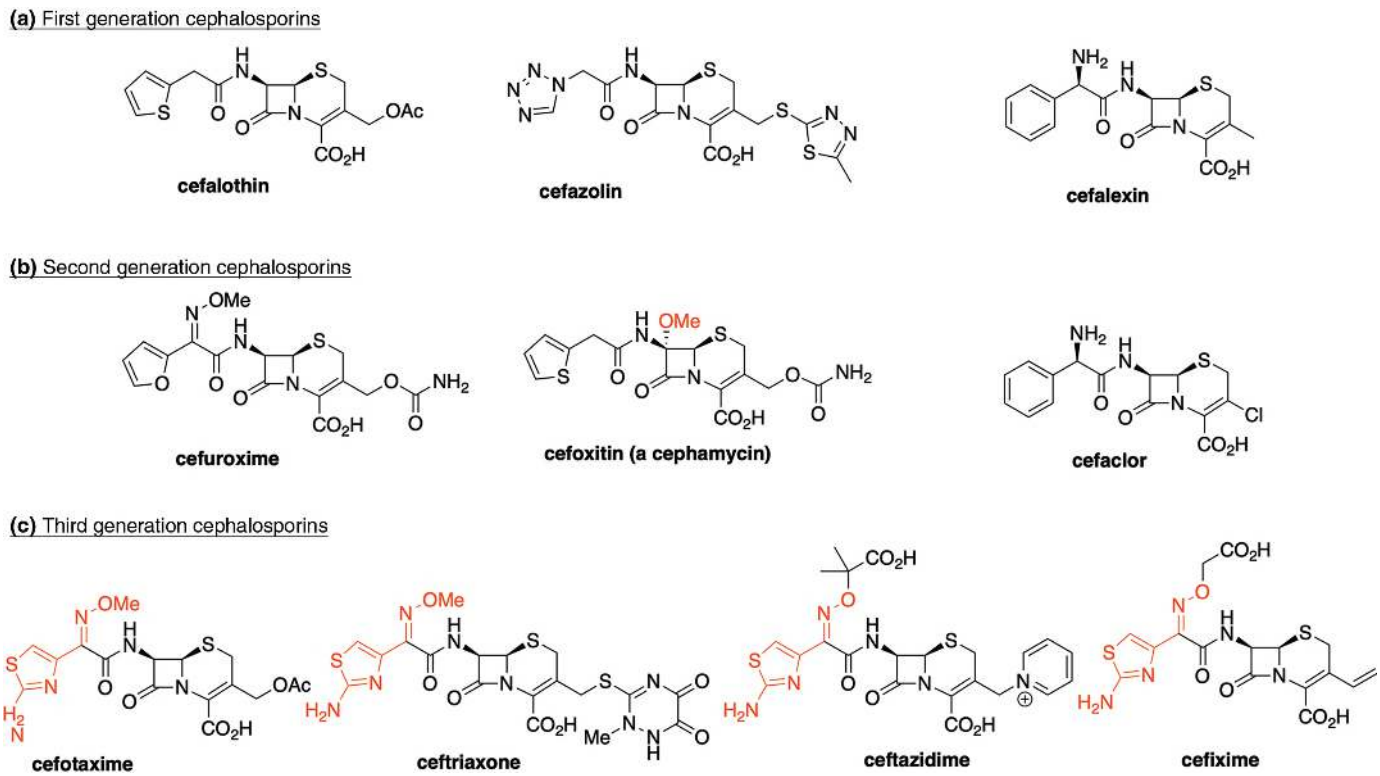


Figure 12.11 First- through third-generation cephalosporin structures. Generations correspond roughly to the era of introduction as well as the spectrum of activity. First-generation drugs (1964–1970s) contain penicillin-like side chains and have activity against gram-positive organisms; second-generation cephalosporins, such as cefuroxime and cefaclor, introduced in the late 1970s, have improved activity toward gram-negative bacteria, with cephamycins, such as cefoxitin (methoxy group, shown in red), adding anaerobic activity against organisms typically found in the GI tract. Third-generation cephalosporins (1980s) have improved activity against gram-negative bacteria; many contain a common aminothiazole oxime, shown in red.

Further discovery efforts in cephalosporins were driven by a focus on achieving a broader spectrum of activity, including activity against resistant organisms, and enhancements in pharmacokinetics. Major improvements in Gram-negative activity were seen with the third-generation cephalosporins developed in the 1980s. The major structural change involved adding an aminothiazole ring to the oxime amide side chain that had proved successful in the second-generation analog cefuroxime. This combination is found in all third-generation cephalosporins. These changes both broaden the spectrum of activity and improve β-lactamase stability. Examples include cefotaxime, ceftriaxone (which has a long half-life and can be dosed once a day), ceftazidime (which has improved activity against *P. aeruginosa*), and cefixime. Many of these drugs are not orally available, with the exception of cefixime.

The newest cephalosporins are considered members of the fourth and fifth generations (**Figure 12.12**). Those in the fourth generation have a broad spectrum of activity, as well as activity against problem organisms, such as the ESKAPE pathogens *P. aeruginosa* and *K. pneumoniae*. These compounds retain the aminothiazole oxime-containing amide side chains, and most, such as cefepime and cefpirome, are zwitterions with improved penetration into Gram-negative organisms. The most recent include ceftaroline, cefiderocol, and ceftobipirole, fifth-generation cephalosporins. Ceftaroline has a broad spectrum of activity, including activity against MRSA, due to the 1,3-thiazole at the C3´-position. Ceftobiprole medocaril is a prodrug of ceftobiprole, with activity against resistant *S. aureus* blood infections, as well as pneumonia. Cefiderocol is classified as a siderophore cephalosporin; it contains a catechol side chain that chelates iron and facilitates entry into the bacteria.

Resistance to cephalosporins occurs through the production of β-lactamase and altered PBPs, as for penicillins. A third prominent mechanism in Gram-negative

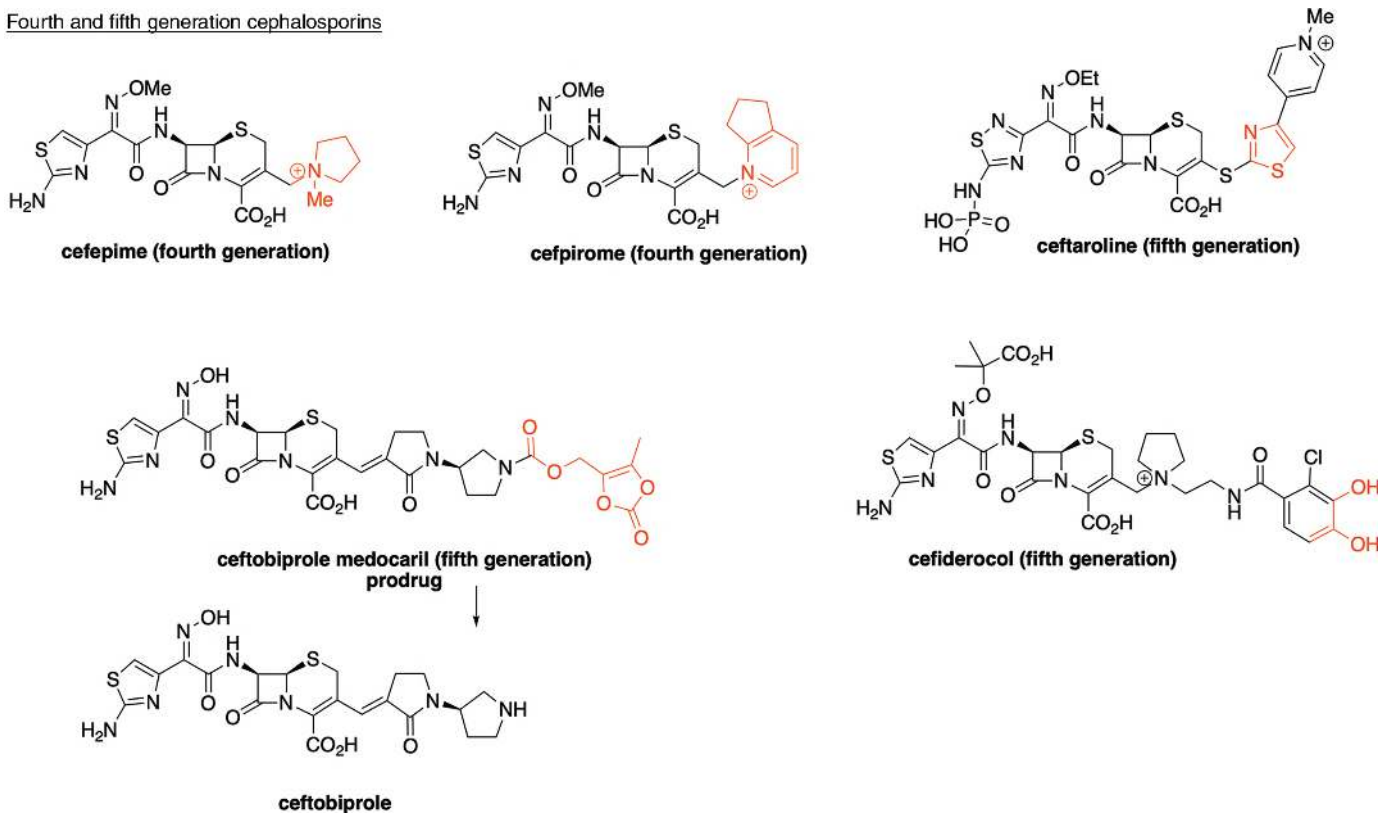


Figure 12.12 Structures of fourth- and fifth-generation cephalosporins. Fourth-generation compounds (1990s) cefepime and cefpirome have an extended spectrum of activity against Gram-negative organisms and some resistant organisms and are characterized by a cationic C3′-substituent (red). The fifth-generation cephalosporin ceftaroline is a broad-spectrum antibacterial agent, with Gram-positive and Gram-negative activity, as well as activity against resistant organisms, such as MRSA, due to the 1,3-thiazole ring (red) at C3′. Ceftobiprole medocaril is a prodrug (red) of ceftobiprole, and cefiderocol is the first siderophore cephalosporin (catechol in red).

organisms is the production of mutant porins, which reduces the ability of cephalosporins to penetrate the bacterial cell wall. Porins are transmembrane channels in the outer membrane of Gram-negative bacteria, through which antibacterial compounds, including β-lactams, enter the cell. Mutations in the genes that encode porins produce modified channels that restrict the entry of antibacterial drugs.

The **penems and carbapenems** are a much smaller family of bicyclic β-lactam antibacterial drugs than the penicillins and cephalosporins (Figure 12.13). While the penems are synthetic, the carbapenems constitute both natural and semisynthetic products. Both contain a β-lactam fused to an unsaturated five-membered ring. The penem structure was originally proposed as a hybrid of penicillin (4,5 ring system) and cephalosporin (unsaturated sulfur-containing ring) by Woodward, and analogs with and without an amide side chain were prepared but were found to be unstable. Carbapenems are the corresponding analogs in which the sulfur atom is replaced with a carbon. The discovery of the naturally occurring carbapenem thienamycin in 1976, which had the unusual hydroxyethyl side chain, inspired the synthesis of the more stable and less toxic amidine analog imipenem. Both penems and carbapenems are susceptible to hydrolysis by the kidney enzyme renal dehydropeptidase (DHP); therefore, one strategy is to co-administer a DHP inhibitor, such as cilastatin (see Figure 12.12). The combination of imipenem and cilastatin is highly effective and is marketed as Primaxin. The addition of a methyl group on the five-membered ring in a carbapenem affords analogs such as meropenem with improved DHP stability. Faropenem, an example of a penem, was the only approved member of the penem class until 2024, when sulopenem was approved to treat urinary tract infections.

As a class, both natural and semisynthetic carbapenems, as well as hydroxyethyl penems, have a broad spectrum of antibacterial activity and are particularly effective against many types of anaerobic bacteria. They have also recently been found to be active against *M. tuberculosis*, unlike other β-lactam antibacterial drugs.

Monobactams are monocyclic β-lactams (Figure 12.14) first reported by groups at Takeda and Squibb Institute for Medical Research; aztreonam is the one example on the market. As a class, they are particularly potent against Gram-negative bacteria. In the late 1970s and early 1980s, scientists at these organizations screened various sources of natural products for antibacterial activity using bacterial strains that were genetically engineered to be overly sensitive to β-lactams. One of the earliest reported was SQ 26,445 (see Figure 12.14). Most of the naturally occurring monobactams contain a sulfonate at position 1, and many possess a methoxy substituent at C3, along with simple amide side chains.

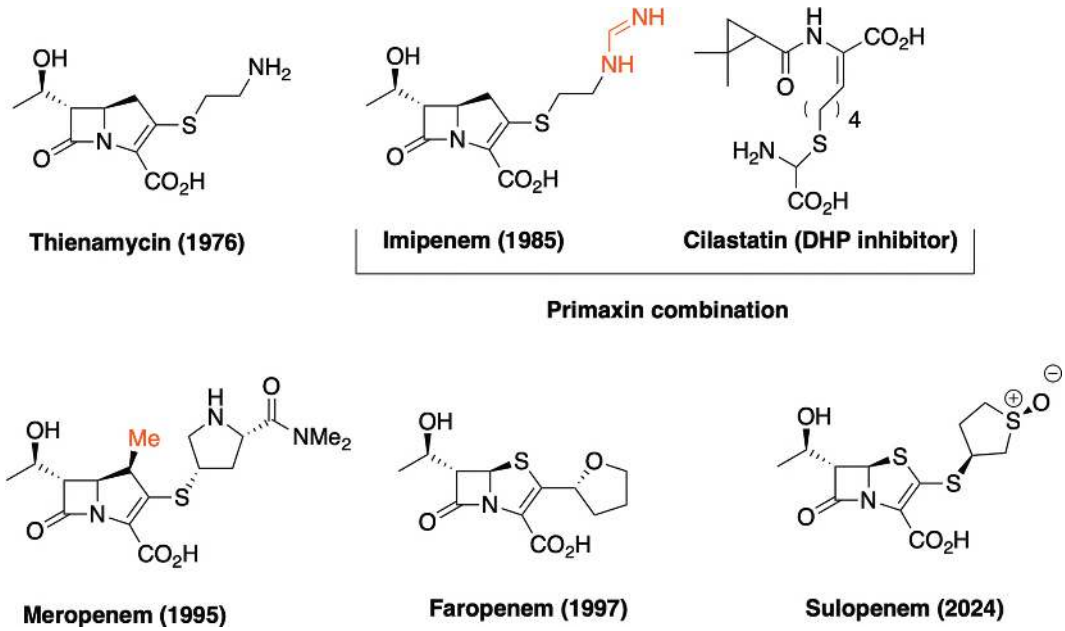
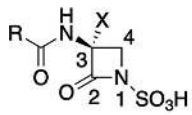
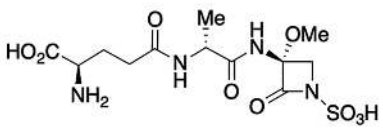


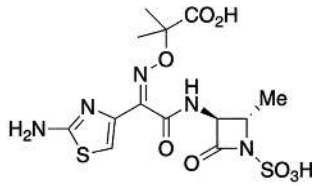
Figure 12.13 Penem and carbapenem structures. Thienamycin, a carbapenem natural product, provided insight into how to generate less toxic and more stable synthetic analogs, such as imipenem and meropenem. The combination of imipenem and cilastatin, an inhibitor of renal dehydropeptidase (DHP), sold as Primaxin, extends the half-life of imipenem. The methyl analog meropenem (methyl group, shown in red) has improved stability against DHP. Until 2024, faropenem was the only approved member of the penem class, with the orally active sulopenem approved in 2024.



monobactam structure
X = OMe, H



SQ 26,445



aztreonam

Figure 12.14 Monobactam structures. All monobactams have the general structure on the left with either a proton or methoxy at the 3-position (red). SQ 26,445 was one of the first monobactams reported. The semisynthetic aztreonam, containing a third-generation cephalosporin-like aminothiazole oxime, has improved activity compared to the natural product as well as β-lactamase stability and was approved in 1986.

Although the early monobactams typically had only modest antibacterial activity, many synthetic analogs bearing both penicillin- and cephalosporin-like side chains were investigated. The N-sulfonate was found to be necessary for activity; however, the C3 ´-methoxy group could be eliminated. A substituent at C4 improved both potency and resistance to β-lactamase, as did the incorporation of the aminothiazole oxime amide side chain found in third-generation cephalosporins. A 2-methyl-2-propionic acid moiety on the oxime imparts improved potency against *Pseudomonas*. All these SAR features are evident in aztreonam (see **Figure 12.14**), a monobactam introduced in 1986 and still used today to treat a variety of infections caused by Gram-negative bacteria. Its combination with the β-lactamase inhibitor avibactam to treat certain complicated Gram-negative infections was recently approved.

12.4.2 Glycopeptide Antibacterial Agents Inhibit Cell Wall Synthesis by a Different Mechanism Than β-Lactams

Vancomycin and teicoplanin are dalbaheptide glycopeptide antibacterial agents; the term *dalbaheptide* was derived from D-Ala-binding antibiotics having a heptapeptide structure. Vancomycin was isolated in 1953 from a soil bacterium found in Borneo, as a result of a search by Eli Lilly for antibacterial agents with activity against penicillin-resistant *Staphylococcus*. Although vancomycin was approved in 1958, toxicity attributed to impurities in early preparations, as well as the introduction of methicillin in the same year, led to infrequent use of this intravenous drug. In the 1980s, vancomycin began to be used much more frequently to treat postoperative colitis, usually caused by *S. aureus* or *Clostridium difficile*, as well as infections caused by MRSA. Teicoplanin was isolated from a soil bacterium in 1984 and has the same dalbaheptide core as vancomycin (**Figure 12.15**). It contains an additional diphenyl ether bridge, as well as two glycosides (D-mannose and *N*-Ac-glucosamine) and an amide group in place of the L-vancosaminyl unit found in vancomycin. This amide substituent is typically a C8 or C9 chain in place of the L-vancosaminyl carbohydrate of vancomycin. Dalbavancin is very similar in structure to teicoplanin but has one less carbohydrate unit (*N*-Ac-glucosamine) and converts the acid to an amide, as well as other minor variations. Other semisynthetic analogs include telavancin and oritavancin,

Like the β-lactam antibacterial drugs, vancomycin, teicoplanin, and dalbavancin inhibit the synthesis of the peptidoglycan component of the bacterial cell wall. However, they are not enzyme inhibitors but bind to the peptidoglycan substrate through a series of hydrogen bonds, thus preventing the enzyme from binding to its substrate.

An X-ray structure of vancomycin bound to L-Lys-D-Ala-D-Ala is shown in **Figure 12.16**. This mode of binding includes five hydrogen bonds between vancomycin and the carboxylate and the backbone of the D-Ala-D-Ala fragment and can be considered an example of a drug interrupting a protein–protein interaction, such as those described in Chapter 8.

Vancomycin resistance is conferred by mutations that change the target D-Ala-D-Ala terminus of the peptidoglycan strand to D-Ala-D-lactate, replacing the peptide bond with an ester or D-Ala-D-Ser, sterically extending the methyl group. These simple changes result in reduced affinity of the target for the drug by 1,000-fold. Most mutations that result in resistance to vancomycin do not confer resistance to teicoplanin.

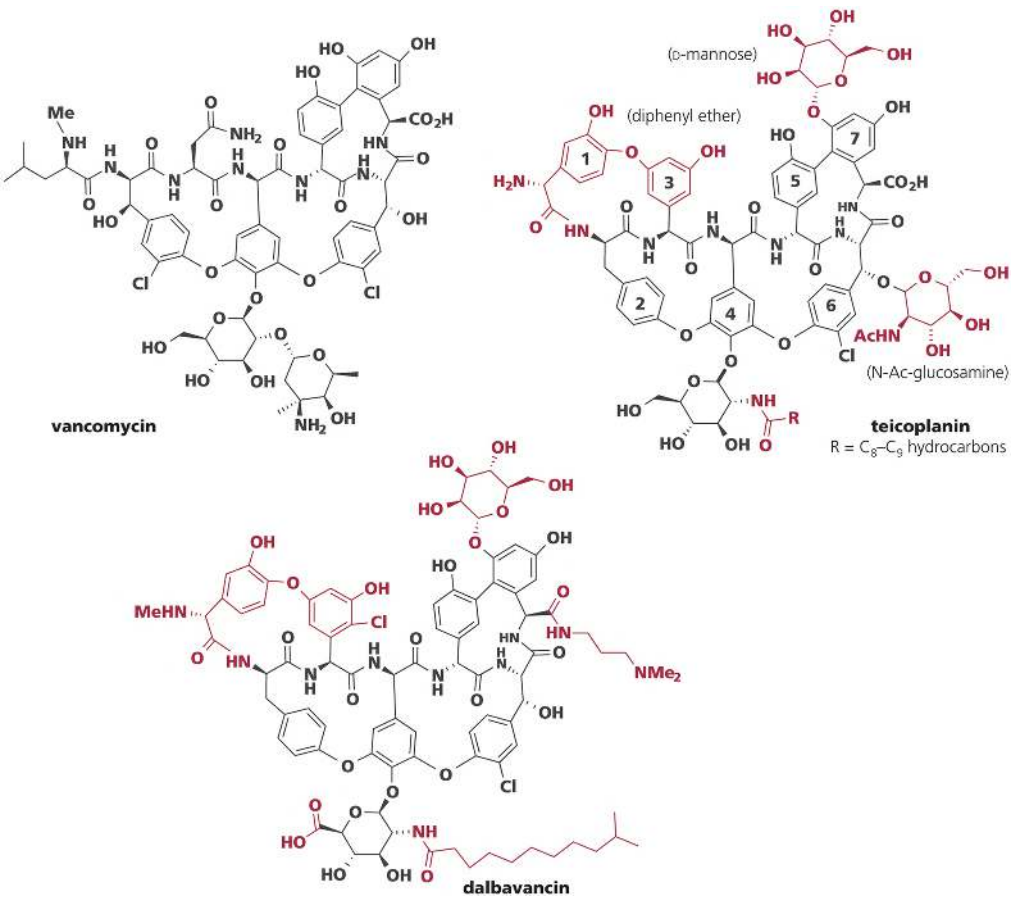


Figure 12.15 Glycopeptide antibiotics. Structures of vancomycin, teicoplanin, and dalbavancin are presented (the aryl amino acid numbering system is shown for teicoplanin). Structural changes compared to vancomycin are shown in red, with the cores in teicoplanin and dalbavancin that are common to vancomycin in black.

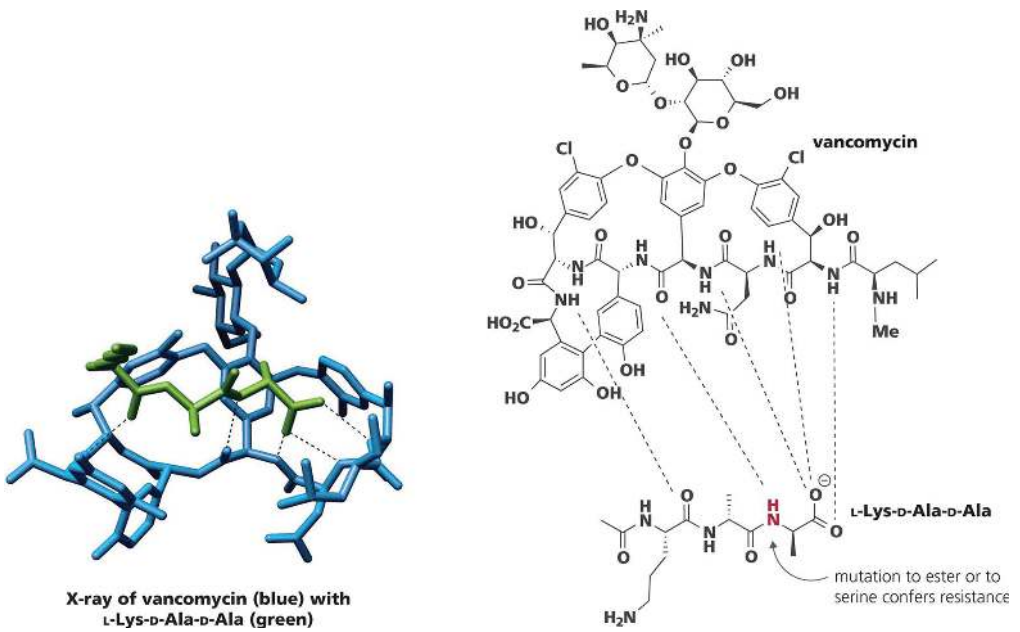


Figure 12.16 X-ray structure of vancomycin and details of binding interactions. *Left*, vancomycin (blue) hydrogen-bonded to L-Lys-D-Ala-D-Ala (green). *Right*, schematic of the five hydrogen bonds (shown as dotted lines) between four amide N-H groups and one carbonyl of vancomycin to the L-Lys-D-Ala-D-Ala substrate. Resistant organisms produce mutant D-Ala-D-Ala-lactate with an ester linkage, resulting in the loss of a hydrogen bond and reduced affinity, or a change to serine.

Source: Photo by M. Stone/Wikimedia Commons. CC BY-SA 3.0.

12.5 BACTERIAL PROTEIN SYNTHESIS INHIBITORS

While many antibacterial agents target the structure of the organism, another class affects the ability to replicate and survive by inhibiting bacterial protein synthesis. There are four major families of drugs in this class: aminoglycosides, tetracyclines, macrolides, and oxazolidinones. All but the oxazolidinones are natural or semisynthetic products, and all act by a similar mechanism, binding to bacterial ribosomes. This binding leads to the synthesis of faulty bacterial proteins so that eventually protein synthesis is arrested and bacterial cell death occurs.

The process of protein synthesis is outlined in **Figure 12.17**, showing messenger RNA (mRNA) binding to each incoming aminoacyl-transfer RNA (tRNA) in the ribosome. As each tRNA enters the ribosomal machinery, carrying an individual amino acid, it binds to the aminoacyl (A) site. Here, proofreading occurs to ensure the correct amino acid is matched to the corresponding codon. The tRNA then moves to the peptidyl (P) site, where its amino acid is transferred to the growing peptide chain. Finally, the tRNA shifts to the exit (E) site, where it leaves the ribosome.

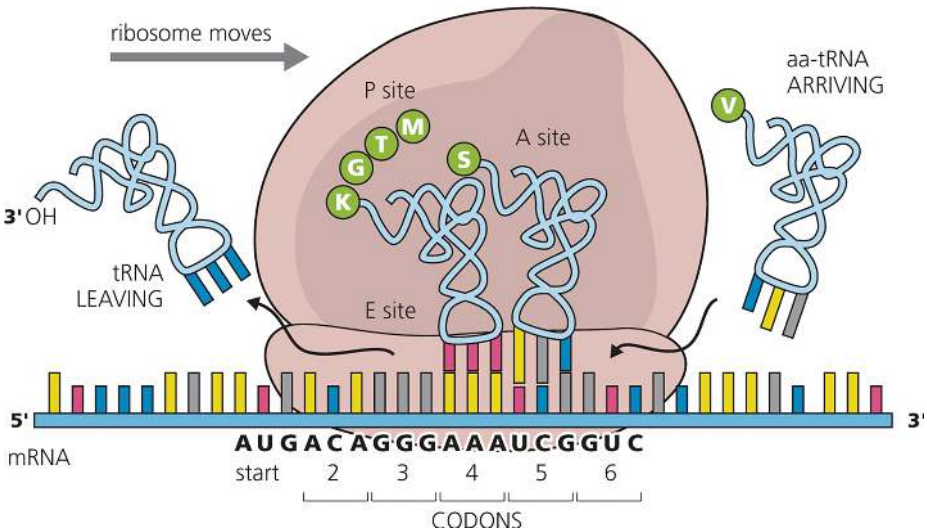


Figure 12.17 Process of protein synthesis on the ribosome. The ribosome, where protein synthesis occurs, consists of multiple proteins and several ribosomal RNA molecules. A template mRNA guides protein synthesis; the two subunits of the bacterial ribosome are shaded pink: The lower part is the 30S subunit, and the upper part is the 50S subunit. On the ribosome, charged aminoacyl-tRNA enters the A site carrying an amino acid; the amino acid is transferred to the growing peptide strand in the P site, and then the tRNA moves to the E site, where uncharged tRNA exits.

Source: Adapted from Kuriyan, J.; Konforti, B.; and Wemmer, D. (2012). *The Molecules of Life: Physical and Chemical Principles*. Garland Science.

Ribosomes are complexes of ribosomal RNA (rRNA) and proteins and are characterized in terms of Svedberg units (S). Bacteria contain 70S ribosomes, consisting of a 30S and 50S subunit. The function of the 30S subunit, which consists of protein and 16S rRNA, is to decode the incoming tRNA so that the growing peptide is translated properly. Drug selectivity for bacterial ribosomes vs. human ribosomes is conferred by both location and structure. Eukaryotic ribosomes are larger (80S) than bacterial ribosomes and have a different rRNA and protein composition. They are located in the cytosol or on the endoplasmic reticulum, whereas in bacteria, the ribosomes are located in the cytoplasm.

12.5.1 Aminoglycosides, Tetracyclines, and Erythromycins Are Protein Synthesis Inhibitors Based on Natural Products

Aminoglycosides include the natural products streptomycin, paromomycin, neomycin, gentamicin, and tobramycin, as well as the semisynthetic amikacin

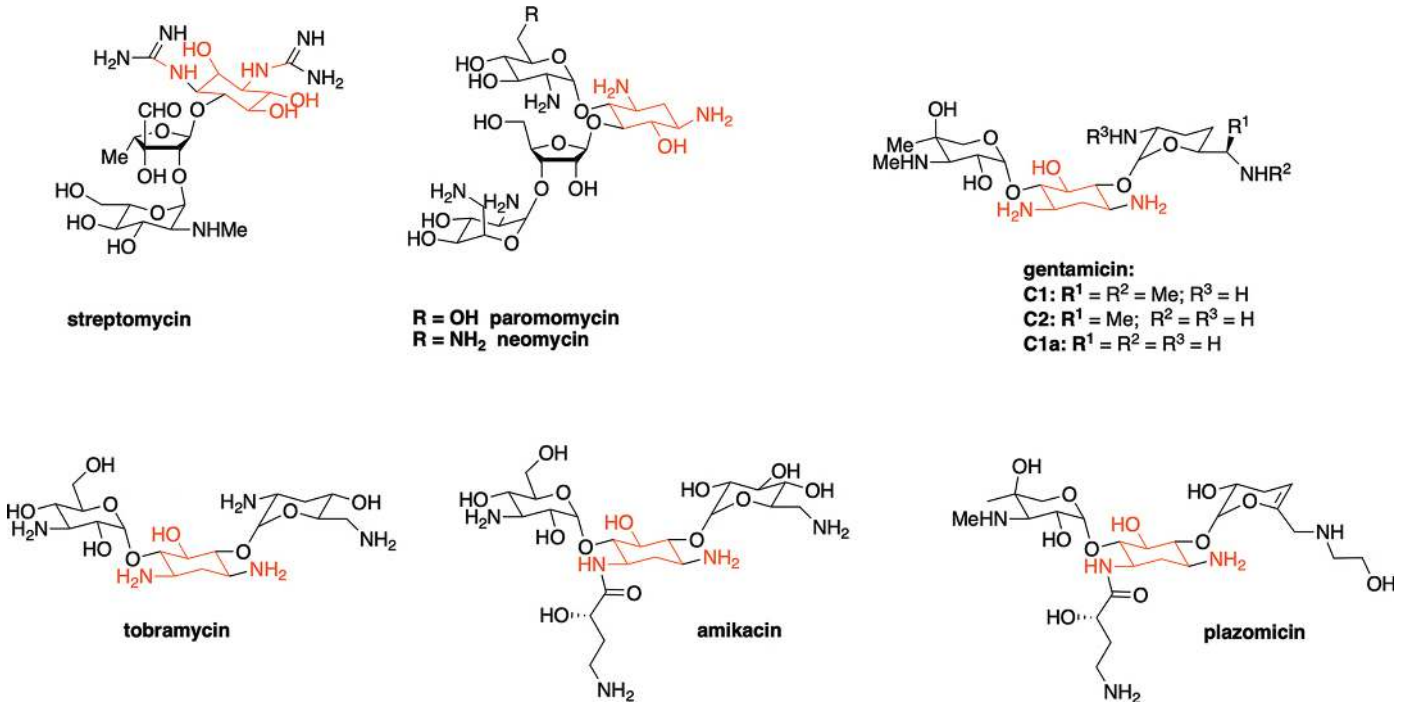


Figure 12.18 Structures of aminoglycoside antibiotics. Common structural features include an aminocyclitol unit called 2-deoxystreptamine (shown in red; note that streptomycin is hydroxylated at position 2) and amino sugars. Streptomycin, paromomycin, and neomycin contain a furanose and a pyranose unit; gentamicin, tobramycin, and the semisynthetic amikacin (from kanamycin) are substituted with pyranose units only, as is plazomicin.

(**Figure 12.18**). Members of this family are characterized by one or more amino sugars and an aminocyclitol, which is a cyclohexane substituted with amino and hydroxyl groups. Streptomycin, isolated by Selman Waksman in 1943 from the soil microbe *Streptomyces griseus*, was the first aminoglycoside reported. It is particularly useful in treating mycobacterial infections, primarily tuberculosis. This was the first example of a natural product isolated from a soil microbe, and it initiated an era of exploration of these organisms as sources of other antibiotics. Paromomycin was introduced in 1960 and is used to treat intestinal infections as well as the parasitic infection leishmaniasis; the closely related neomycin is primarily used to treat topical infections. Gentamicin is a mixture of three closely related natural products (C1, C2, and C1a). Gentamicin, amikacin, tobramycin, and plazomicin are used to treat *E. coli* and *K. pneumoniae* infections. Although aminoglycosides exhibit broad activity against a range of Gram-negative organisms, their use is limited by kidney and ear toxicity (ototoxicity). Structural requirements for activity include the central aminocyclitol and the pendant amino sugars. Resistance occurs by the production of aminoglycoside transferase enzymes that modify the amine and alcohol substituents of the drug. Both tobramycin and the semisynthetic amide amikacin have activity against resistant organisms.

The aminoglycosides bind to ribosomal RNA and associated proteins, specifically the A site of the 16S rRNA portion of the 30S subunit of bacterial ribosomes, resulting in inaccurate translational reading. Insight into the mechanism of aminoglycosides comes from an X-ray co-structure with the 30S subunit. **Figure 12.19** shows hydrogen bonding and ionic interactions between streptomycin and specific residues in the A site. The hydroxyl, aldehyde, amino, and guanidine groups of streptomycin make interactions with phosphates and hydroxyl groups of the RNA and with a lysine residue in the S12 protein region of the ribosome.

Accuracy in recognition of tRNA in the A site is related to two conformations, termed accuracy-prone and error-prone. One segment proposed to be involved in the conformational change between the two is helix 27 (yellow in **Figure 12.19b**), sometimes called the accuracy helix, which adopts different conformations in each state. When tRNA binds to the error-prone conformation, it leads to faulty protein translation, while binding to the other conformation leads to accurate translation. Streptomycin binds to and stabilizes the error-prone form, leading to mistakes in the translation of proteins, the arrest of protein synthesis, and bacterial growth inhibition.

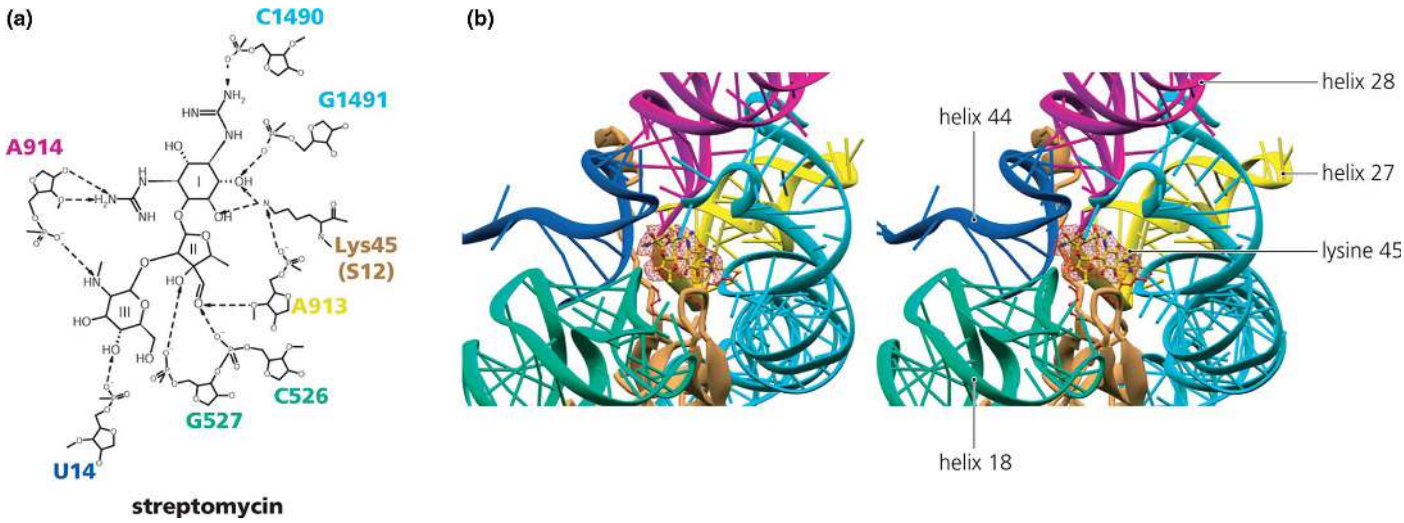
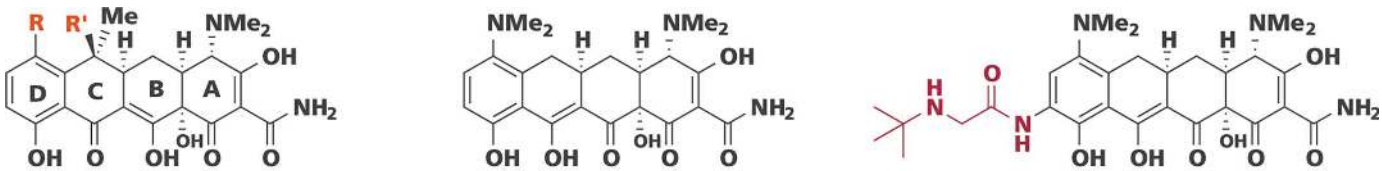


Figure 12.19 X-ray structure of streptomycin in the 30S ribosome subunit. (a) Schematic of specific binding interactions, including ionic interactions and hydrogen bonds, between streptomycin and the 30S ribosomal subunit, which is composed of protein and 16S rRNA. Lys45 of ribosomal protein S12 (brown) forms an ionic interaction with the phosphate of A913 (adenine 913) on helix 27 of rRNA (yellow), and also hydrogen bonds to hydroxyl groups of the aminocyclitol ring of streptomycin, stabilizing the error-prone form. A913 interacts with the aldehyde carbonyl group of streptomycin; this aldehyde group also binds to a phosphate between C526 and G527. Other nucleotides that participate in binding include C1490 and G1491 of helix 44 (blue), A914 of helix 28 (purple), and G527 and C526 of helix 18 in rRNA (teal). (b) Stereoview of streptomycin bound in the ribosome, showing helix 27, known as the accuracy helix (depicted in yellow), and the S12 ribosomal protein component (depicted in brown); other helices colored as described in (a).

Source: From Carter, A. P.; Clemons, W. M.; Brodersen, D. E., et al. (2000) *Nature* 407:340–348. With permission from Macmillan Publishers Ltd.

Tetracyclines, characterized by four fused six-membered rings, were discovered in 1948 when chlortetracycline was isolated from the soil bacteria *Streptomyces aureofaciens* (**Figure 12.20**). They are exemplified by oral activity as well as a broad spectrum of activity against Gram-positive, Gram-negative, and some parasitic infections. As such, they are used to treat sinusitis, bronchitis, and urinary tract infections (UTIs) caused by *H. influenzae*, pneumococci, and *E. coli*, as well as gonococci and *Chlamydia*. This class is particularly useful as a treatment for *Borrelia burgdorferi*, the causative agent of Lyme disease. Tetracyclines are most likely complexed to magnesium ions and then transported into cells through porins. They bind to the 16S RNA of the 30S subunit of the bacterial ribosome and prevent tRNA from binding at the A site.

Structural requirements for activity include the linear fused tetracyclic ring system, the naturally occurring stereochemistry found in the A and B rings, including the *cis* ring fusion, and the extended keto–enol system of the B, C, and D rings. The first-generation tetracyclines (1962–1973) include the natural products chlortetracycline and tetracycline and the semisynthetic compound doxycycline, which vary only



chlortetracycline: R = Cl, R' = OH
tetracycline: R = H, R' = OH
doxycycline: R = R' = H

minocycline

tigecycline

Figure 12.20 Structures of tetracyclines. First-generation analogs are represented by chlortetracycline, tetracycline, and doxycycline and differ in the substituents on the C and D rings. Minocycline is a second-generation derivative, and semisynthetic tigecycline is a representative of a third-generation derivative. The glycylcycline substituent of tigecycline is shown in red and may contribute to its activity against resistant organisms.

in the substituents on the C and D rings. A broader spectrum of activity characterizes the second-generation (semisynthetic) tetracyclines (1973–1990), such as minocycline. With the emergence of resistance to tetracyclines, substitution on the D ring led to glycylcyclines (third-generation), such as tigecycline, also a semisynthetic analog (see **Figure 12.20**). The two major mechanisms for tetracycline resistance are ribosomal protection and drug efflux. Ribosomal protection proteins, such as Tet(M) and Tet(O), act by dislodging tetracyclines from the ribosome, and the drug efflux systems are membrane-associated proteins that pump tetracyclines from the cell. The glycylcyclines, including newer analogs eravacycline and omadacycline, evade both these resistance mechanisms.

Erythromycin was the active ingredient in a soil sample containing *Streptomyces erythraeus*, found in the Philippines by Abelardo Aguilar. The sample was sent to Eli Lilly, and erythromycin, originally named ilosone, was isolated, its structure was elucidated, and the drug was approved and marketed in 1952. The complex macrolide is polyhydroxylated with a desosamine sugar at C5 and a cladinose sugar unit at C3. It has potent activity against Gram-positive and some Gram-negative bacteria and is used to treat respiratory, skin, and soft tissue infections, as well as infections caused by *Mycoplasma* and *Legionella*. One liability is its instability in the acidic environment of the stomach. Under those conditions, a significant fraction of the equilibrium mixture is present as the inactive anhydroerythromycin shown in **Figure 12.21**. Formation is initiated when the C12 hydroxyl attacks the C9 ketone to form a 9,12-hemiacetal. This reversible, acid-catalyzed process leads to the formation of the 6,9-acetal anhydroerythromycin, limiting the bioavailability of active erythromycin.

One solution to this problem was the development of the C6 methoxy analog clarithromycin, which cannot undergo the ketalization reaction (**Figure 12.22**). Another semisynthetic analog, roxithromycin, contains the C6 methoxy group as well as an oxime (both shown in red) at the original C9 ketone. This modification maintains the activity of the parent but has enhanced stability. Azithromycin, also lacking the C9 ketone, was prepared from erythromycin by a Beckmann rearrangement of a C9 oxime at Pliva Pharmaceuticals in Croatia and was subsequently licensed to Pfizer. This analog replaces the ketone with an amine and expands the ring size of the macrocycle from 14 to 15. Azithromycin has improved potency, acid stability, and spectrum of activity over erythromycin, with a low incidence of side effects.

With the increased use of erythromycin, roxithromycin, and azithromycin has come a corresponding increase in the emergence of resistant organisms. For example, these drugs are frequently prescribed to treat common infections, such as strep throat and upper respiratory infections; however, there are now many *Streptococcus* strains that are resistant to azithromycin.

Macrolides bind to 23S ribosomal RNA of the 50S subunit of ribosomes and prevent translocation of tRNA from the A site to the P site. One mechanism of resistance is modification of the ribosome via mono- or dimethylation of an adenine on the 23S ribosomal RNA component of the 50S subunit, which decreases binding to the cladinose of erythromycins. The enzymes carrying out this methylation are known as erythromycin ribosome methylation (Erm) enzymes and are found in resistant organisms.

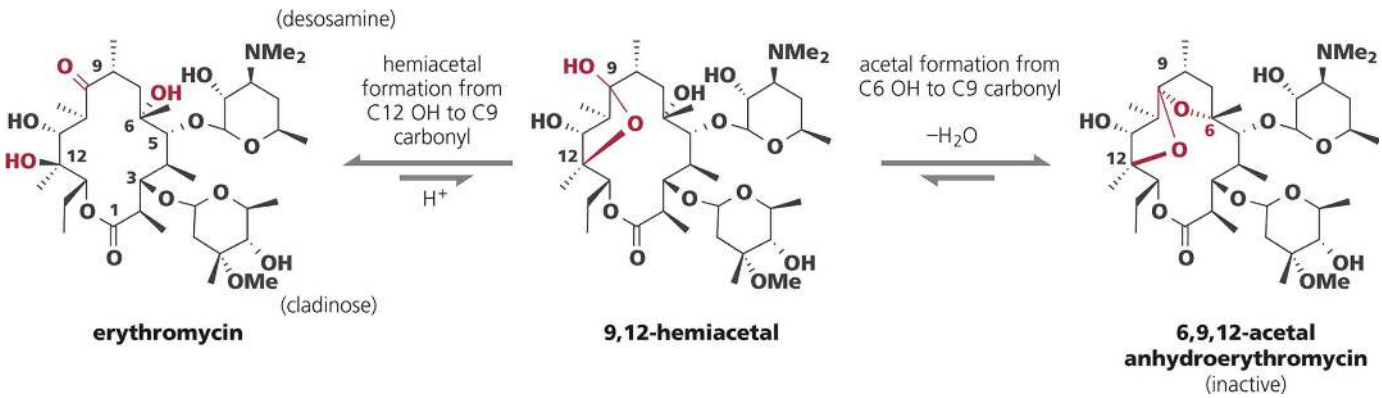
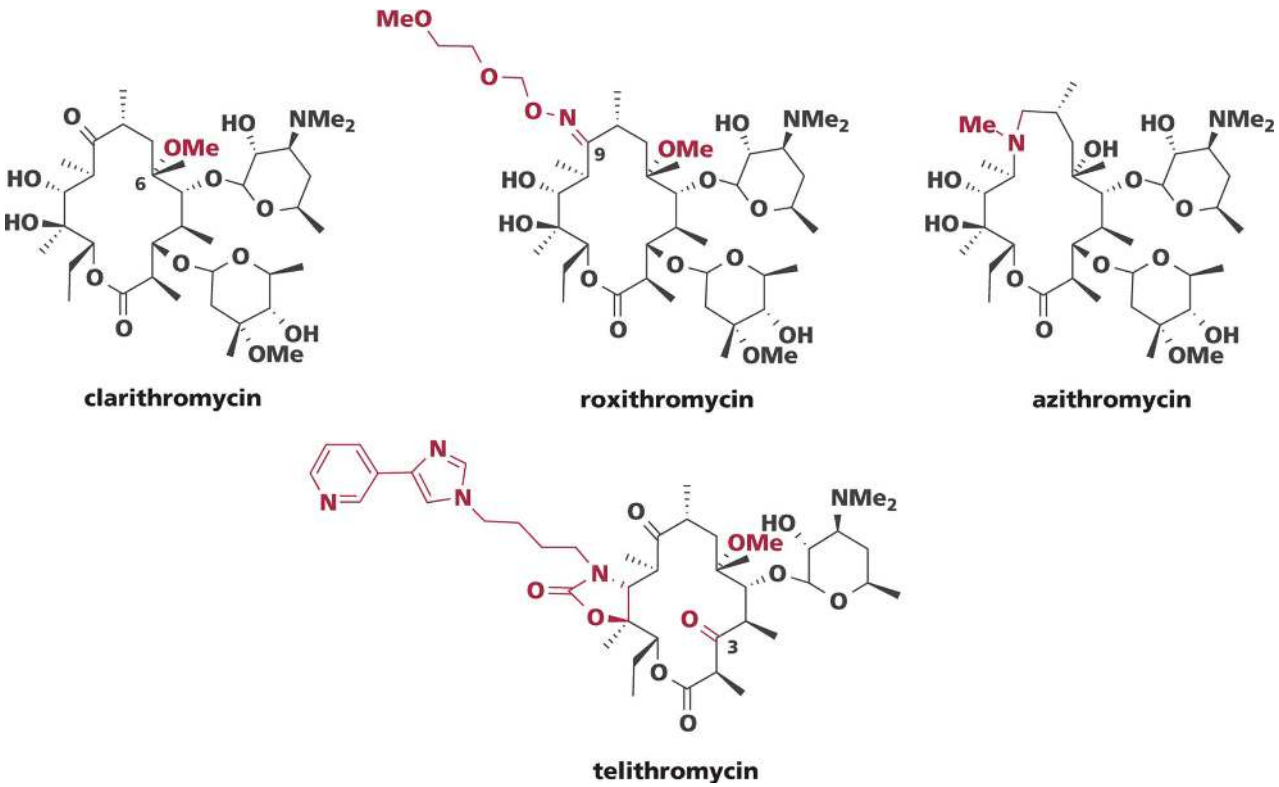


Figure 12.21 Erythromycin and its conversion to hemiacetal and acetal forms. Erythromycin is in equilibrium with its 9,12-hemiacetal; loss of H_2O occurs with the formation of the 6,9-acetal to generate anhydroerythromycin, which is inactive.



One approach to address this type of resistance is the use of Erm inhibitors, several of which have been identified, although none are used clinically. Another resistance mechanism that has been identified is the production of macrolide efflux proteins, which effectively reduces the amount of macrolide that is absorbed by the bacteria. Further medicinal chemistry optimization of this scaffold found that the addition of a cyclic carbamate improved activity against resistant organisms, and the replacement of the C3 cladinose ring with a ketone improved activity against efflux-mediated resistance. These modifications led to the ketolide telithromycin (Figure 12.22).

Figure 12.22 Structures of semisynthetic macrolide antibacterial drugs. Clarithromycin, roxithromycin, azithromycin, and telithromycin are shown, with key structural features that are modified compared to erythromycin highlighted in red. These structural changes result in improved stability and, in some cases, activity against erythromycin-resistant organisms.

The **pleuromutilins** (Figure 12.23) are antimicrobial compounds from *Pleurotus mutilus*, an edible mushroom. They were initially discovered in 1950 and were found to have modest potency against Gram-positive organisms. Two thioether analogs, tiamulin and valnemulin, containing extended side chains instead of the hydroxyl group, have been used in veterinary medicine since the late 1990s.

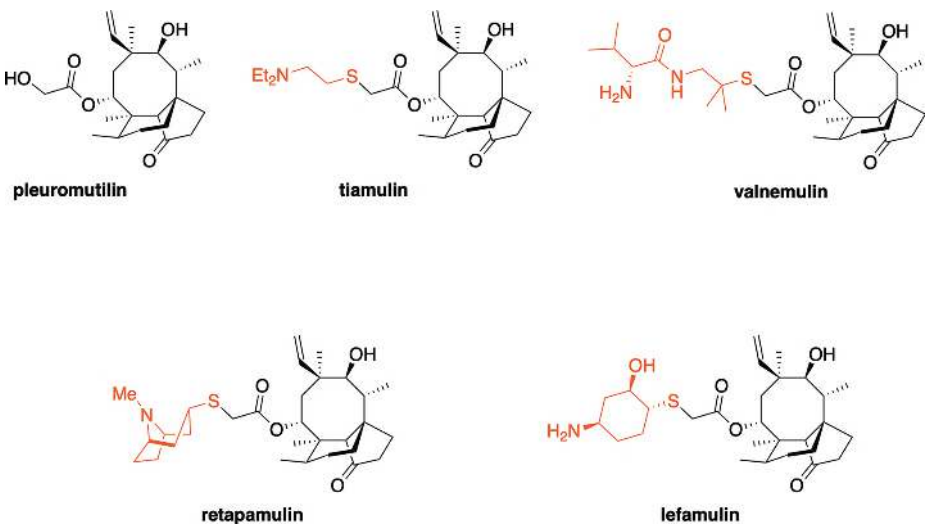


Figure 12.23 Structures of pleuromutilin and thioether analogs. Pleuromutilin is a natural product with modest antibacterial activity. Thioether modifications (shown in red) generated the drugs tiamulin, valnemulin, and retapamulin. Tiamulin and valnemulin are used in veterinary medicine, with retapamulin and lefamulin approved for use in humans. All inhibit protein synthesis by inhibiting the peptidyl transferase reaction.

The pleuromutilins inhibit protein synthesis by binding to nucleotides in the 23S site and inhibiting the peptidyl transferase reaction. While pleuromutilins and macrolides both bind to the 23S site, their specific binding sites and mechanisms of action are different. X-ray crystallography of tiamulin bound in the ribosome indicated that further variations of the side chain could be tolerated, such as in the rigid analog retapamulin, which was approved for human use in 2006 as a topical treatment for skin infections caused by MRSA. Lefamulin was approved in 2019 for bacteria pneumonia.

12.5.2 Oxazolidinones Are Inhibitors of Protein Synthesis

Oxazolidinones are a relatively new class of synthetic antibacterial drugs. The first member of the family, linezolid, was approved in 2000. As of 2017, only one additional oxazolidinone, tedizolid (formerly known as torezolid), is marketed (**Figure 12.24**). This class exhibits potent activity against Gram-positive organisms, particularly against MRSA and vancomycin-resistant enterococci (VRE), as well as activity against *M. tuberculosis*.

On the basis of an X-ray crystal co-structure (**Figure 12.25**), linezolid was shown to bind to the A site of the 50S subunit and is proposed to prevent tRNA from entering the site. Hydrophobic interactions are observed between linezolid's aromatic ring and the bases of C2452 and A2451 in 23S ribosomal RNA. Although, as a new class of drugs, resistance has been slow to develop, there were reports of linezolid-resistant organisms within 4 years of its introduction. The resistance mechanism appears to be mainly conferred by mutations to the 23S RNA.

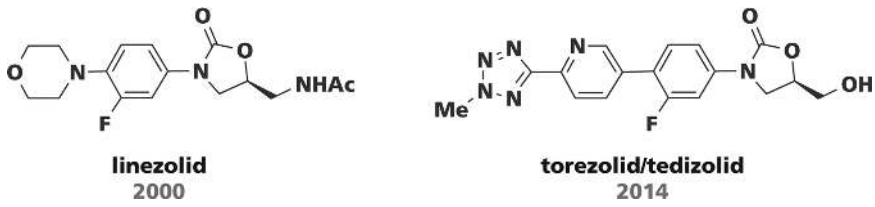


Figure 12.24 Structures of oxazolidinone antibacterial drugs. Linezolid and tedizolid (formerly known as torezolid) are shown.

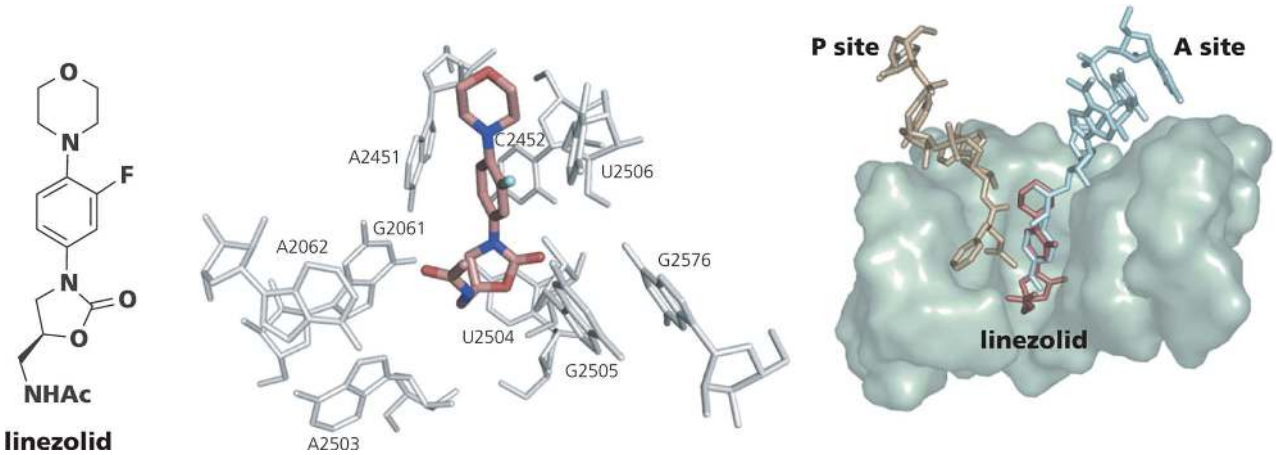


Figure 12.25 X-ray structure of linezolid bound in ribosome. *Left*, structure of linezolid. *Center*, representation of linezolid (depicted in pink) bound in ribosome. Specific interactions include hydrophobic interactions between the aromatic ring of linezolid and nucleobases of the ribosomal RNA. *Right*, surface features of ribosome with bound linezolid (depicted in pink). The A and P sites are defined by tRNA substrates shown in blue and tan, respectively. Linezolid occupies the same site as incoming tRNA, blocking its entry.

Source: Adapted from Leach, K. L.; Swaney, S. M.; Colca, J. R.; et al. (2007). *Mol Cell* 26:393–402. With permission from Elsevier.

12.6 DRUGS TARGETING BACTERIAL DNA REPLICATION

Two main categories of drugs, DNA gyrase inhibitors (quinolones) and dihydropteroate synthase inhibitors (sulfonamides), specifically prevent bacterial DNA replication and, ultimately, the ability of bacteria to reproduce and proliferate. Both classes inhibit bacterial enzymes that have human homologs. DNA gyrase is a topoisomerase involved in supercoiling and uncoiling of DNA during the bacterial replication process. The mechanism of topoisomerases was discussed in Chapter 10, and inhibitors of human topoisomerase are effective anti-cancer agents. Dihydropteroate synthase is an enzyme in the bacterial folic acid biosynthesis pathway. Folic acid is required for DNA synthesis, specifically as the source of the methyl group in thymidine, as was also described in Chapter 10. While humans do not efficiently synthesize folic acid, inhibitors of the related human dihydrofolate reductase, required for biosynthesis of purines and pyrimidines in the DNA replication process, are also used as anti-cancer agents.

12.6.1 Quinolone Antibacterial Agents Inhibit DNA Gyrase

Quinolone antibacterial agents, characterized by the scaffold in **Figure 12.26**, are synthetic compounds that were discovered in 1962. Nalidixic acid is a naphthyridone that was identified by George Lesher at Sterling Pharmaceuticals. This by-product, formed during the synthesis of the antimalarial drug chloroquine, was found to have potent activity against Gram-negative organisms and was the starting point for the development of the more useful class of fluoroquinolones. Since the original discovery of nalidixic acid, thousands of quinolones have been reported in the literature, and over 20 are used clinically. Requirements for activity include the quinolone or naphthyridone core, a carboxylic acid or equivalent at position 3, and an aliphatic or aromatic substituent on the N-1 nitrogen. Early calculations indicated that either an ethyl group or the constrained cyclopropyl moiety was the optimal size for the N-substituent, although it was later found that aryl substituents are also tolerated. The addition of a nitrogen heterocycle at C7 results in increased potency against Gram-positive bacteria and improves the compound's half-life. A fluorine atom at the C6 position, which characterizes the fluoroquinolines, improves potency toward DNA gyrase inhibition as well as cell penetration. More recent quinolones include those containing a substituent at C8 (chlorine, fluorine, or methoxy), which improves the overall spectrum of activity. A summary of the quinolone structure–activity relationships is shown in **Figure 12.26**.

Quinolones are categorized by generation, based primarily on the spectrum of activity. Nalidixic acid, oxolinic acid, and cinoxacin are examples of first-generation compounds with activity against Gram-positive bacteria (**Figure 12.27**). The 1980s saw the advent of the fluoroquinolones (second generation), which typically contain a cyclic amine at C7. Examples include ciprofloxacin, ofloxacin, levofloxacin, and norfloxacin. These analogs maintain Gram-positive activity and add activity against Gram-negative bacteria. The third-generation agents, closely related in structure to the second-generation fluoroquinolones, have an expanded spectrum of activity against Gram-positive bacteria, especially pneumococci. Examples include sparfloxacin and temafloxacin. The most recent drugs, the fourth generation, have additional activity against anaerobic bacteria and pneumococci. This generation includes compounds such as moxifloxacin, gemifloxacin, and trovafloxacin. The addition of larger, more structurally complex substituents at C7 both decreases susceptibility to efflux and increases activity against certain bacteria. The importance of this class of antibacterial drugs is evidenced by the fact that eight of the nine synthetic antibacterial agents introduced between 2000 and 2009 were fluoroquinolones. However, this class is also

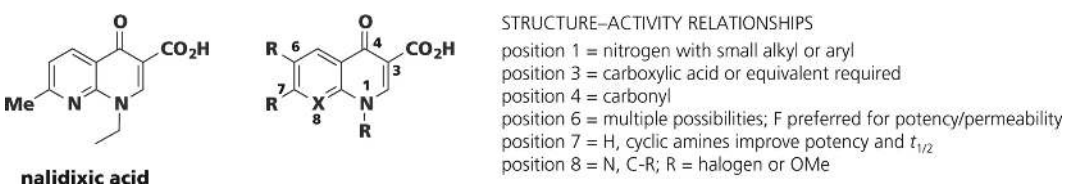


Figure 12.26 Structure of nalidixic acid and structure–activity relationships of quinolone antibacterials. Nalidixic acid was an early example of a DNA gyrase inhibitor and served as the basis for the class of quinolone antibiotics. Active compounds contain a quinolone or naphthyridone core and an acid at C3. Small alkyl or aryl substituents are preferred at N1. A variety of substitutions at C6 are possible, but F is preferred (fluoroquinolones). Cyclic amines at C7 improve potency and duration. C8 may be substituted with halogen or methoxy, which improves the spectrum of activity.

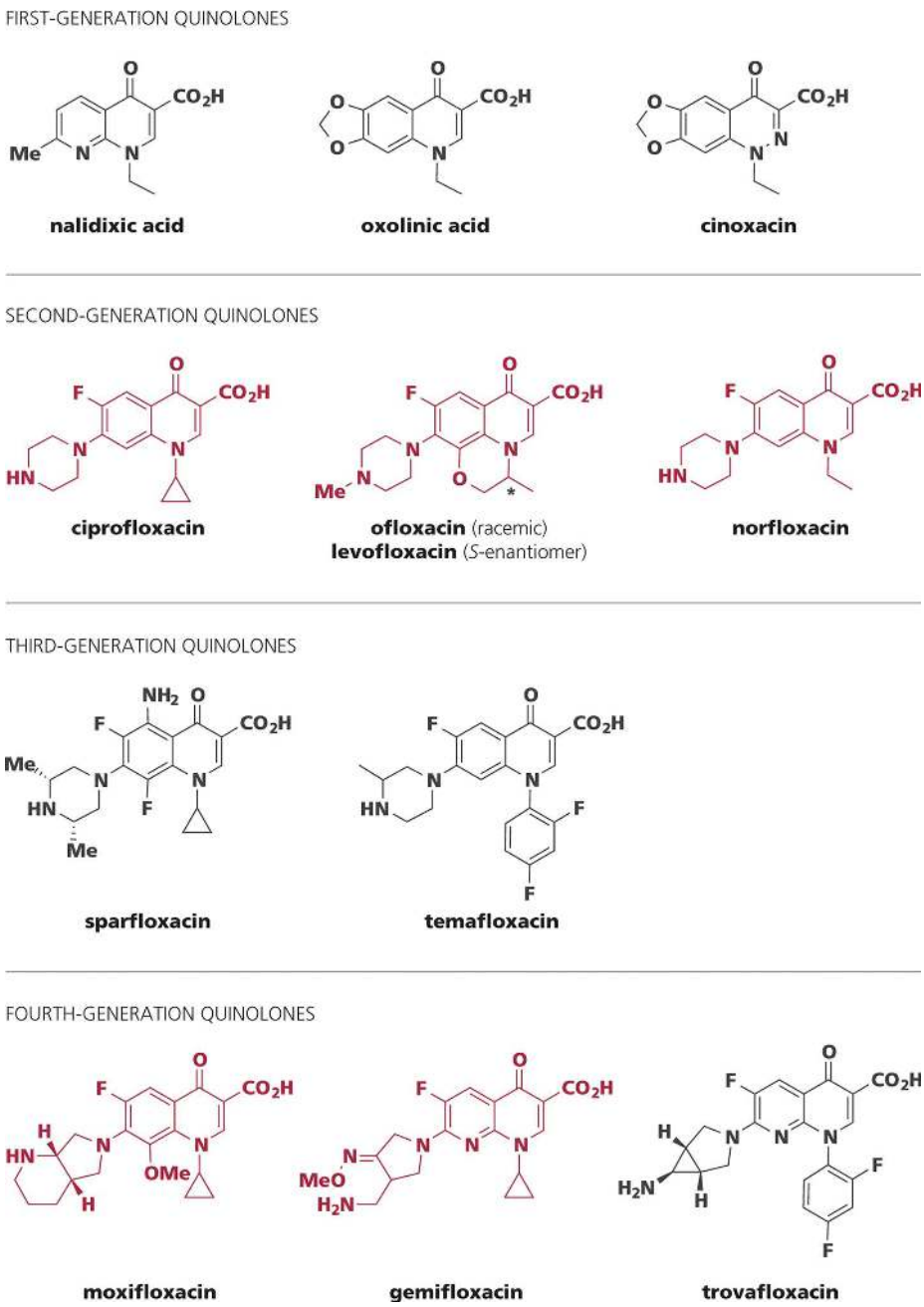


Figure 12.27 Examples of quinolone antibacterial drugs according to generation. First-generation quinolones include nalidixic acid (technically a naphthyridone) as well as oxolinic acid and cinoxacin, none of which are still used. Second-generation drugs are all fluoroquinolones, with a fluorine atom at C6 and cyclic amine at C7, which imparts improved potency and permeability. Ciprofloxacin, ofloxacin, levofloxacin, and norfloxacin are still used clinically. Third-generation compounds have branching on the C7 heterocycles and have improved Gram-positive activity; examples include sparfloxacin and temafloxacin, although both have been withdrawn. Examples of fourth-generation compounds include moxifloxacin, gemifloxacin, and trovafloxcin (withdrawn). These have bicyclic or complex C7 heterocycles and are active against anaerobes. Those remaining on the market in the United States as of 2017 are shown in red.

plagued by safety issues due to musculoskeletal, cardiovascular, and CNS side effects, leading to the removal of several of them from the market.

The quinolone antibacterial drugs are inhibitors of the bacterial enzyme DNA gyrase, or bacterial topoisomerase II. In Gram-positive bacteria, topoisomerase IV is also a target, and it is the primary target of fourth-generation quinolones. DNA gyrase and topoiosmerase IV catalyze the cleavage of double-stranded DNA and introduce negative superhelical turns. The enzyme is a tetrameric (A_2B_2) 400 kDa protein. Subunit A is responsible for the cleavage and resealing of the phosphate backbone, and Tyr124 of this subunit makes a covalent bond to DNA. The ATPase activity, which resides in the B subunit, hydrolyzes ATP to produce energy for supercoiling. Quinolones do not bind directly to DNA but to the DNA–gyrase complex to form a ternary structure. The mechanism is similar to that described for human topoisomerases and their inhibitors found in Chapter 10 (see **Figure 10.6**). Two molecules of the quinolone bind to the DNA–enzyme complex by intercalation into DNA, increasing the concentration of the ternary cleavage complexes and acting as poisons.

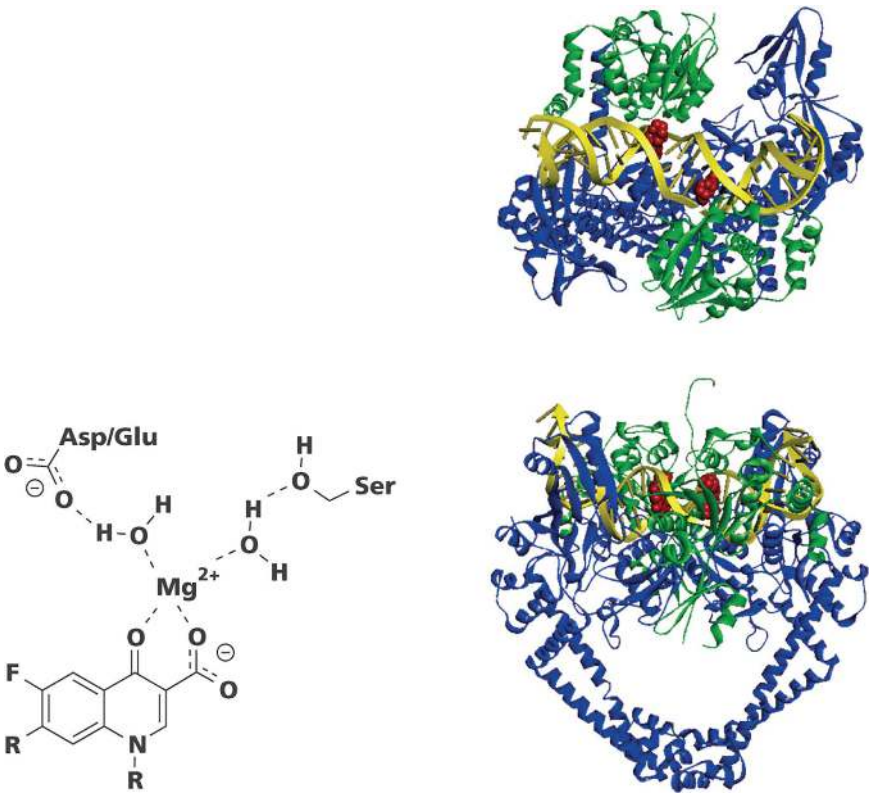


Figure 12.28 Representation of quinolone binding to topoisomerase IV. *Left*, representation of Mg^{2+} complex with a quinolone, with binding interactions to serine (Ser) and the side chain of an acidic residue (Glu or Asp) mediated through water molecules. *Right*, X-ray structure of two molecules of moxifloxacin (red) bound to the DNA (yellow)–topoisomerase IV (green/blue) complex. Two views are shown.

Source: Adapted from Aldred, K. J.; Kerns, R. J.; and Osheroff, N. (2014). *Biochemistry* 53:1, 565–1,574. With permission from the American Chemical Society.

Recent X-ray crystallography has provided detailed insight into the binding and mechanism of action of the quinolones. For example, it had long been proposed that there was a role for Mg^{2+} in quinolone binding, and this was confirmed in 2010. The carbonyl at C4 and the acid at C3 form a complex with a magnesium ion, which forms a bridge to serine and acidic (Asp, Glu) residues on the enzyme, mediated by water molecules (**Figure 12.28**). These residues are absent in human topoisomerase, providing some explanation for the selectivity of quinolones for bacterial DNA gyrase and topoisomerase IV. The planar aromatic quinolone also intercalates into DNA, distorting the helix. An X-ray structure of bound moxifloxacin shows two molecules intercalating into the DNA–enzyme complex, four base pairs apart (see **Figure 12.28**).

Mechanisms of resistance to quinolones include modification of the target, particularly mutations of the serine and acidic residues that bind to the magnesium complex. Resistance can also be conferred through plasmids, encoding both efflux pumps and proteins that metabolize or acetylate heterocyclic amines present at C7 in second-generation drugs and beyond, thereby reducing its activity. Porin mutations in Gram-negative bacteria are also an important resistance mechanism, resulting in both lower expression and decreased ability of quinolones to permeate bacteria.

12.6.2 Sulfonamide Antibacterial Compounds Inhibit Folic Acid Biosynthesis

The sulfonamides were the first nontoxic synthetic antibacterial drugs. The prototype, the prodrug prontosil, a red dye, was marketed in 1936. Gerhard Domagk and Josef Klarer, at Bayer in Germany, investigated a series of dyes as potential magic bullets that might target bacteria, since they were known to specifically dye bacteria. After a lack of toxicity was demonstrated in mice infected with *Streptococcus*, trials in humans showed both efficacy and safety. Most famously, in 1936, while still a little-known drug, prontosil was used to treat and cure Franklin Roosevelt's son, who was seriously ill from complications of strep throat. The resulting publicity helped market the drug. Although prontosil is active *in vivo*, it is inactive *in vitro* and is converted by the enzyme azoreductase to sulfanilamide, the active form (**Figure 12.29**).

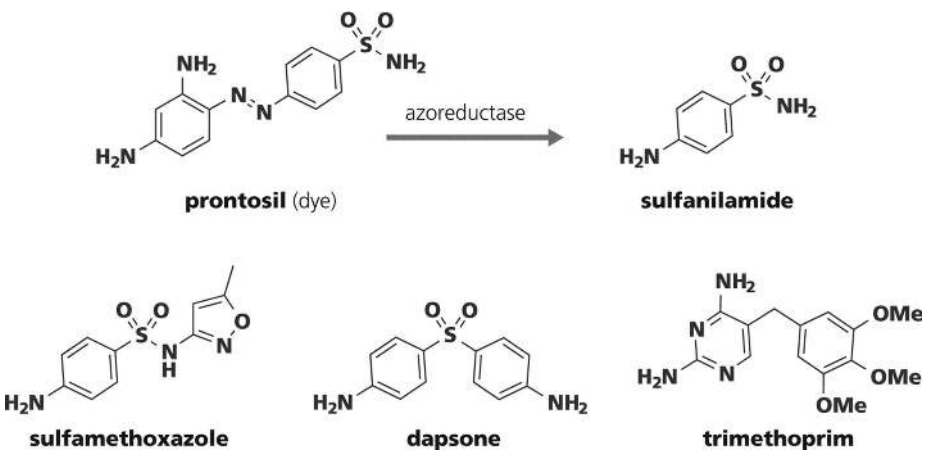


Figure 12.29 Inhibitors of folate biosynthesis. Prontosil is a prodrug that generates sulfanilamide *in vivo*, which inhibits dihydropteroate synthase. Sulfamethoxazole, the only sulfonamide antibacterial drug still in use, and dapsone also inhibit dihydropteroate synthase. Trimethoprim inhibits a different enzyme in the folic acid biosynthesis pathway, dihydrofolate reductase.

Since the active component, sulfanilamide, was a known compound used in dye synthesis, numerous preparations were made and sold. Due to a lack of regulations and oversight at that time, one of these preparations, in sweet-tasting ethylene glycol, led to the deaths of over 100 patients (see Chapter 5). This tragedy resulted in the passage of the Federal Food, Drug, and Cosmetic Act in 1938, and prontosil is no longer used.

The basic structure of all antibacterial sulfonamides includes a *p*-aminosulfonamide moiety, where the sulfonamide nitrogen may be substituted with a variety of groups. Allergic reactions were common, and the sulfa drugs were widely replaced by penicillins in the 1940s. Sulfamethoxazole (see Figure 12.29) is the one example still in use.

The sulfonamide antibacterial drugs are inhibitors of dihydropteroate synthase, an enzyme required for the synthesis of tetrahydrofolic acid, which is a coenzyme required for DNA synthesis in bacteria. In a reaction catalyzed by dihydropteroate synthase, *p*-aminobenzoic acid (PABA) reacts with dihydropteroate diphosphate in a displacement reaction to give dihydropteroic acid. This product is the substrate for a series of additional steps that include the installation of the amide bond (shown in red) found in tetrahydrofolic acid, as shown in Figure 12.30a. The sulfonamide antibacterial drugs are substrates for

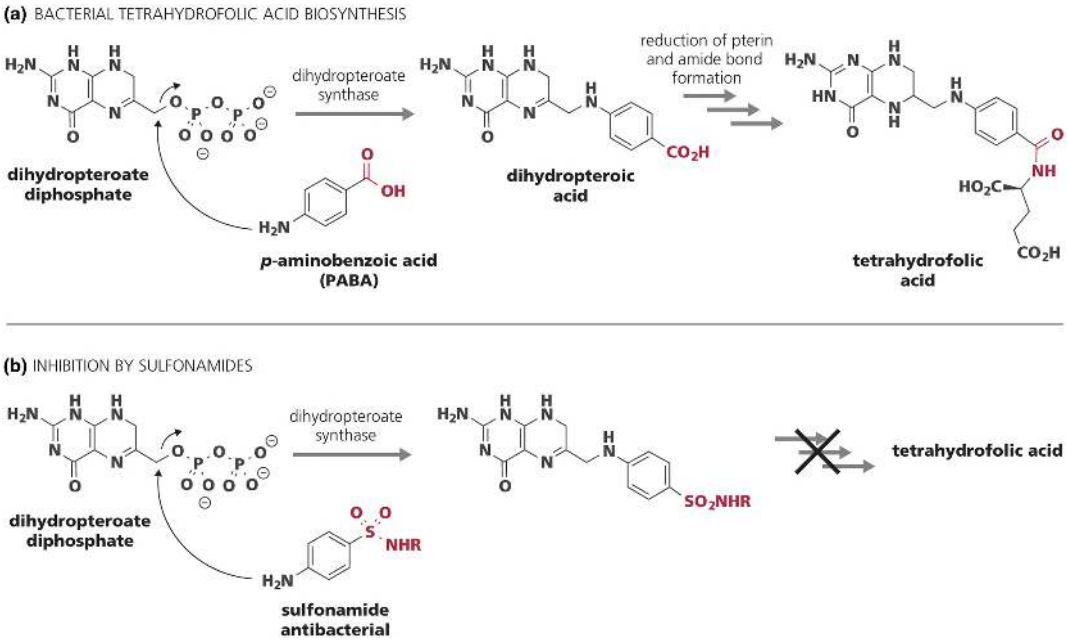


Figure 12.30 Mechanism of action of sulfonamide antibacterial drugs. (a) Generalized biosynthesis of tetrahydrofolic acid, an intermediate in the biosynthesis of folic acid. The amine of *p*-aminobenzoic acid displaces the diphosphate of dihydropteroate diphosphate, leading to dihydropteroic acid; subsequent reactions include conversion of the acid (shown in red) to the amide in tetrahydrofolic acid. Tetrahydrofolic acid eventually is converted to folic acid. (b) Mechanism of inhibition of sulfonamide. In an analogous reaction, the amine of a sulfonamide adds to dihydropteroate diphosphate, but the sulfonamide product cannot participate in the subsequent reactions for the final conversion to tetrahydrofolic acid, leading to the inhibition of folic acid synthesis.

dihydropteroate synthase and compete with the natural substrate, displacing the diphosphate to form the sulfonamide analog of dihydropteroic acid; however, no further reactions proceed since the requisite acid moiety is not present (**Figure 12.30b**).

Dapsone (see **Figure 12.29**), although not a sulfonamide, is a sulfone that also acts as an inhibitor of dihydropteroate synthase by competing with PABA. This drug is particularly effective against mycobacterial infections and is used to treat both tuberculosis (*M. tuberculosis*) and leprosy (*M. leprae*). Another antibacterial that inhibits this pathway is trimethoprim, which inhibits the bacterial enzyme dihydrofolate reductase, blocking the conversion of dihydrofolic acid to tetrahydrofolic acid. It is used most frequently to treat UTIs as a formulation in combination with sulfamethoxazole.

12.7 DRUGS AFFECTING BACTERIAL CELL MEMBRANE PERMEABILITY

Drugs that affect bacterial cell membrane permeability include the polymyxins and daptomycin (**Figure 12.31**). Polyethers, such as monensin and lasalocid, also affect cell membrane permeability and are used as supplements in animal feed; they affect the flora and aid digestion. Polymyxins are cationic cyclic peptides with a long hydrophobic tail that binds to the lipopolysaccharide of Gram-negative bacteria. They are relatively toxic and were not widely used until recently, even though they have been available for years. With the increase in multi-drug-resistant bacteria such as *P. aeruginosa*, *A. baumannii*, and carbapenem-resistant Enterobacteriaceae, there has been renewed interest in polymyxins. Several polymyxins (labeled A–E) were isolated

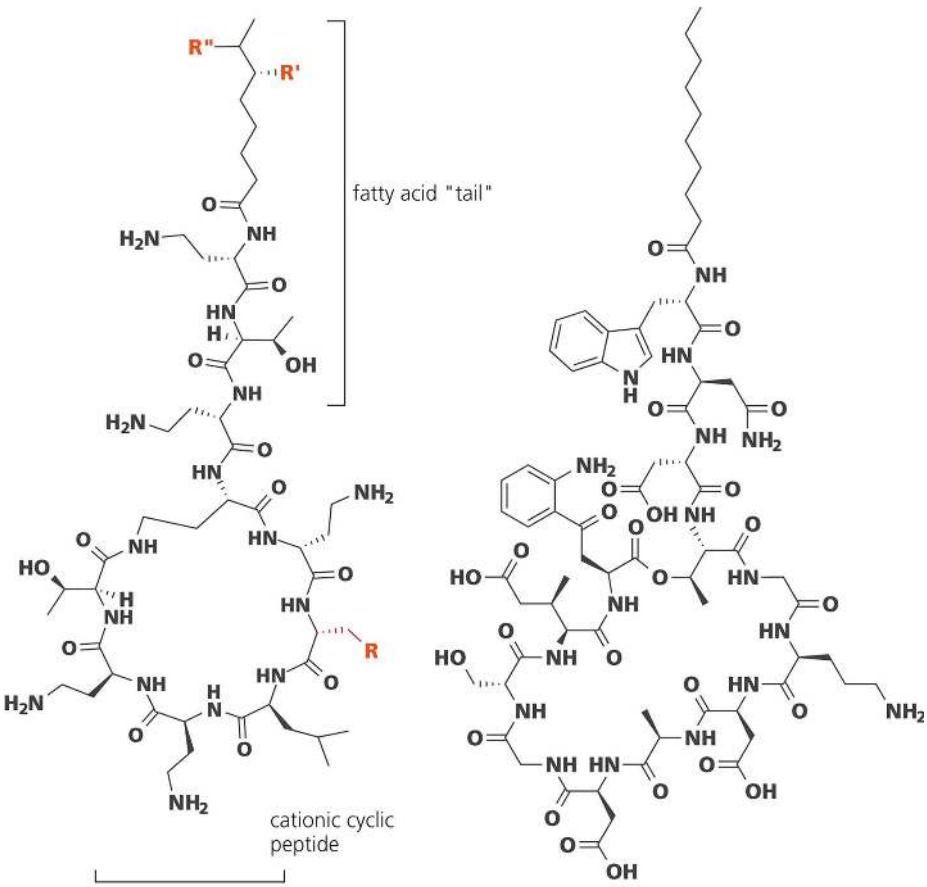


Figure 12.31 Structures of polymyxins and daptomycin. Both are cyclic peptides with lipophilic tails that disrupt bacterial cell membranes; points of variation in the polymyxins are shown in red. Polymyxin B is a mixture of polymyxins B1 and B2, differing in the substitution pattern on the fatty acid side chain. Polymyxin E (colistin) has the same fatty acid side chain as polymyxin B1 but substitutes D-Leu for D-Phe.

in the 1940s from the soil bacterium *Paenibacillus polymyxa*; both polymyxins B and E (also known as colistin) are used clinically. Polymyxin B is a mixture (of B1 and B2) that varies in the composition of the fatty acid side chain. Polymyxin E varies from polymyxin B1 by substitution of D-Leu for D-Phe (see **Figure 12.31**).

The cationic cyclic peptide displaces Ca^{2+} and Mg^{2+} from the membrane, destabilizing it. The fatty acid tail then interacts with LPS, inserting into the membrane and allowing the leakage of cellular components. Due to toxicity, primarily in the kidney, their use was very limited until recently, when they have become a drug of last resort to treat patients infected with multi-drug-resistant Gram-negative bacteria. This mechanism is similar to that proposed for daptomycin (described in detail in Section 8.4), which also interacts with the lipid bilayer of the cell membrane.

12.8 ANTIPARASITIC DRUGS: COMMON PARASITE FEATURES, INFECTIOUS PROCESS, AND IMPACT

A number of infectious diseases are caused not by bacteria but by parasitic organisms. These include both unicellular organisms, formerly called protozoa, and more complex worms (helminths). Protozoan parasites are microscopic, unicellular, eukaryotic organisms that are transmitted to humans through vectors, usually blood-feeding insects, such as mosquitoes. They proceed through several life stages in the infected host cell; some have capabilities for movement or propulsion. Protozoal infections can be relatively mild to life-threatening, with the severity depending on the particular strain of the infectious agent and the health and immune status of the host. Symptoms of protozoal infections can be the results of various effects. As with bacterial infections, host defenses against the infecting agent cause symptoms, and the infectious organism itself can release toxic products that damage cells and tissues. In some cases, the presence of the parasite in the small intestine interferes with the absorption of nutrients. Examples of diseases caused by unicellular organisms include malaria, African sleeping sickness, and Chagas disease.

Helminths are multicellular worms, which in some cases can be quite large (>1 m long). They are transmitted by various methods: by ingestion or penetration of eggs or larvae, through contact with vectors, or by ingestion of infected hosts. In addition to symptoms being the result of the host's own immune system, pathology can also be caused by the helminths' activity or metabolism or by physical blockage of physiological processes. Many helminths are extremely long-lived, and chronic infection can cause irreversible damage.

Parasites are the most common infectious organisms in developing countries in sub-Saharan Africa, Asia, and the Americas, where it is estimated that one-third of the population is infected. The organisms often infect the gastrointestinal tract and cause tropical diseases, such as hookworm, ascariasis, schistosomiasis (snail fever), lymphatic filariasis (elephantiasis), and onchocerciasis (river blindness).

Parasitic diseases affect billions of people in the world, and their incidence is increasing. In addition, certain infections, typically of little consequence, have become more prevalent and serious when they occur in immunosuppressed patients. Finally, the migration of individuals infected with parasites to low-prevalence areas has further contributed to the increase in incidence.

12.9 ANTIPARASITIC DRUG DISCOVERY

Many parasitic diseases are considered neglected diseases. They are prevalent in tropical and subtropical environments and affect billions of people, who often live in poverty, in close contact with infectious vectors, and without adequate sanitation. They are termed *neglected* because these diseases have been largely eradicated in more developed parts of the world. Treatment strategies rely on prevention measures, such as the elimination of vectors and safe food and water sources. In addition, drug therapy needs to consider the specific patients affected and their environment. As such, low cost, short treatment regimens, and ease of administration are primary criteria.

Many older drugs were developed on the basis of known activity against other organisms, relying on phenotypic screens. In fact, a number of antibacterial drugs are also used to treat parasitic diseases. However, the number of drugs that specifically target parasites is far fewer than the number that target bacteria. Challenges include the highly complex life cycle of some parasites and the development of resistance. While there were few drugs approved from the 1970s to the 1990s, there has been an increase in activity in the development of new antiparasitic agents in the 21st century. Advances in sequencing the genomes of these organisms and the formation of private–public partnerships to support drug discovery for neglected tropical diseases are partly responsible for this change. Table 12.2 highlights examples of antiparasitic agents.

Table 12.2 Examples of Pathogenic Parasites, Associated Diseases and Conditions, Drugs or Drug Classes Used to Treat Each, and Their Mechanism of Action

Organism	Disease or Condition	Drug	Mechanism
Protozoal Parasites			
Plasmodium falciparum and other *Plasmodium* species	Malaria	Quinine	Inhibition of hemozoin crystallization
		Artemisinin	ROS generation
		Mefloquine, chloroquine (prevention)	Inhibition of hemozoin crystallization
		Doxycycline (prevention)	Protein synthesis inhibition
Leishmania	Leishmaniasis	Pentamidine	Not clarified; interferes with nuclear metabolism
		Amphotericin B	Cell wall disruption
Trypanosoma brucei gambiense	African trypanosomiasis (sleeping sickness)	Suramin	Inhibition of glycolytic enzymes
		Pentamidine	Not clarified; interferes with nuclear metabolism
Trypanosoma cruzi	Chagas disease	Nifurtimox	ROS generation
		Benznidazole	ROS generation
Helminths			
Onchocerca volvulus	Onchocerciasis (river blindness)	Ivermectin	Increased permeability of invertebrate nerve/muscle cells
Schistosoma mansoni	Schistosomiasis (snail fever)	Praziquantel	Muscle paralysis due to Ca^{2+} influx
Necator americanus	Hookworm	Albendazole, mebendazole	Inhibition of microtubule polymerization
Ascaris lumbricoides	Roundworms (ascariasis)	Albendazole, mebendazole	Inhibition of microtubule polymerization
Strongyloides stercoralis	Strongyloidiasis	Mebendazole	Inhibition of microtubule polymerization
Wuchereria bancrofti	Elephantiasis (lymphatic filariasis)	Ivermectin	Increased permeability of invertebrate nerve/muscle cells

12.9.1 Current Drugs Used to Treat Malaria Are Derivatives of Natural Products

Malaria is a mosquito-borne disease that affects approximately 3 billion people worldwide, resulting in roughly 1 million deaths per year. The disease is caused by various organisms of the *Plasmodium* genus, which have become increasingly resistant to treatment. Humans are infected with the parasite in its sporozoite form by the mosquito vector. The sporozoite then travels to the liver and matures to a schizont, containing merozoites, which then infect other blood cells and produce more merozoites in the blood (erythrocytic) stage. There the parasites degrade hemoglobin, releasing iron-bound heme. However, heme is toxic to the parasite, so the parasites then convert the toxic heme to nontoxic hemozoin crystals (also called "malaria pigment") (Figure 12.32).

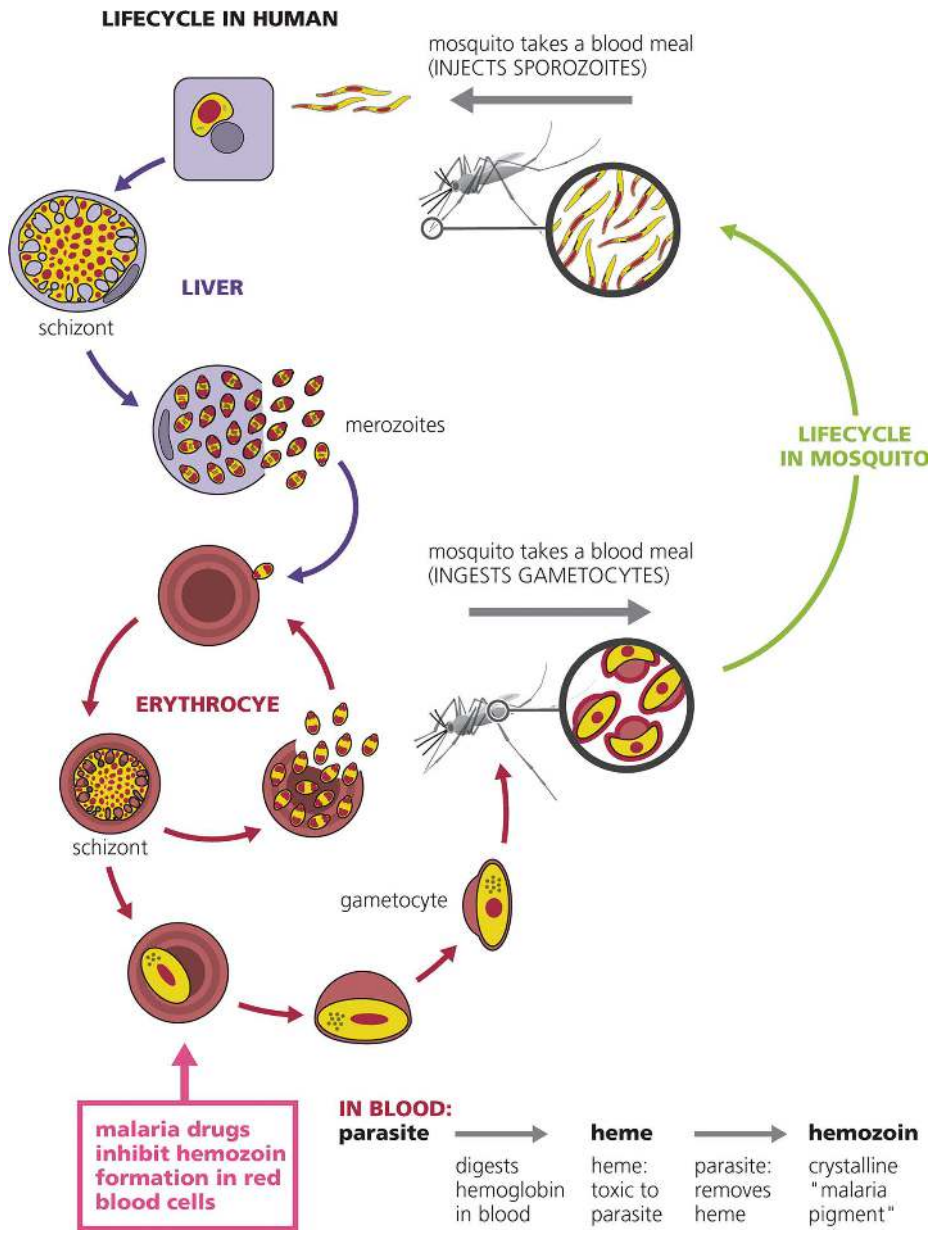


Figure 12.32 Life cycle of malaria parasite. The infected *Anopheles* mosquito injects the parasite as a sporozoite into the blood; the sporozoite enters the liver and matures to a schizont, which contains thousands of merozoites. The release of merozoites into the blood leads to the destruction of blood cells. In this blood stage, the parasite digests hemoglobin, forming heme, which is toxic to the parasite; the parasite then removes heme, forming hemozoin. Malaria drugs inhibit hemozoin formation, leading to the toxic buildup of heme.

Despite renewed efforts to develop antimalarial drugs, there are only two classes in use. Both inhibit hemozoin formation, albeit via different mechanisms. The earliest drug used to treat malaria was quinine, an alkaloid isolated from the bark of the *Cinchona* tree and introduced into Western medicine in the 17th century. More recent analogs of quinine include chloroquine, introduced in the 1950s; mefloquine, developed in the 1970s; and tafenoquine, approved in 2018 (**Figure 12.33**). The mechanism of the quinine-based antimalarial drugs involves inhibition of hemozoin crystallization, resulting in the buildup of heme, which then kills the parasites.

Although the early quinine-based antimalarial drugs have been effective, multi-drug-resistant organisms are a serious threat. To combat this issue, in 2006, the WHO advocated the use of artemisinin-based therapies, which are now used in many countries. Artemisinin was discovered in the 1970s as a result of Project 523, an effort by the Chinese government to find new antimalarial drugs during the Vietnam War (see Case Study 1 in Chapter 1). It contains an unusual peroxide moiety and has poor solubility. The reduced semisynthetic analog artemether is most widely used, as well as the succinic ester artesunate. Unlike the quinine derivatives, which act in the blood stage, artemisinin and artemether are active at all stages of the parasite life cycle. The mechanism of action is still not fully defined but involves bioactivation by heme iron

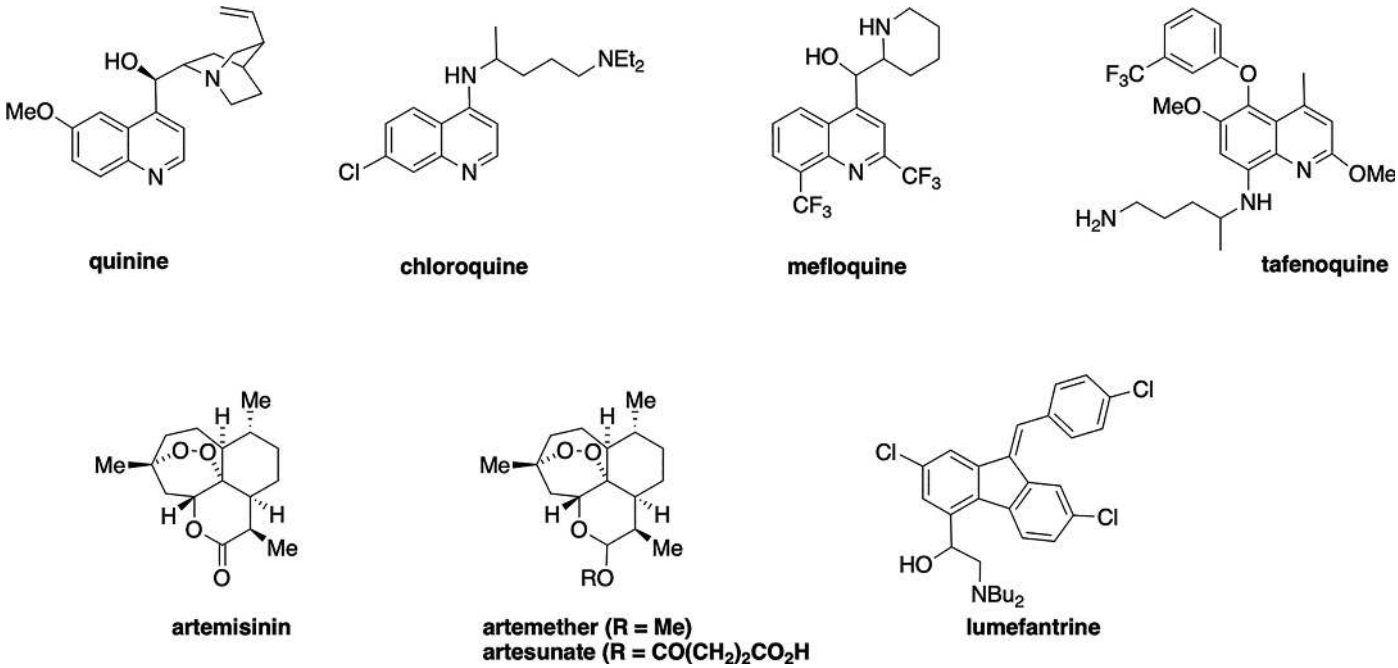


Figure 12.33 Structures of quinine- and artemisinin-based antimalarial drugs. The natural product quinine has been in use since the 17th century to treat malaria. More recent synthetic analogs include chloroquine, mefloquine, and tafenoquine, all of which are administered as racemic mixtures. The natural product artemisinin was discovered in the 1970s, and its semisynthetic analogs artemether and artesunate are widely used, in combination with lumefantrine.

to generate reactive oxygen species (ROS). The free radicals inhibit hemozoin formation and also damage cellular targets in the parasite. Artemether has a short half-life and is usually given in combination with a longer-acting drug, such as lumefantrine. Lumefantrine was also discovered in China, as part of Project 523, and interferes with the conversion of heme to hemozoin. The Nobel Prize in Physiology or Medicine 2015 was awarded to Youyou Tu for her work on the isolation of artemisinin.

In addition to treatment of malaria, several drugs are used prophylactically, so that they will be present in the body if a person is infected with the parasite by an insect bite. One commonly used prophylactic drug is Malarone, a combination of atovaquone and proguanil (**Figure 12.34a**). As part of the Malaria Research Committee, hundreds of hydroxynaphthoquinones, structurally related to atovaquone, were synthesized in the 1940s in an effort to find quinine alternatives due to shortages of it during World War II. Although many of these compounds were found to inhibit respiration of the organism, they were not metabolically stable and therefore not an effective quinine alternative.

These hydroxynaphthoquinones were re-investigated years later when, in the late 1960s, it was found that malarial ubiquinones (coenzymes present in eukaryotes) differed from their mammalian counterparts at the polyprenyl chain. The ubiquinones (see **Figure 12.34a**) are involved in electron transfer reactions required for cell respiration, and therefore, it was suggested that ubiquinone analogs might selectively inhibit electron transfer in the malaria parasite and be an effective therapeutic. In the 1990s, the investigation of atovaquone showed that it irreversibly blocks ubiquinol–ubiquinone redox cycling.

The ubiquinol–ubiquinone process is involved in pyrimidine biosynthesis, and blockage leads to death of the organism. Although as a single agent, atovaquone had a 70% cure rate in infected patients, it was found to act synergistically with another drug, proguanil, which had been discovered in the 1940s; when used in combination, cure rates are nearly 95%. Proguanil is a dihydrofolate reductase (DHFR) inhibitor, with selectivity for the parasite form of the enzyme.

In 1999, the Medicines for Malaria Venture (MMV), a public–private foundation, was launched to develop novel medicines to treat malaria. Several compounds from this initiative progressed to Phase II clinical trials, including the synthetic ozonide artefenomel, a peroxide that works by the same mechanism as artemisinin, and the

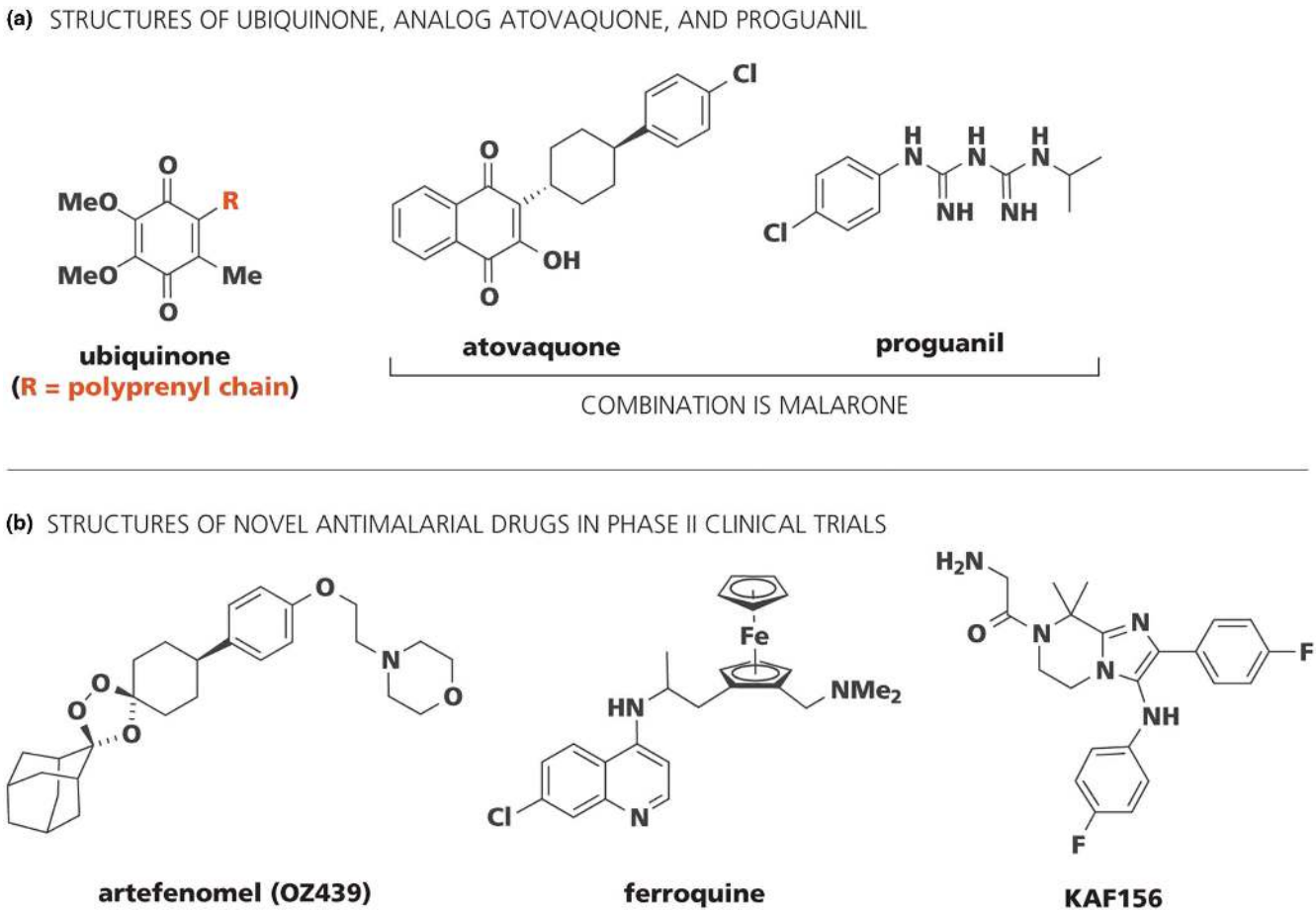


Figure 12.34 Structures of quinone-based antimalarial prophylactic drugs and novel drugs in clinical trials. (a) The malaria parasite relies on a unique ubiquinone for respiration. The prophylatic antimalarial drug atovaquone is an analog of ubiquinone and is used in combination with the DHFR inhibitor proguanil as Malarone. (b) Structures of several novel anti-malarial drugs that have advanced to clinical trials that work through various mechanisms.

quinine-based organometallic drug ferroquine (Figure 12.34b). Although artefenomel was promising through Phase IIb trials, it was discontinued due to formulation and solubility problems. In addition, Novartis has developed KAF156 (ganaplacide), a novel imidazolopiperazine that was optimized from a hit identified from a high-throughput screening campaign of 2 million compounds to detect compounds that inhibited proliferation of *Plasmodium falciparum* (see Journal Club 1 in Chapter 4). KAF156 has advanced to Phase IIb clinical trials, in combination with lumefantrine.

12.9.2 There Are Few Effective Treatments for Trypanosomal Diseases

Trypanosomes are a type of protozoal parasite transmitted by insect vectors, in a similar fashion to malaria. *Trypanosoma brucei gambiense* is transmitted by the tsetse fly and is seen mainly in sub-Saharan Africa, where it causes African sleeping sickness, or human African trypanosomiasis (HAT). A similar organism, *Trypanosoma cruzi*, causes American trypanosomiasis, or Chagas disease, and is transmitted by the kissing bug. The disease is found throughout South America and Central America and in the Southern United States.

In response to an epidemic of HAT in the Congo and Kenya in the late 1800s, a cure was sought by colonial governments. Ehrlich investigated a number of dyes, which led to the identification of a series of compounds such as trypan blue, which was effective in mice but not in other mammals, and the closely related trypan red, which was effective in humans but dyed the skin red. Inclusion of a urea linking the structural elements eventually led to the dimeric compound suramin, which did not have trypan red's pigmenting side effects. Suramin was developed at Bayer in Germany in 1916, where it was originally named Germanin. Its value was so high that it was actually offered to the British in return for some German colonies that had been lost in World War I, without disclosing the structure. The British declined, and the

structure was determined independently by the French pharmacist Ernest Fourneau. Suramin is still used to treat the early stages of HAT (**Figure 12.35a**).

Another class of compounds investigated during the early 1900s was the arsenicals. Although toxic, particularly to the optic nerve, tryparsamide was used until the 1960s, when it was replaced by the less-toxic analog melarsoprol, which is still used to treat the later stages of HAT. Pentamidine, invented in 1940, is an analog of the early anti-diabetic drug synthalin, or decane-1,10-bis(guanidine), and was found to have anti-trypanosomal activity and has been used to treat HAT and Chagas disease. Its mechanism has not been clarified but involves interfering with nuclear metabolism (see **Figure 12.35a**).

The connection between anti-diabetics and drugs to treat trypanosome infection came from the observation that trypanosomes need glucose to reproduce. Synthalin, discovered in the 1920s, was based on a folk remedy, the French lilac *Galega officinalis*, used in medieval Europe for the diabetic condition. This plant contains both guanidine and *N*-isoprenylguanidine, and alkyl analogs such as synthalin were investigated in the 1920s for diabetes.

Eflornithine (**Figure 12.35b**) was designed as a suicide inhibitor of ornithine decarboxylase, where inhibition leads to reduced cell proliferation. It was ineffective as an anti-tumor agent, but testing in the 1980s showed it to be an effective inhibitor of

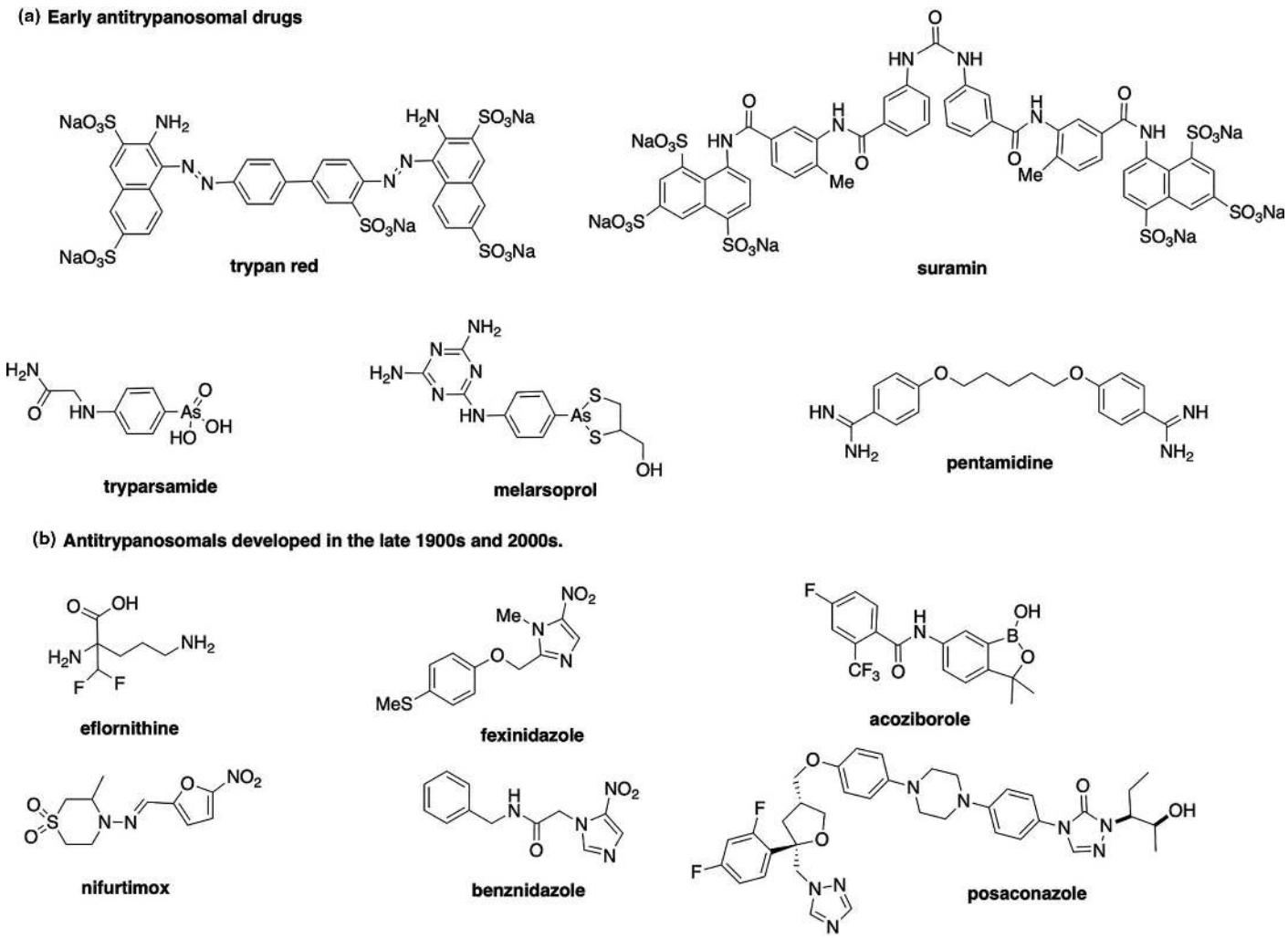


Figure 12.35 Structures of anti-trypanosomal drugs. (a) Early anti-trypanosomal drugs. Structure activity relationship studies of trypan red served as the basis for suramin, which is used to treat HAT. Other early drugs include the arsenicals tryparsamide and melarsoprol (still used to treat later stages of HAT), as well as pentamidine. (b) Later anti-trypanosomal drugs. Eflornithine inhibits ornithine decarboxylase and is used to treat HAT. Fexinidazole, an analog of benznidazole, was approved in 2021 to treat HAT. The nitrofuran nifurtimox and the nitroimidazole benznidazole likely work through ROS generation and are effective in Chagas disease. Acoziborole is in late clinical trials and inhibits pre-mRNA cleavage. Posaconazole, an anti-fungal agent, inhibits Cyp51 and is effective in animal models as a treatment for Chagas disease.

the proliferation of *T. brucei gambiense*. It was approved in 1990 to treat African sleeping sickness. Fexinidazole, approved to treat HAT in 2021, was first synthesized in the 1980s and is closely related to benznidazole. It was rediscovered due to the efforts of the public–private partnership Drugs for Neglected Diseases Initiative (DNDi) and is on the WHO's List of Essential Medicines.

Like the MMV for malaria drugs, DNDi was launched in 2003, with the goal of identifying novel accessible treatments for HAT, leishmaniasis, Chagas disease, pediatric HIV, and helminth infections. In 2010, researchers, with support from DNDi, as well as others, including the Bill and Melinda Gates Foundation, reported the identification and optimization of benzoxaboroles that cross the blood–brain barrier and cured HAT infections in preclinical models after a single dose. Acoziborole (**Figure 12.35b**) has shown very promising results in advanced clinical trials; its mechanism of action appears to be through inhibition of the RNA cleavage and polyadenylation specificity factor subunit 3 (CPSF3), a member of the metallo-beta-lactamase family of proteins that cleaves pre-mRNA.

Drugs used to treat Chagas disease include nifurtimox and benznidazole. The mechanism of these aromatic nitro compounds is believed to involve the reduction of the nitro group to afford reactive oxygen species (ROS), which then cause cell damage. The increased interest in addressing neglected diseases has led to several compounds that have entered clinical trials. Antifungal compounds such as posaconazole (**Figure 12.35b**) and others from this class are effective in models of Chagas disease; they act by CYP51 (sterol 14α-demethylase) inhibition. Since *T. cruzi* requires ergosterol, it is susceptible to these inhibitors, more commonly used as antifungal drugs (see Chapter 11).

12.9.3 Drugs That Treat Helminthic Diseases Are Used in Both Veterinary and Human Medicine

Although diseases caused by helminths (worms) have existed for thousands of years and affect many millions of people, there are very few drugs to treat these conditions. Like protozoal parasites, helminths have complex life cycles and can grow in both mammals and invertebrates. Only four new drugs were introduced between 1975 and 2004; however, their effectiveness has nearly eradicated some conditions.

The azole drugs, such as mebendazole and albendazole, were discovered in the 1970s. These carbamates are less-toxic analogs of thiabendazole, first reported in 1961 as a veterinary medicine. Mebendazole and albendazole bind selectively to parasitic β-tubulin, blocking microtubule formation in the parasite (**Figure 12.36**).

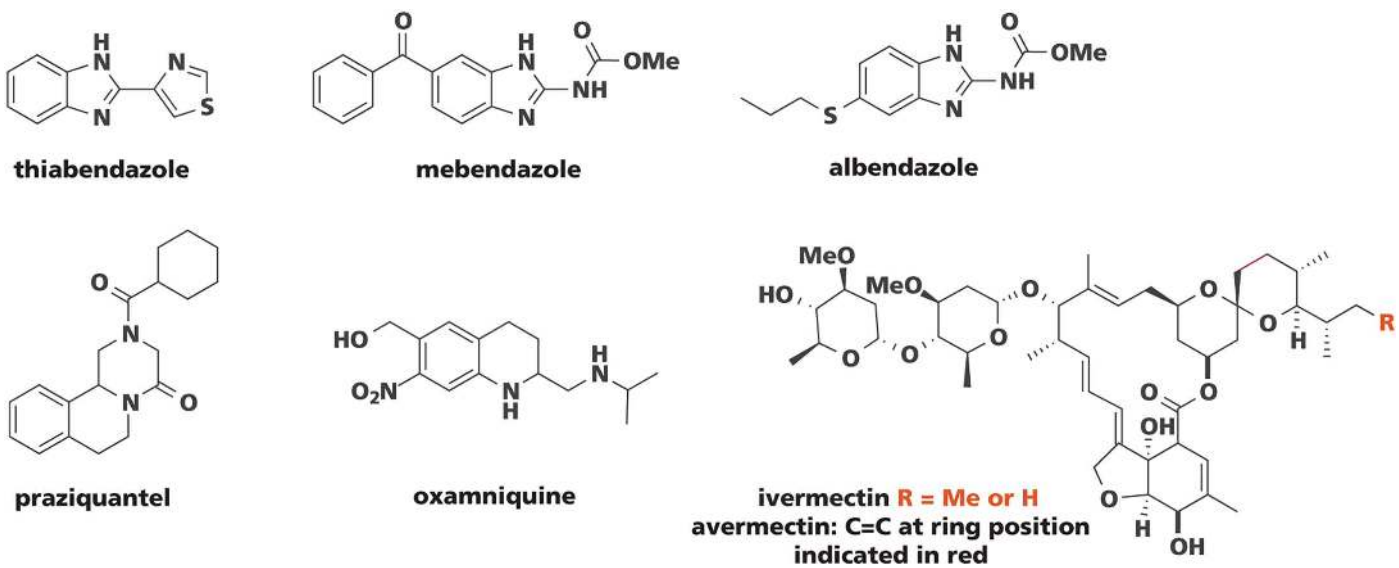


Figure 12.36 Structures of drugs used to treat helminthic diseases. The veterinary medicine thiabendazole was the basis of less-toxic analogs mebendazole and albendazole, used for human disease. The synthetic quinolines praziquantel and oxamniquine are effective against schistosomiasis. Avermectin, a natural product isolated from soil bacteria, and ivermectin, a semisynthetic analog, are used to treat river blindness.

Praziquantel and oxamniquine, introduced in the early 1980s, were discovered by the collaborative efforts of Bayer and Merck to find treatments for schistosomiasis, a disease caused by parasitic flatworms that infects 250 million and kills >200,000 people per year worldwide. The schistosomes infect both mammalian and invertebrate (river snail) hosts and are found in contaminated water, where they penetrate human skin. The parent pyrazinoisoquinolines were originally investigated at Merck as potential tranquilizers in the early 1970s. While inactive for that use, the compounds were shared with Bayer for veterinary screening, where praziquantel was found to be effective against a number of parasitic helminths. Testing in humans indicated high effectiveness against schistosomes (parasitic flatworms) in a single-dose oral formulation and with low toxicity. Its mechanism involves increasing the parasite's membrane permeability toward Ca^{2+}, leading to death of the organisms. Oxamniquine is effective against only one species, *Schistosoma mansoni*.

Ivermectin is a semisynthetic derivative of avermectin and differs only by a reduction of a carbon–carbon double bond in the spiroketal ring (shown in red in **Figure 12.36**). Avermectin was discovered by Satoshi Ōmura in the soil bacterium *Streptomyces avermitilis*. It was purified and identified at Merck Pharmaceuticals by William Campbell, who had also discovered thiabendazole. Ivermectin acts against many parasitic helminths at low doses and with low toxicity. The mechanism involves the modulation of parasite-specific chloride ion channels, leading to the death of the organism. It has become one of the most widely used drugs and is an important part of a program to eradicate river blindness and elephantiasis. Ōmura and Campbell were awarded the Nobel Prize in Physiology or Medicine in 2015 for the discovery of avermectin.

12.10 SUMMARY

Antibacterial and antiparasitic drugs are some of the oldest drugs that have been developed, starting with suramin and Salvarsan in the early 1900s and the sulfonamides in the 1930s. The introduction of antibacterial drugs has had a profound effect on lifespan, and penicillin is considered a wonder drug. Unlike viruses (see Chapter 11), bacteria have a number of unique targets, and these are acted on by different classes of antibacterial drugs. These targets are primarily involved in DNA replication and the integrity of the cell wall. The discovery of antibacterial drugs has most often come about by screening of both natural products and synthetic compounds, rather than *a priori* drug design, although a vast majority of modern antibacterial drugs are derived from medicinal chemistry optimization of natural products. Resistance toward antibiotics is a continual issue, and newer drugs are those that are not prone to those mechanisms. In addition, strategies to avoid common resistance mechanisms constitute an emerging approach. The development of antiparasitic drugs lags behind that of antibacterial drugs. Despite the fact that malaria affects over 3 billion people, only a few effective therapies are available, and resistance toward those is becoming a concern. Older drugs were discovered on the basis of phenotypic screening, and many drugs that are used for other infections (fungal or bacterial) have found use as antiparasitic agents. Drug discovery efforts for neglected diseases, however, have increased due to the significant need, and novel therapeutics from these initiatives are currently in development.

CASE STUDY 1: DISCOVERY OF LINEZOLID

Linezolid was discovered at Upjohn Pharmaceuticals (now part of Pfizer), and when introduced in 2000, it was the first entirely new class of antibacterial drug introduced in 30 years. It was based on a series of oxazolidinone-containing compounds, represented as DuP 105 and DuP 721, first reported by chemists from DuPont Merck in 1987 (**Figure 12.37**).

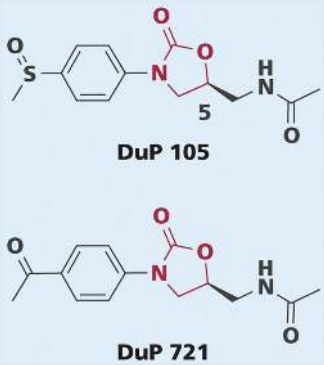


Figure 12.37 Initial oxazolidinone leads for linezolid. Two oxazolidinone antibacterial lead compounds, DuP 105 and DuP 721, were reported by DuPont Merck Pharmaceuticals in 1987. The oxazolidinone group is shown in red.

Promising characteristics of these leads included good *in vitro* and oral activity vs. MRSA and *M. tuberculosis*, as well as good pharmacokinetics in rats. In addition, there was no induction of resistance observed in *in vitro* studies. Structure–activity relationships reported an aryl substituent containing a para electron-withdrawing substituent on the oxazolidinone nitrogen, and a C5 (*S*)-acetamidomethyl group being important for potency. However, efforts in this series were suspended due to toxicity issues.

On the basis of these reports, chemists at Upjohn initiated a program to identify novel oxazolidinones. One strategy was to use conformational constraints to improve activity as well as to impart novelty. Compound **1** is a bicyclic analog of DuP 721 and exhibited potent MRSA *in vitro* and *in vivo* activity (in mice), as well as favorable pharmacokinetics (**Figure 12.38**). Surprisingly, although DuP 721 causes bone marrow toxicity, **1** had no such effects, but since it was covered in the DuPont patent, further structure elaboration was pursued. A second series, in which the ketone was replaced with an acylated amine, provided the novel and patentable 5′-indolinyloxazolidinones **2** and **3**, which also exhibited potent activity and low toxicity.

The project goal was to further optimize the compounds to develop two clinical candidates, each with a vancomycin-like profile, that were effective both intravenously and orally. Three teams of chemists were assigned, and they developed three series of

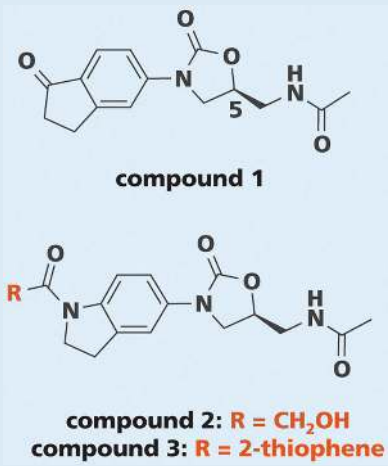


Figure 12.38 Structures of constrained analogs of oxazolidinones. Chemists at Upjohn investigated the oxazolidinones further, including compound **1**, a bicyclic analog of DuP 721. Compounds **2** and **3** replace the ketone of **1** with acylated amines, and both were potent and had low toxicity.

compounds: tropinone, indolinyl, and piperazinyl (Figure 12.39). In the tropinone series, fluorination on the aryl group improved potency, but the compounds had poor solubility and were difficult to synthesize. The indolinyl compounds had a good safety profile but were less potent than the original series. The optimal series proved to be the piperazinyl compounds, which displayed good pharmacokinetics, solubility, potency, and toxicity profiles.

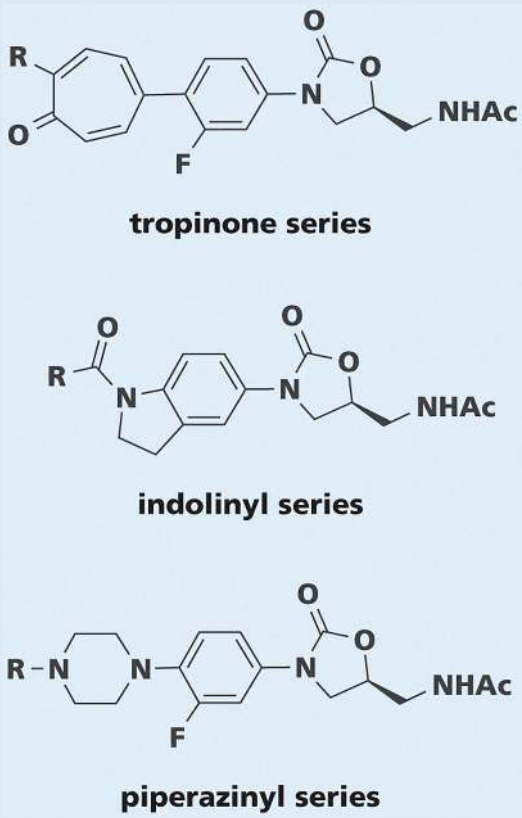


Figure 12.39 Three series of oxazolidinone antibacterial leads. Three series of leads were developed: the tropinone, indolinyl, and piperazinyl structures; the piperazinyl series had the optimal combination of properties.

Focus on this series led to eperezolid, which entered into Phase I trials in 1994 (Figure 12.40). The compound had high solubility and was well tolerated up to 2,000 mg/kg when dosed orally. The fluorine substituent on the aryl ring doubled potency *in vitro* as compared to the unsubstituted analog, and while the addition of a second fluorine atom to the aromatic ring increased potency, that derivative had lower solubility.

Linezolid, the morpholine analog, has a similar profile to eperezolid and entered phase I trials at the same time. Although very similar in potency, linezolid progressed further due to a superior PK profile in humans. It had 100% oral bioavailability, with rapid absorption and minimal metabolism. Minimal toxicity was also seen in clinical

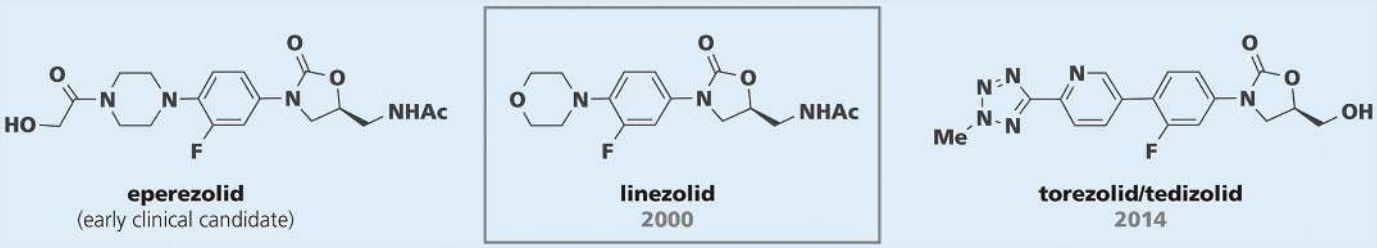


Figure 12.40 Structures of oxazolidinone antibacterial drugs. Eperezolid and linezolid both advanced to phase I clinical trials, with linezolid having better pharmacokinetic properties and advancing to an approved drug. Torezolid (tedizolid) is the only other approved oxazolidinone antibacterial drug.

trials, and very slow development of resistance to linezolid was observed. Linezolid is bacteriostatic for enterococci and staphylococci and bactericidal for streptococci.

While a number of other oxazolidinones have entered clinical trials, as of 2017, only one other has received FDA approval. Torezolid (now called tedizolid) is a biaryltetrazole that has some improvements over linezolid and received approval in 2014.

CASE STUDY 2: SQ109 FOR MULTI-DRUG-RESISTANT TUBERCULOSIS

Tuberculosis is the world's leading cause of death from a single microorganism and kills over 3 million people per year. The infectious agent is *Mycobacterium tuberculosis* (Mtb), and many of the common antibacterial drugs are ineffective against it. Mycobacteria are Gram-positive bacteria which are also acid-fast bacteria; they possess small amounts of peptidoglycan but have a waxy outer layer of glycolipids composed of mycolic acid and arabinogalactan (**Figure 12.41**).

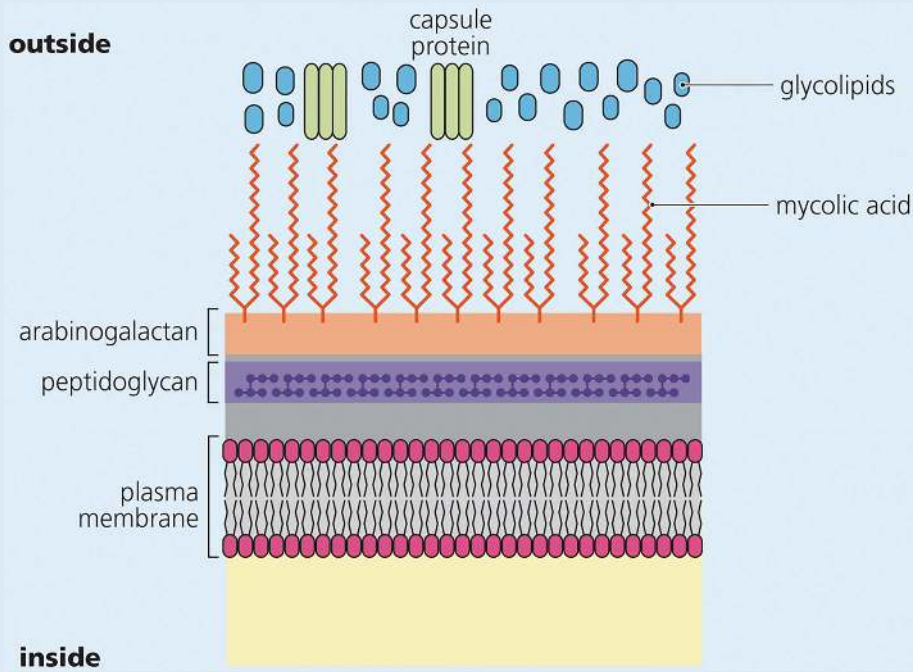


Figure 12.41 Structure of the cell wall of mycobacteria. These bacteria possess an outer capsule (blue/green) of proteins and glycolipids, a waxy layer of mycolic acid (orange), arabinogalactan (orange), and peptidoglycan (purple).

The standard drug treatment for tuberculosis includes a combination of isoniazid (INH), rifampicin (RIF), pyrazinamide (PZA), and ethambutol (EMB), all of which have been used for over 40 years. The combination must be given for 2 months, followed by just INH and RIF for 4–6 months. This lengthy regimen, along with the toxicity of the drugs, results in poor compliance. Noncompliance has led to multi-drug-resistant tuberculosis (MDR-TB), which is resistant to both drugs, as well as extensive-drug-resistant tuberculosis (XDR-TB), which is resistant to second-line drugs, such as quinolones, kanamycin, and capreomycin.

In the 1990s, renewed TB drug discovery efforts focused on both identifying new drugs and investigating repurposed drugs. In 1999, scientists at Sequella began a combinatorial chemistry approach to synthesizing analogs of EMB (**Figure 12.42**). Ethambutol had been discovered in the 1950s, and no extensive SAR was reported. Although it is part of the standard regimen, the drug causes toxicity to the eye, which limits the dose that can be administered.

By use of split-pool combinatorial chemistry techniques (see Section 2.5.3), a library of 63,238 compounds was created, based on the ethylenediamine core (see

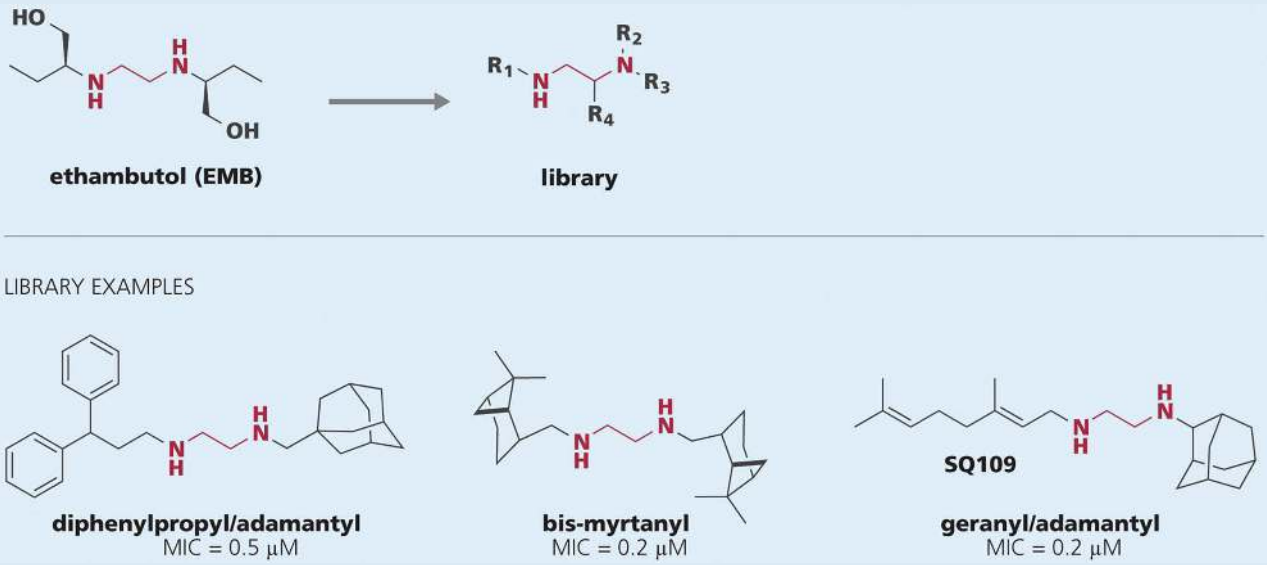


Figure 12.42 Ethambutol scaffold and library design. Ethambutol was the basis for combinatorial synthesis of 63,238 compounds (*top*). Three examples of the 170 active compounds from combinatorial synthesis are shown and include representatives of diphenyl propyl, myrtanyl amine, and isoprenoid scaffolds. The geranyl/adamantyl analog SQ109 advanced to clinical trials. The common ethylene diamine core in each is shown in red.

Figure 12.42). Compounds were synthesized in groups of ten and evaluated as mixtures in a broth microdilution assay to determine the minimum inhibitory concentration (MIC) vs. Mtb as well as in a bioluminescence assay for inhibition of Mtb cell wall formation. Components of active ten-compound pools were resynthesized individually to identify active compounds, and the SAR was evaluated. They identified 2,796 ten-compound mixtures as active, and deconvolution generated 170 individual compounds that had MICs in the <0.6 μM range in both screens.

Most of the active compounds were lipophilic, with examples being extensively α-branched aliphatic, diphenylethyl, and diphenylpropyl fragments; tricyclic skeletons from adamantanes, myrtanylamine, and isopinocamphylamine; and isoprenoids (select examples shown in Figure 12.42).

The top 27 compounds, based on potency, were evaluated more extensively in cytotoxicity assays in HepG2 cells to determine a selectivity index (SI). SQ109 had the lowest MIC and highest SI in this group. The top 13, based on potency and high SI, were also tested in secondary assays for effects in Mtb-infected macrophages. The top 11, based on potency and efficacy in that intracellular assay, proceeded to *in vivo* trials in mice. The three most promising were investigated in further *in vivo* studies (mice) to determine their maximum tolerated dose (MTD) and evaluate pharmacokinetic (PK) parameters. Overall, SQ109 exhibited the most potent activity, lowest toxicity, and best MTD/PK profile, and it had activity against MDR strains. It was also found to be effective in a chronic model of TB in mice at 10 mg/kg, and it became the lead candidate for clinical trials after PK and toxicity studies in rats, dogs, and primates.

The ADMET profile in preclinical trials showed the highest concentrations of the drug in the liver, but also significant levels in the lung and spleen, which are TB target organs. Metabolism was primarily by cytochrome P450 enzymes from oxidation, epoxidation, and N-dealkylation. No significant toxicity was observed at doses well above the effective dose. SQ109 received Fast Track designation by the FDA in 2007, as well as an orphan drug status. Safety testing in Phase I clinical trials showed no adverse effects, and as of 2025, it had advanced to Phase III clinical trials.

SQ109 works by inhibition of MmpL3, which is different from the mechanism by which EMB works. EMB inhibits the synthesis of the arabinogalactan of the cell wall and is bacteriostatic. In contrast, with SQ109 treatment, there is a decrease not in synthesis of cell wall components but in the incorporation of mycolic acids into the cell wall, specifically, trehalose dimycolate and arabinogalactan mycolates. There is evidence that SQ109 targets the transport of mycolates, which is supported by the fact that mutants resistant to SQ109 had a mutant TMM transporter, called MmpL3. This transporter is found only in mycobacteria.

12.11 REVIEW QUESTIONS

1. Outline three mechanisms of action of antibacterial drugs.
2. Both quinolones and sulfonamides target bacterial DNA synthesis. Explain how their mechanisms are different.
3. Describe three classes of cell wall synthesis inhibitors that act by covalent binding.
4. Describe two different mechanisms that are utilized by bacteria to resist penicillins.
5. Describe the main difference between Gram-positive and Gram-negative bacteria and why drugs might be effective against one and not the other.
6. What are neglected tropical diseases, and what specific requirements should be considered for drug discovery and development programs for these diseases?
7. What are the two classes of drugs used to treat malaria, and what are their mechanisms?
8. What drug discovery projects have contributed to increased numbers of drugs to treat parasitic disease?
9. Identify three antiparasitic drugs that are either natural products or derived from natural products, and for each, briefly describe what the drug is used for and its mechanism of action.
10. Explain why some drug discovery targets for antibacterial agents are similar to those used in anti-cancer drug discovery.

12.12 APPLICATION QUESTIONS

1. Draw the products of the reaction of PBP (particularly the active site serine OH) with cefuroxime.
2. Using the examples in the chapter, identify two different compounds that are conformationally constrained.
3. Using the examples in the chapter, identify a compound that uses a prodrug strategy, and research the advantage of using the prodrug.
4. Using the examples in the chapter, identify three different compounds that incorporate isosteric replacements.
5. The Nobel Prize in Physiology or Medicine 2015 was awarded for the discovery of medicines for infectious diseases. Research the awardees and the medicines to which they contributed.
6. PBP is a serine protease. On the basis of your knowledge of serine protease inhibitors from Chapter 7, list other functional groups that could be used in place of the β-lactam in compounds such as penicillin.
7. Review the structures in this chapter, and identify at least two that do not obey Lipinski's rule. For these two, calculate their Lipinski parameters. Use public data (such as PubChem) or other programs (such as ChemDraw) to approximate log$P$. Investigate whether they are orally active.
8. Research the structure and activity of anthracycline anti-cancer agents. Compare them to the tetracyclines described in this chapter. What are the similarities and differences in structure and activities?
9. Vancomycin resistance can occur when the D-Ala-D-Ala terminus of the substrate is mutated to D-Ala-D-lactate (substitution of oxygen for N-H). Design an analog of vancomycin that would be effective against the D-Ala-D-lactate mutant.
10. Many third- and fourth-generation quinolones contain a fluorine atom at several different positions. Given what you have learned about the properties of fluorine (see Chapter 4), speculate on how the introduction of that atom changes and improves the properties of these drugs.

FURTHER READING

Butler MS, Blaskovich MA & Cooper MA (2013) Antibiotics in the clinical pipeline in 2013. *J Antibiot* 66:571–591.

Kohanski MA, Dwyer DJ & Collins JJ (2010) How antibiotics kill bacteria: from targets to networks. *Nat Rev Drug Discov* 8:423–435.

Neu H (1992) The crisis in antibiotic resistance. *Science* 257:1064–1073.

Cell Wall Synthesis Inhibitors

Fleming A (1929) On the antibacterial action of cultures of a *Penicillium*, with special reference to their use in the isolation of *B. influenzae*. *Br J Exp Pathol* 10:226–236.

Knox JR & Pratt RF (1990) Different modes of vancomycin and D-alanyl-D-alanine peptidase binding to cell wall peptide and a possible role for the vancomycin resistance protein. *Antimicrob Agents Chemother* 34:1342–1347.

Levine DP (2006) Vancomycin: a history. *Clin Infect Dis* 42 (Suppl 1): S5–S12.

Llarrull LI, Testero SA, Fisher JF & Mobashery S (2010) The future of β-lactams. *Curr Opin Microbiol* 13:551–557.

Papp-Wallace KM, Endimiani A, Taracila MA & Bonomo RA (2011) Carbapenems: past, present, and future. *Antimicrob Agents Chemother* 55:4943–4960.

Protein Synthesis Inhibitors

Beringer M & Rodnina MV (2007) The ribosomal peptidyl transferase. *Mol Cell* 26:311–321.

Brodersen DE, Clemons WM, Carter AP et al. (2000) The structure basis for the action of the antibiotics tetracycline, pactamycin, and hygromycin B on the 30S ribosomal subunit. *Cell* 103:1143–1154.

Carter AP, Clemons WM, Brodersen DE et al. (2000) Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature* 407:340–348.

Chopra I & Roberts M (2001) Tetracycline antibiotics: mode of action, applications, molecular biology, and epidemiology of bacterial resistance. *Microbiol Mol Biol Rev* 65:232–260.

Connell SR, Tracz DM, Nierhaus KH & Taylor DE (2003) Ribosomal protection proteins and their mechanism of tetracycline resistance. *Antimicrob Agents Chemother* 47:3675–3681.

Kotra LP, Haddad J & Mobashery S (2000) Aminoglycosides: perspectives on mechanisms of action and resistance and strategies to counter resistance. *Antimicrob Agents Chemother* 44:3249–3256.

Pal S (2006) A journey across the sequential development of macrolides and ketolides related to erythromycin. *Tetrahedron* 62:3171–3200.

Rafie S, MacDougall C & James CL (2010) Cethromycin: a promising new ketolide antibiotic for respiratory infections. *Pharmacotherapy* 30:290–303.

Recht MI, Douthwaite S & Puglisi JD (1999) Basis for prokaryotic specificity of action of aminoglycoside antibiotics. *EMBO J* 18:3133–3138.

Zuckerman JM (2004) Macrolides and ketolides: azithromycin, clarithromycin, telithromycin. *Infect Dis Clin N Am* 18:621–649.

DNA Synthesis Inhibitors

Aldred KJ, Kerns RJ & Osheroff N (2014) Mechanism of quinolone action and resistance. *Biochemistry* 53:1565–1574.

Andriole VT (2005) Quinolones: past, present, and future. *Clin Infect Dis* 41:S113–119.

Chu CT & Fernandes PB (1980) Structure activity relationships of the fluoroquinolones. *Antimicrob Agents Chemother* 33:131–135.

Gleckman R, Blagg N & Joubert DW (1981) Trimethoprim: mechanisms of action, antimicrobial activity, bacterial resistance, pharmacokinetics, adverse reactions, and therapeutic indication. *Pharmacotherapy* 1:14–19.

Wohlkonig A, Chan PF, Fosberry AP et al. (2010) Structural basis of quinolone inhibition of type II topoisomerases and target-mediated resistance. *Nat Struct Mol Biol* 17:1152–1153.

Antiparasitic Drugs

Andrews KT, Fisher G & Skinner-Adams TS (2014) Drug repurposing and human parasitic protozoan diseases. *Int J Parasitol: Drugs Drug Resist* 4:95–111.

Chabala JC, Mrozik H, Tolman RL et al. (1980) Ivermectin, a new broad-spectrum antiparasitic agent. *J Med Chem* 23:1134–1136.

Headrick DR (2014) Sleeping sickness epidemics and colonial responses in east and central Africa, 1900–1940. *PLoS Neglected Trop Dis* 8:e2772.

Hotez PJ, Brindley PJ, Bethony JM et al. (2008) Helminth infections: the great neglected tropical disease. *J Clin Invest* 118:1311–1321.

Jacobs RT, Nare B, Wring SA et al. (2011) SCYX-7158, an orally-active benzoxaborole for the treatment of stage 2 human African trypanosomiasis. *PLoS Neglected Trop Dis* 5:e1151.

Pink R, Hudson A, Mouries M-A & Bendig M (2005) Opportunities and challenges in antiparasitic drug discovery. *Nat Rev Drug Discov* 4:727–740.

SQ 109

Lee RE, Protopova M, Crooks E et al. (2003) Combinatorial lead optimization of [1,2]-diamines based on ethambutol as potential antituberculosis preclinical candidates. *J Comb Chem* 5:172–187.

Tahlan K, Wilson R, Kastrinsky DB et al. (2012) SQ109 targets MmpL3, a membrane transporter of trehalose monomycolate involved in mycolic acid donation to the cell wall core of *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* 56:1797–1809.



Drugs Acting on the Central Nervous System

13

LEARNING OBJECTIVES

- Learn the history of treatment for psychiatric and neurodegenerative conditions.
- Become familiar with CNS-acting drugs and how they were developed.
- Identify mechanisms of drugs used to treat psychiatric disorders.
- Describe targets and approaches to treat neurodegenerative diseases.
- Develop an appreciation for the challenges of CNS drug discovery.
- Learn some novel approaches to future CNS drugs.

13.1 INTRODUCTION

Disorders of the central nervous system (CNS) affect many millions of people worldwide and, according to the World Health Organization, account for over 11% of total deaths. They are typically divided into neuropsychiatric (or mental health) disorders and neurodegenerative diseases. Examples of mental health disorders include anxiety, depression, bipolar disorder, and schizophrenia. Neurodegenerative diseases, such as Parkinson's and Alzheimer's diseases, whose incidence increases with age, affect over 7 million people in the United States, and this number is expected to increase as baby boomers age. In addition, pain is also considered a CNS disorder; whether chronic or more transient, it can be debilitating and affects roughly 17% of the population in the United States alone.

Some of the oldest medicines known were used to treat CNS disorders, such as plant preparations from the opium poppy or willow barks that were used to alleviate pain. However, other CNS conditions, particularly neuropsychiatric disorders, such as psychoses and depression, which have been known for thousands of years, were not treatable until the 20th century, when psychotherapy and drug therapy were introduced. Even today, while there are many drugs available to treat these diseases, patients find varying levels of success with them. Some are effective, but only in a subset of the patient population, and/or are accompanied by significant side effects that impact compliance. Others improve some symptoms, but not all. Finally, some work for a limited time, after which patients become refractory. It is generally believed that there is still considerable room for improvement in drugs that treat most psychiatric disorders.

Similarly, while symptoms of certain neurodegenerative diseases were reported in ancient times, there are few drugs that effectively treat them. While some neurodegenerative disorders are the result of injury or genetic factors, a majority are associated with advanced age. With an increase in the number of people reaching their 80s, 90s,

DOI: 10.1201/ 9781003465607-16

and even 100 years old, the numbers of patients who suffer from neurodegenerative diseases are expected to increase dramatically. In fact, in the United States, it has been estimated that, as of 2025, over 7 million people suffer from Alzheimer's disease, and that number is expected to triple by 2050. Therefore, there has been increased urgency to identify drugs that treat diseases such as Parkinson's and Alzheimer's.

Many drugs in use to treat psychiatric and neurodegenerative disorders affect neurotransmitter (NT) levels, either increasing or decreasing their concentration. These NTs include dopamine, serotonin, norepinephrine, acetylcholine, glutamate, and γ-aminobutyric acid (GABA). The role of neurotransmitters is to relay messages along their corresponding neurons; this is accomplished through their release from the presynaptic neuron into the synapse and interaction with receptors on the postsynaptic neuron. This release affects the transmission of ions, such as Na^+, Ca^{2+}, K^+, and Cl^-, and changes the electrical potential. The resting potential of a neuron is –70 mV; **excitatory neurotransmitters** (NTs) increase the concentration of cations, causing depolarization and relay of an electrical signal along the axon (action potential). **Inhibitory neurotransmitters** increase the concentration of anions, causing hyperpolarization and decreasing the relay. Important drug targets are the GPCR neurotransmitter receptors, ion channels, and neurotransmitter transporters (see Chapter 6) involved in reuptake. The pre- and postsynaptic neurons and associated receptors, channels, and transporters of these neurotransmitters are illustrated in Figure 13.1.

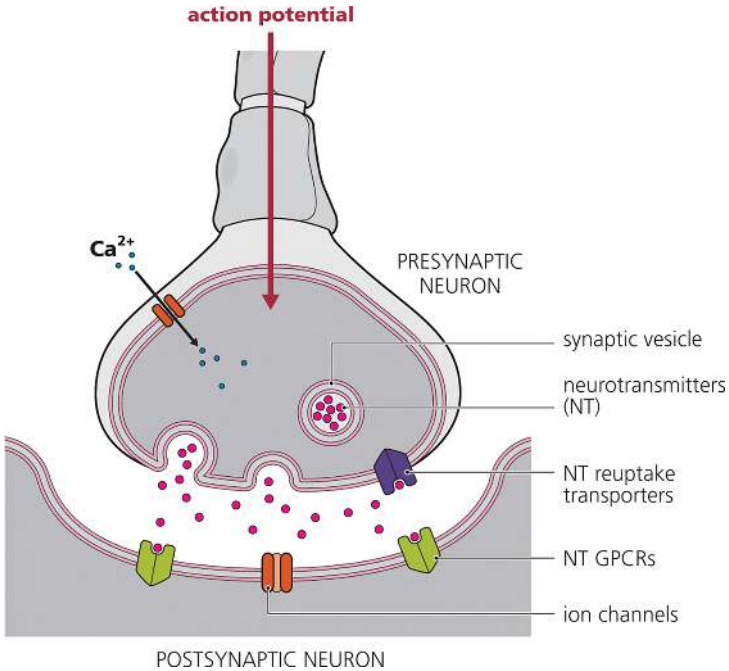


Figure 13.1 Schematic drawing of two neurons (presynaptic and postsynaptic) and the synapse between them. The action potential travels along the axon, ultimately effecting the release of neurotransmitters (NTs, shown as pink circles) from the synaptic vesicles. NTs bind to receptors, such as GPCRs (green) on the postsynaptic neuron, and to NT reuptake transporters (purple) on the presynaptic neuron. They also affect voltage- and ligand-gated ion channels (orange) on either neuron, causing ions such as Ca^{2+} to enter (shown) or be released.

Synaptic vesicles containing neurotransmitters release them when their membranes fuse with the outer cell membrane of the neuron. Once released, some neurotransmitter molecules cross the synaptic cleft and bind to receptors, including ligand-gated ion channels (LGICs) and G protein–coupled receptors (GPCRs) on the postsynaptic neuron. Binding to these receptors results in changes in ion concentration, either directly for LGICs or indirectly through GPCRs.

When concentrations of NTs in the synaptic cleft reach certain levels, NT transporters remove them in a process called reuptake as one mechanism to regulate the concentration of NTs and their signaling. Once removed from the cleft, the NTs are repackaged in vesicles.

The release and reuptake of various NTs are controlled by multiple factors, and there is interplay between individual NTs. Neurological diseases are sometimes caused by an imbalance in NT levels, with NT roles in the disease state depending on their location in the brain. Structures of some small-molecule neurotransmitters are shown in Figure 13.2, and an overview of the role of specific neurotransmitters involved in CNS disorders is given here. While all are amines, serotonin, dopamine, and norepinephrine are commonly referred to as **monoamine neurotransmitters**.

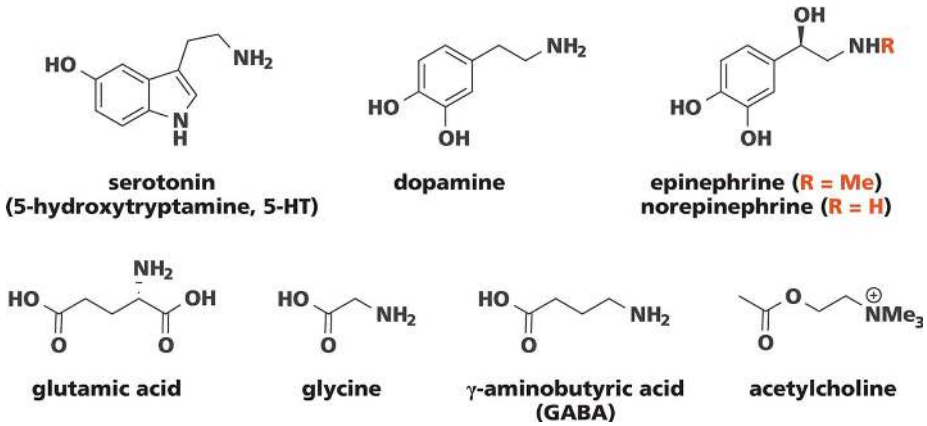


Figure 13.2 Structures of neurotransmitters involved in CNS disorders. Small molecules that interact with GPCRs, reuptake transporters, and ion channels to regulate synaptic transmission include serotonin, dopamine, epinephrine, norepinephrine, glutamic acid, glycine, γ-aminobutyric acid, and acetylcholine.

Serotonin, or **5-hydroxytryptamine** (5-HT), is an inhibitory neurotransmitter that regulates mood as well as sleep, appetite, memory, body temperature, and sexual function. Low levels of serotonin are associated with both anxiety and depression, and its dysregulation is also thought to have a role in schizophrenia, as well as bipolar disorder. **Dopamine** is both inhibitory and excitatory and plays roles in motor behavior, motivation, and emotion. Two conditions that are caused by an imbalance in dopamine levels, although in opposite directions, are Parkinson's disease and schizophrenia. Low levels of dopamine and loss of dopaminergic neurons (neurons primarily releasing dopamine) are implicated in Parkinson's disease, while increased dopamine levels and dopaminergic neuron activity are thought to contribute to the psychoses experienced by schizophrenics. Dopamine levels are regulated in part by other neurotransmitters, and nearly all drugs of abuse raise dopamine levels either directly or indirectly. This relationship is the basis for the **dopamine reward system**, also known as the mesolimbic dopamine system in the brain. This circuit connects the ventral tegumental area (VTA) and the nucleus accumbens (NAc), and activation controls responses to rewards such as food and sex.

Epinephrine and **norepinephrine** (also known as noradrenaline and adrenaline) are excitatory neurotransmitters. They are involved in the regulation of sleep and alertness, and in the periphery, they also have roles in controlling heart rate and bronchodilation. In the CNS, high levels of these NTs cause anxiety and can induce paranoid psychosis. **Glutamic acid** is also an excitatory neurotransmitter with roles in synaptic plasticity, which refers to changes in the connections between neurons. Hyperfunction of its receptors causes excitotoxicity that may lead to seizures. **Glycine** is an inhibitory neurotransmitter and plays a role in movement. In fact, several movement disorders are associated with mutations in the glycine receptor. Glycine is also a co-agonist at the *N*-methyl-D-aspartate (NMDA) subtype of glutamate receptors, and in that role, it is excitatory. **γ-aminobutyric acid** (GABA) is an inhibitory neurotransmitter, which balances the excitatory NTs. Low levels can lead to anxiety, depression, and convulsions. **Acetylcholine** affects muscle contraction and has a role in memory. Table 13.1 summarizes these neurotransmitters and their roles in psychiatric and neurodegenerative disorders.

The brain's architecture relies on a unique barrier, the **blood–brain barrier** (BBB), to protect it. Paul Ehrlich demonstrated the existence of the BBB in the late 1880s when he injected a dye into the bloodstream of mice and saw that all the organs become stained except the brain. In a related experiment almost 30 years later, injection of the same dye into mouse brains only stained the brains, and no other organs, proving the existence of a barrier between the brain and the bloodstream. The BBB is made up of tightly packed endothelial cells that are denser than other membrane structures. This architecture protects the brain from toxins and infectious agents but also provides an additional barrier for drugs to penetrate. A number of studies have attempted to characterize the physical properties of drugs that are active in the CNS, and therefore must cross the BBB, and compare them to properties of drugs that affect other organ systems. In general, CNS-active agents are of lower molecular weight, higher lipophilicity (log*P*), contain fewer H-bond donors and acceptors, and have smaller polar surface areas compared to drugs that act in the periphery. Of these, lipophilicity is considered

Table 13.1 Neurotransmitters and Drug Targets That Modulate NT Levels, Disorders Implicated in NT Imbalance, and Examples of Drug Affecting NT Levels

Neurotransmitter	Drug Target or Mechanism	Disorder	Drug Examples
Serotonin (5-HT)	5-HT transporter and receptors	Depression	Fluoxetine, citalopram
		Migraine	Sumatriptan
		Schizophrenia	Aripirazole, quetiapine
		Social anxiety	Paroxetine, citalopram
		Obesity	Lorcaserin
Dopamine	Dopamine receptors and transporter; biosynthesis	Parkinson's disease	L-dopa
		Schizophrenia	Thorazine, clozapine
GABA	GABA receptor	Anesthesia	Phenobarbital, lorazepam, alprazolam
		Anxiety	Lorazepam, alprazolam
		Epilepsy	Phenobarbital
		Insomnia	Zolpidem
Norepinephrine	Norepinephrine receptors and transporter	Anxiety	Duloxetine
		Depression	Reboxetine
		ADHD	Methylphenidate
Acetylcholine	Acetylcholine receptors, biosynthesis	Alzheimer's disease	Tacrine, donepezil
Glutamate	Glutamate receptors	Alzheimer's disease	Memantine
	NMDA receptors	Schizophrenia, depression	Pomaglumetad* ketamine
Glycine	Glycine transporters	Schizophrenia	Iclepertin*

*Indicates research probes or drugs not yet approved.

the most important determinant. In addition to the tightly packed cells that prevent diffusion across the BBB, the brain also expresses high levels of the P-glycoprotein transporter (also called MDR1), which expels molecules that have diffused into the brain. Assays measuring the permeability of compounds using cells that overexpress MDR1 are one way to predict brain penetration of a drug candidate.

Many drugs targeting neuropsychiatric disorders are based on neurotransmitters and are either agonists or antagonists at their receptors. Ligand-based drug design and the use of pharmacophore models have been widely used in the development of most of these agents. There are far fewer drugs approved to treat neurodegenerative diseases, and while some affect neurotransmitter levels, others are in development that inhibit enzymes involved in these diseases. The challenges in understanding the complex (and often multifactorial) basis for many CNS disorders make drug discovery particularly challenging.

13.2 DRUGS TARGETING ANXIETY AND DEPRESSION

Anxiety disorders are the most common mental health disorder, affecting about 19% of Americans, with a higher prevalence in women than in men. Depression is a little less common but still affects approximately 8% of the US population. These are related conditions, with about half of those being diagnosed with depression also suffering from anxiety disorders, and both are challenging to treat. Some patients may respond to drug treatment, while others are refractory, and drug trials are complicated by a

large placebo effect. In addition, some drugs have a multiweek lag time before they take effect, resulting in lower patient compliance.

Neurotransmitters believed to be involved in anxiety disorders include GABA, glutamate, serotonin, and norepinephrine. GABA binds at the $GABA_A$, $GABA_B$, and $GABA_C$ receptors: The A and C receptor subtypes are chloride ion channels, and the B subtype is a GPCR. Of these, $GABA_A$ is the most important target in terms of treating anxiety and depression, where increased GABA activity is correlated with reduced anxiety. Glutamate is an excitatory neurotransmitter that counterbalances the inhibitory GABA. Low levels of serotonin and norepinephrine are both implicated in depression, and many drugs used to treat depression raise levels of these NTs. A prominent mechanism for raising levels is the inhibition of the NT transporters, which raises levels by blocking reuptake of the NTs.

13.2.1 Positive Allosteric Modulators at the $GABA_A$ Receptor Are Sedatives and Anticonvulsants

The $GABA_A$ receptor is a ligand-gated chloride ion channel, and positive allosteric modulators (PAMs) act as sedatives as well as anticonvulsants. Binding of either GABA or PAMs increases intracellular chloride ion concentration, resulting in membrane hyperpolarization and reduced transmission of the action potential.

The basic structure of the $GABA_A$ receptor is a hetero-oligomer, with different combinations of various subunits resulting in a variety of subtypes (**Figure 13.3**). However, the majority of $GABA_A$ receptors identified in the brain are composed of two α, two β, and one γ subunit. The natural ligand GABA binds at two extracellular sites on the α/β interface. Different classes of drugs acting at the GABA receptor bind at different allosteric sites on the receptor.

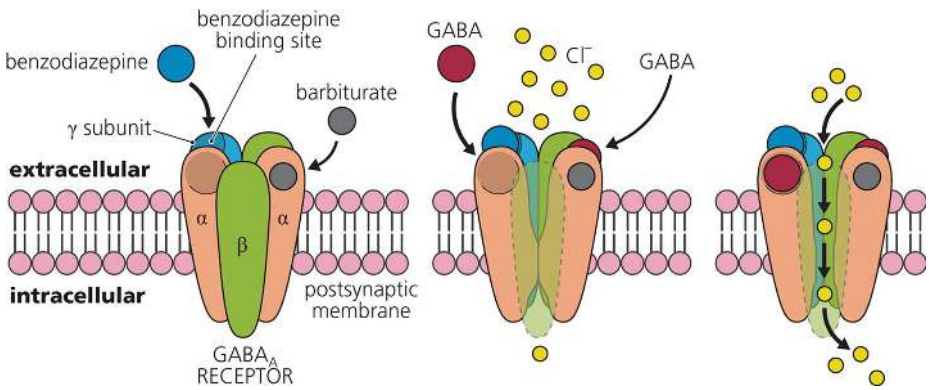


Figure 13.3 Simplified structure and binding sites of small molecules on the $GABA_A$ receptor chloride ion channel. *Left*, the $GABA_A$ receptor, located on the postsynaptic neuron, is a ligand-gated channel for chloride ions (shown as yellow circles). It consists of two α subunits (orange), two β subunits (green), and one γ subunit (blue). *Center*, two molecules of GABA (red circles) bind between the α and β subunits, resulting in channel opening. *Right*, benzodiazepines (blue circles) bind at the α/γ interface, resulting in increased binding of GABA and increases in channel opening, and barbiturates (gray circles) are thought to bind on the α subunit.

Barbiturates, as a class, act as sedatives and hypnotics. The first approved member of this class, barbital, was marketed in 1904 (**Figure 13.4**). It is a derivative of barbituric acid, synthesized in the 1860s as a condensation product of urea and malonic acid. Although barbituric acid is not pharmacologically active, it is the parent for the barbiturates. In Germany, chemists Josef von Mering and Emil Fischer made barbital, the diethyl analog of barbituric acid, in part on the basis of the observation that the bis(diethylsulfone), sulfonal, had activity as a hypnotic. Prior to that time, the most widely used sedative was the salt potassium bromide, which was introduced in 1857 as a sedative and anticonvulsant. Barbital was used to treat insomnia, sedate psychotic patients, and induce anesthesia. Its use to treat epileptic seizures was also described in 1912. Many analogs of the

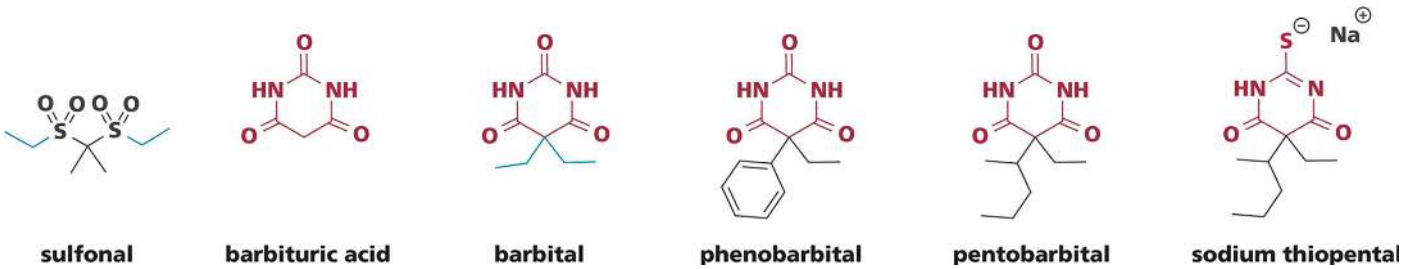


Figure 13.4 Structural basis and examples of barbiturate drugs, in order of increasing complexity. Sulfonal was an early hypnotic, and combining some of its structural features (ethyl moieties) with inactive barbituric acid formed the basis for the barbiturate analog barbital. Drugs used today include phenobarbital, pentobarbital, and sodium thiopental. The ethyl groups of sulfonal and barbital are shown in blue, and the barbituric acid core is shown in red.

barbiturates were made in industrial and academic laboratories. The barbital scaffold is the basis of a number of drugs that are used for various indications.

The derivative phenobarbital is one of the most widely used agents to treat seizure disorders, and pentobarbital is primarily used as a sedative and in anesthesia. Sodium thiopental, a sulfur analog, is used as a rapid-onset, short-acting anesthetic. It is also one of a cocktail of drugs used for lethal injection and has been used as a truth serum, although the lowered inhibitions may lead to false as well as true statements.

With the wide use of barbiturates, it became obvious that their frequent use could result in dependence. In fact, by 1960, there were an estimated 250,000 barbiturate addicts in the United States. In addition, these drugs have a low therapeutic index, sometimes resulting in fatal overdose. While their use to treat anxiety has been supplanted with much safer drugs, the barbiturates are still used as anticonvulsants and in anesthesia. They are thought to bind near the GABA binding sites on the α subunit (see **Figure 13.3**). The mechanism of barbiturates was not fully understood until the late 1970s, when they were shown to potentiate inhibitory effects on neurotransmission of GABA.

The **benzodiazepines** (BZs) were discovered in the 1950s in Leo Sternbach's lab at Hoffmann–La Roche. As with other synthetic drugs discovered in this era, the program was driven by ease of synthesis and *in vivo* activity rather than a specific target; the GABA receptor activity was identified much later. Although sometimes erroneously classified in popular literature as accidental, the discovery was driven by a rational search for tranquilizers. Sternbach synthesized many compounds that were derivatives of structures originally assigned in the 1850s as benzheptoxdiazines (**Figure 13.5a**). During the course of his work, these structures were reassigned as quinazoline-3-oxides. None of the compounds synthesized were active in mouse models of sedation, and after several years, the project was abandoned. At the time of the work, purification methods were limited, and only compounds that could be recrystallized (as a method for purification) were subjected to testing. During a cleanup of the lab, Sternbach's associate Earl Reeder found an untested sample that had crystallized, and upon isolation and assay in a mouse model, it was found to have promising activity as a tranquilizer. Reinvestigation of the structure showed that this particular compound actually contained a 1,4-benzodiazepine scaffold, which was formed by a rearrangement during the synthesis (**Figure 13.5b**). During the reaction of the starting α-chloromethylquinazoline-3-oxide with methylamine, the expected S_N2 displacement of the chloride did not occur. Instead, methylamine added to the imine to give an unstable intermediate that led to a ring-opened oxime. Displacement of the chloride occurred at this point by reaction of the oxime nitrogen to give chlordiazepoxide.

Chlordiazepoxide was found to have activity as a muscle relaxant, tranquilizer, and anticonvulsant. It was devoid of the hypnotic effects of the barbiturates and had very low toxicity. It was marketed in 1960 as Librium, the first benzodiazepine (BZ) to be used as a tranquilizer.

Librium was soon replaced by the more stable diazepam (Valium), which was the top-selling drug between 1968 and 1982 (**Figure 13.6**). While their low toxicity and profile of muscle relaxation without sedation meant that benzodiazepines rapidly replaced barbiturates, their mechanism of action remained unknown until 1977, when they were found to interact with specific receptors in the brain. These were cloned and identified as the $GABA_A$ receptors in 1987. Benzodiazepines bind between the α and γ subunits of the receptor and modulate binding of GABA (see **Figure 13.3**). Positive allosteric modulators (PAMs), such as diazepam, enhance GABA activity and are

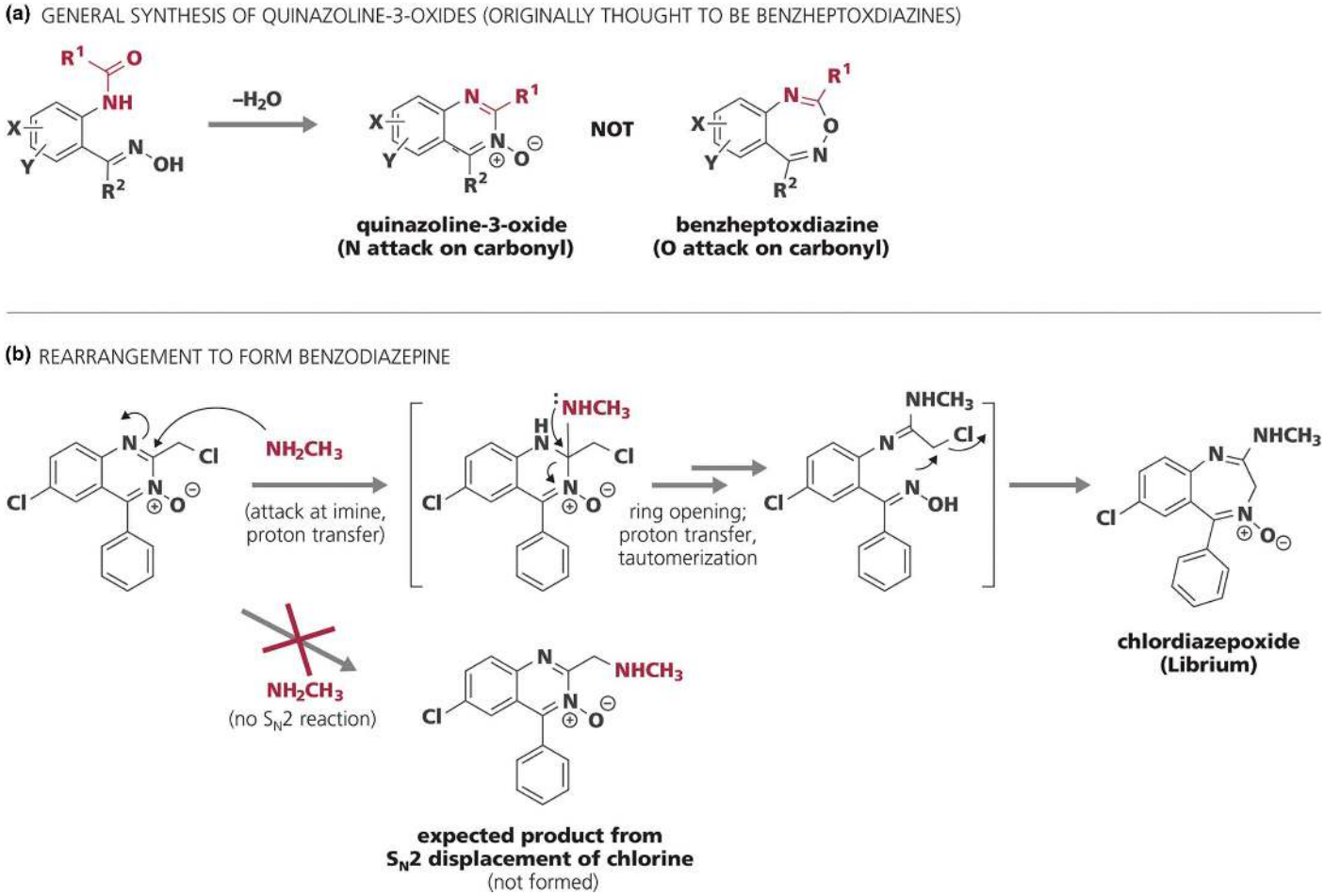


Figure 13.5 **History of chlordiazepoxide synthesis.** (a) General synthesis of quinazoline-3-oxides. Compounds with various aryl (X, Y) and amide substituents R^1 (amide shown in red) were prepared. Attack of the hydroxylamine nitrogen at the amide carbonyl of the starting material gives the quinazoline-3-oxide. The structure was originally assigned as the benzheptoxdiazine that would result from attack by the oxime oxygen on the amide. (b) Rearrangement of a specific quinazoline-3-oxide (X = 6-chloro, Y = H, $R^1 = CH_2Cl$, R^2 = Ph) to benzodiazepine scaffold. Reaction of 6-chloro-2-chloromethyl-quinazoline-3-oxide with an amine was expected to generate the simple displacement product via S_N2 reaction. In the case of methylamine, addition to the imine occurred instead of the expected S_N2 displacement. Ring opening, proton transfer, and ring closure produced chlordiazepoxide (abbreviated mechanism shown).

used to treat anxiety and convulsive disorders, such as epilepsy. As of 2025, clonazepam, lorazepam, and alprazolam are the most widely prescribed benzodiazepines for anxiety and panic disorders. The short-acting drug midazolam is commonly used for the induction of anesthesia. Antagonists of benzodiazepine binding, such as flumazenil, can be used to treat benzodiazepine overdose. Although benzodiazepines were originally believed to be nonaddictive, and therefore much safer than barbiturates, dependence is indeed seen with their long-term use. The structure–activity relationships (SAR) for benzodiazepines PAMs include an aromatic A ring with an electron-withdrawing group at C7, a carbonyl or equivalent unsaturation at C2, and a nitrogen at position 1. Pharmacophore models for the benzodiazepines have been developed and used in virtual screening programs to identify new ligands for the GABA BZ binding site (see Section 3.3.4).

In addition to the benzodiazepines, zolpidem and eszopiclone also have been found to bind at the benzodiazepine site on the $GABA_A$ receptor. These drugs have hypnotic effects at a low dose, resulting in less impact on cognition and psychomotor reflexes than benzodiazepines, and they are primarily used to treat insomnia. Unlike BZs, which bind to multiple subtypes of the α subunit, zolpidem binds at the BZ site only on certain α subunits, specifically the α_1 subtype; it is believed this selectivity results in its hypnotic effects and differentiates it from the BZs that bind to the α_1, α_3, and α_4 subtypes, a characteristic of drugs that have anti-anxiety and anticonvulsant effects.

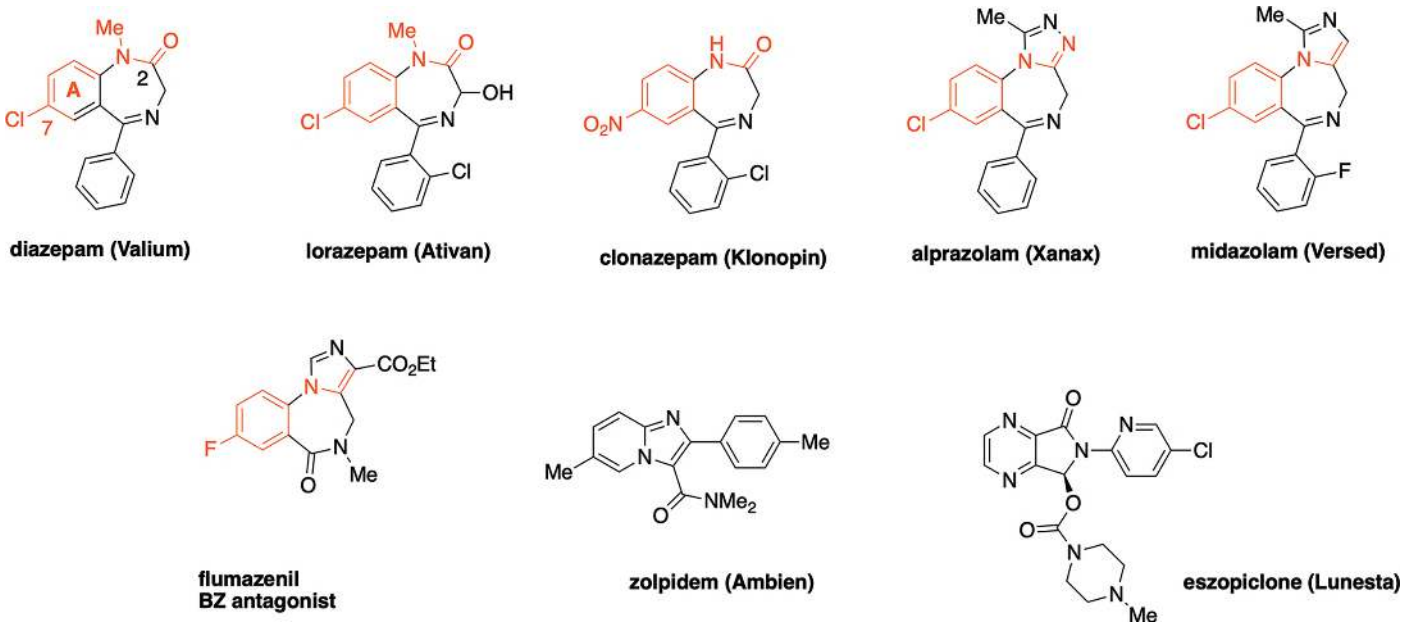


Figure 13.6 Structures of benzodiazepines and other drugs that bind at the BZ site of $GABA_A$ receptor. Common structural features of the benzodiazepines, required for activity, are shown in orange and include an electron-withdrawing substituent on an aromatic ring, a carbonyl equivalent at C2, and a nitrogen atom at position 1. Benzodiazepine PAMs, such as diazepam, lorazepam, clonazepam, alprazolam, and midazolam, are sedatives; flumazenil is an antagonist used for the reversal of benzodiazepine effects. Zolpidem and ezsopiclone are used to treat insomnia and have a different profile based on binding to specific α subtypes.

Dating back to the 1990s, there were proposals that the NMDA receptor, an ionotropic glutamate receptor, was involved in depression, with NMDA antagonists having antidepressant effects. Glutamate is an excitatory neurotransmitter, and the balance of glutamate with the inhibitory neurotransmitter GABA has a role in depression. Most recently, two NMDA antagonists have been approved to treat depression. (S)-ketamine has been used as an anesthetic since the 1970s but has also been a drug of abuse due to its dissociative effects (Figure 13.7). However, at low doses, it is a rapid-acting antidepressant and has been approved as a nasal spray to treat depression. Dextromethorphan is another older drug that has common usage as a cough suppressant but, like ketamine, has been abused due to its dissociative effects. It was recently approved to treat depression, in combination with bupropion. Bupropion is a dopamine reuptake and norepinephrine reuptake inhibitor that is also an antidepressant, but it is also an inhibitor of CYP2D6, the major metabolizing enzyme for dextromethorphan, and its addition results in higher bioavailability of dextromethorphan.

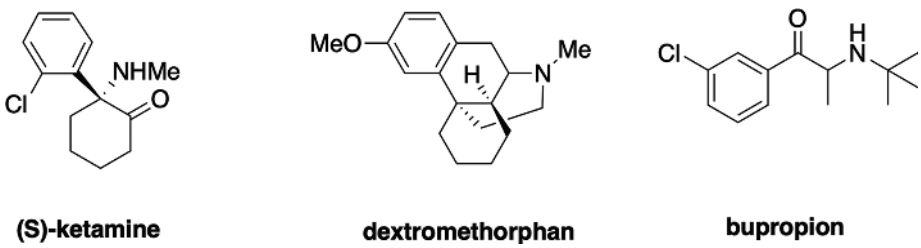


Figure 13.7 Structures of NMDA antagonists (S)-ketamine and dextromethorphan and bupropion.

13.2.2 Drugs That Increase Monoamine Neurotransmitter Levels or Act as Partial Agonists at Neurotransmitter Receptors Are Antidepressants

Depression is caused, in part, by an imbalance in levels of monoamine neurotransmitters, including serotonin, norepinephrine, and dopamine, in specific brain regions. Support for this hypothesis is based on observations that drugs that are effective in humans modulate brain concentration of these monoamines in animal models. The mechanisms of these drugs include monoamine oxidase (MAO) inhibition, inhibition of monoamine reuptake, and mimicking the action of these neurotransmitters. Because monoamine

oxidases are also involved in the metabolism of many drugs, MAO inhibitors are not widely used due to their potential to cause drug–drug interactions. Currently, the most common treatment for depression relies on increasing levels of serotonin or mimicking its effects. This may be accomplished by use of partial agonists or, more often, reuptake inhibitors. Since there are multiple subtypes of serotonin (5-HT) receptors (see Chapter 6), with effects on mood, appetite, energy, sexual function, vasodilation, and locomotion, selectivity for specific subtypes is desirable. Partial agonism at the 5-HT_{1A} receptor subtype is particularly effective in treating depression.

Selective serotonin reuptake inhibitors (SSRIs) are currently the most widely used drugs to treat depression. In the transmission of messages by nerve cells, neurotransmitters, such as serotonin, are released into the synaptic cleft and bind to receptors on the postsynaptic cell. As neurotransmitters are released from the receptors, they are taken up by monoamine transporters (see Section 6.5.1) back into the presynaptic cell in a process known as **reuptake** (refer to **Figure 13.1**). Reuptake inhibitors maintain concentrations of monoamine neurotransmitters in the synapse for a longer period of time. Early drugs in this category were the tricyclic antidepressants, and newer examples include the selective serotonin reuptake inhibitors (SSRIs).

The earliest **tricyclic antidepressant** approved was imipramine, developed in the 1950s at Geigy. Imipramine was synthesized as an analog of the antipsychotic drug chlorpromazine but was found to be effective as an antidepressant. The carbon analog amitriptyline and its metabolite, nortriptyline, were introduced by Merck in the early 1960s (**Figure 13.8a**). The tricyclic antidepressants all act as serotonin reuptake inhibitors, although they are not selective and also act as norepinephrine reuptake inhibitors. In fact, their antidepressant effects were originally thought to be due to their effects on norepinephrine. They also block adrenergic, muscarinic acetylcholine, and histamine receptors, leading to undesirable cardiac and sedative effects, and therefore, they have a narrow therapeutic index.

The concept of developing SSRIs was based on ideas that the mode of action of tricyclic antidepressants involved blocking the reuptake of serotonin. Arvid Carlsson and co-workers identified antihistamines, which also had serotonin and norepinephrine reuptake inhibition (termed SRI and NRI, respectively) properties, and began to optimize for serotonin reuptake inhibition activity. Among a group of pheniramine antihistamines, brompheniramine was identified as having a promising profile based on some selectivity for serotonin reuptake inhibition (**Figure 13.8b**). Modification of the nitrogen heterocycle and conformational constraint by the introduction of a rigid alkene led to zimelidine. Zimelidine is considered the first SSRI to be approved (1982, Europe) but was quickly withdrawn due to its association with Guillain–Barré syndrome.

The first SSRI to be introduced in the United States was fluoxetine (**Figure 13.8c**), and it has a considerably improved safety profile compared to zimelidine. The starting point for fluoxetine was diphenhydramine, a potent histamine antagonist with a similar structure to pheniramine that also had activity as a monoamine reuptake inhibitor. The success of fluoxetine led to the development of additional members of this class of drugs that have selectivity for the serotonin transporter compared to other transporters and receptors (see also supplemental Case Study in Chapter 5). Fluoxetine was first reported in 1974, approved in 1986, and marketed as Prozac. It became the most widely prescribed antidepressant, and within 10 years of its introduction, its sales reached $2.5 billion per year. As of 2025, there are five SSRIs approved for the treatment of depression. Other examples shown in **Figure 13.8c** include racemic citalopram, its single S-enantiomer escitalopram, sertraline, and paroxetine.

Citalopram's structure was based on an older antidepressant called talopram that worked through selective norepinephrine reuptake inhibition (SNRI) rather than as an SSRI. Talopram entered clinical trials, but development was stopped due to side effects, such as increased psychomotor activity. Like the benzodiazepines, talopram was the result of a synthesis that did not yield the expected product (**Figure 13.9**). In 1965, Klaus Bøgesø at Lundbeck was trying to synthesize an analog of tricyclic antidepressants, called melitracen, but the reaction followed an alternate pathway, and a phenylphthalane was produced instead. Unexpectedly, the alternate structure was also effective as an antidepressant and was a potent NRI. Modification of the structure led to talopram, which was a highly selective inhibitor of norepinephrine reuptake (SNRI). Converting the secondary amine to a tertiary amine, removing the geminal

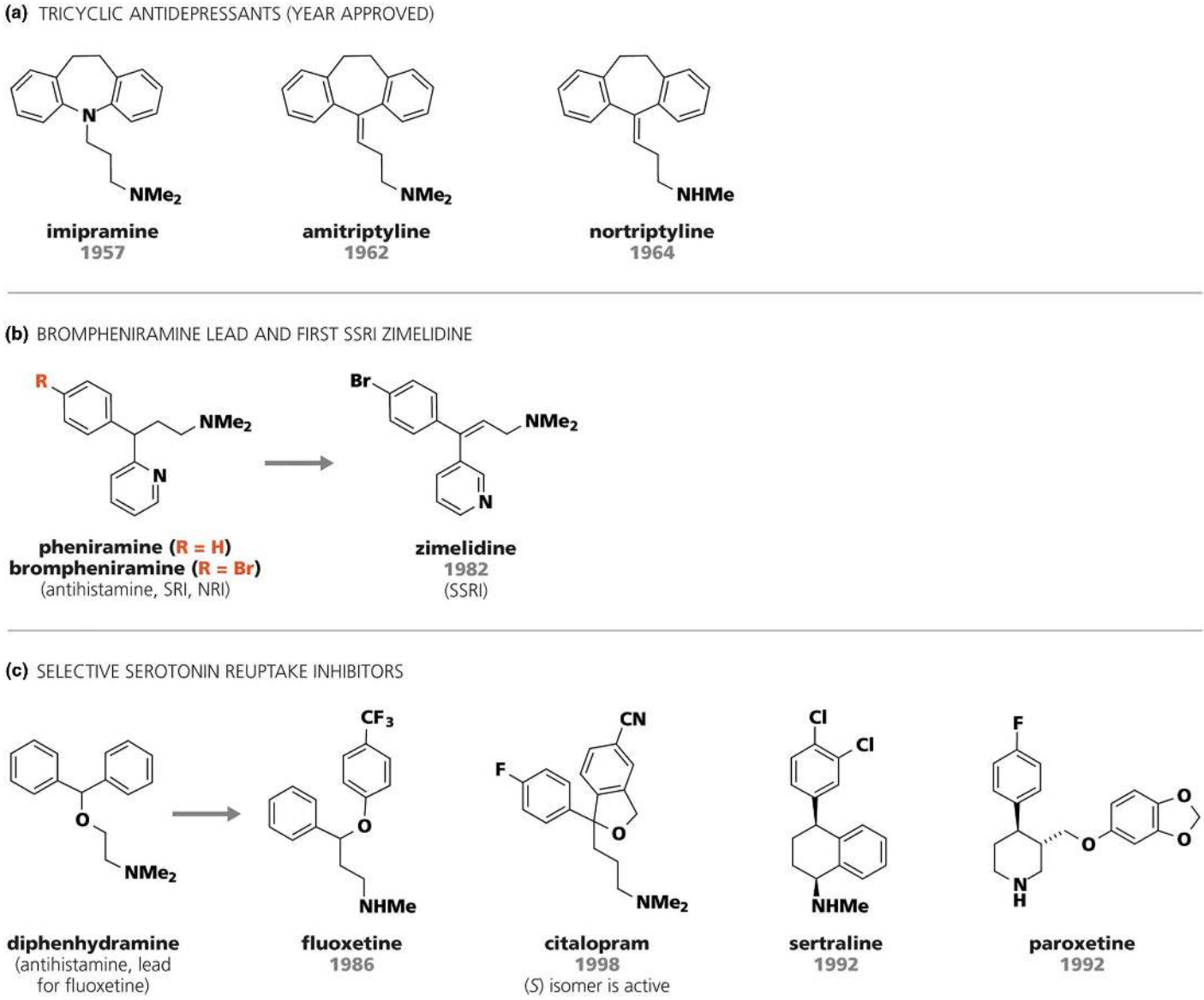


Figure 13.8 Examples of drugs used to treat depression that work through serotonin reuptake inhibition. (a) *Tricyclic antidepressants.* These include drugs such as imipramine, amitriptyline and nortriptyline, which are serotonin reuptake inhibitors but are not selective and exhibit significant side effects. (b) *Brompheniramine lead and first SSRI zimelidine.* Pheniramine and brompheniramine were the basis for the first SSRI introduced (in Europe), zimelidine. (c) *Selective serotonin reuptake inhibitors.* Drugs such as fluoxetine, racemic citalopram, its S-enantiomer escitalopram, sertraline, and paroxetine are relatively safe and are the most widely prescribed class of antidepressants. The antihistamine diphenhydramine served as the structural basis for fluoxetine.

di-methyl groups from the bicyclic structure, and introducing aryl substituents led to citalopram and changed the reuptake inhibition selectivity from norepinephrine to serotonin. While it was originally marketed as the racemate, the *S*-enantiomer of citalopram (escitalopram) is the active stereoisomer and is also approved.

Sertraline, initially reported in 1973, was synthesized by William Welch at Pfizer as part of a program to develop antipsychotics. A series of 1-aminotetralins based on the antipsychotic chlorpromazine was investigated, and while they were not active as antipsychotics, they did exhibit anti-anxiety effects. This observation led to the synthesis of a number of analogs, including the *trans*-aryl-substituted tametraline, which exhibited activity as both a norepinephrine and dopamine reuptake inhibitor (DRI). Interest in pursuing SSRIs in the early 1980s led Kenneth Koe at Pfizer to reinvestigate this series. Changing the ring substituent stereochemistry to the *cis* orientation and introducing aryl substituents ultimately led to sertraline.

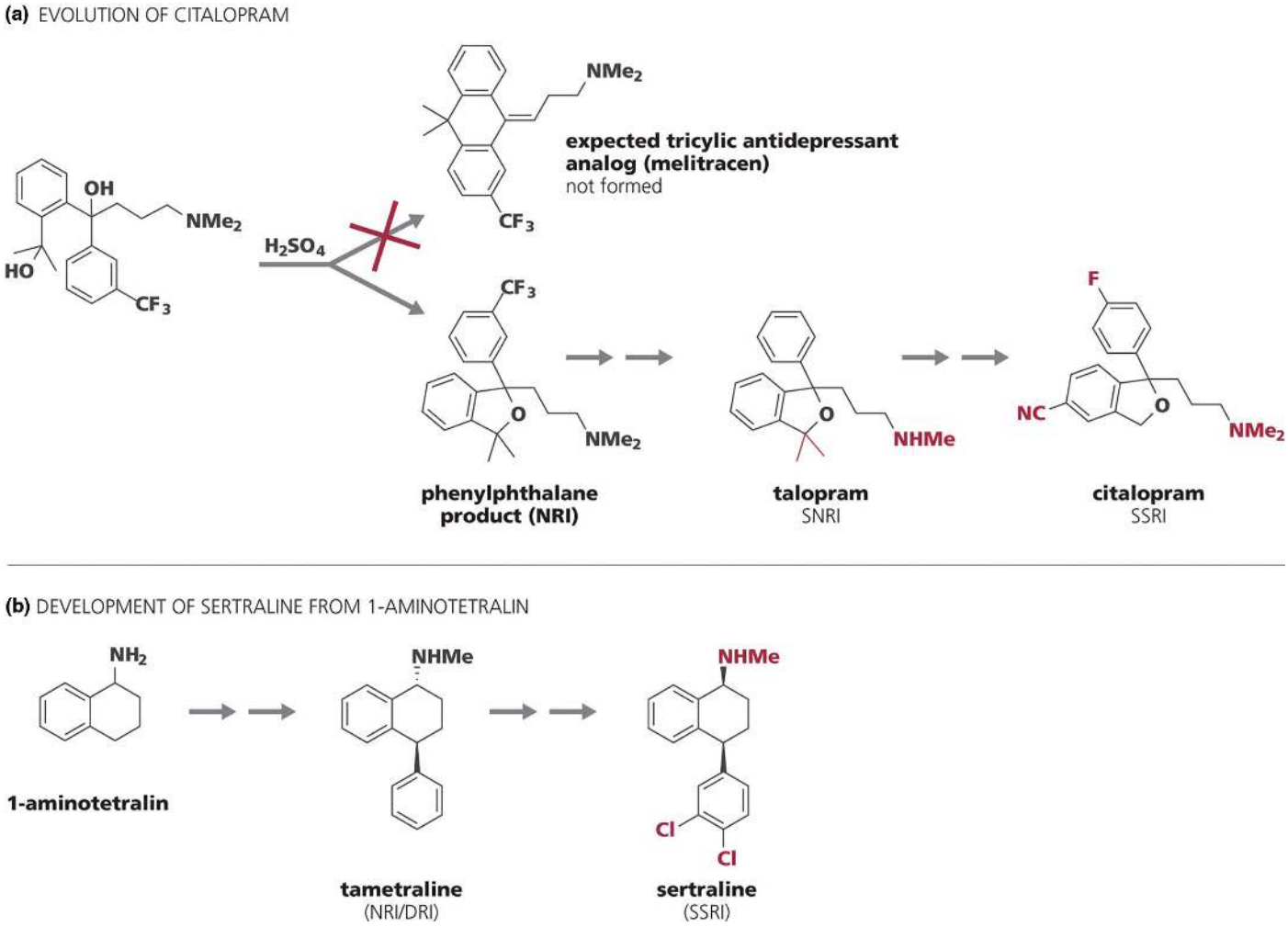


Figure 13.9 Development of citalopram and sertraline. (a) Evolution of citalopram. An attempt to synthesize melitracen, an analog of tricyclic antidepressants, via electrophilic aromatic substitution of the diol, resulted instead in the formation of the phenylphthalane cyclic ether. This scaffold also had antidepressant activity, and modifications (shown in red) led to the SNRI talopram and then to the SSRI citalopram. (b) Development of sertraline from 1-aminotetralin. Aminotetralins were found to have anti-anxiety effects. The parent 1-aminotetralin was modified to give the phenyl analog tametraline, the lead for sertraline. Changes included inversion of stereochemistry and incorporation of aryl substituents (shown in red).

Based on the importance of SSRIs as antidepressants, there have been studies to identify the required pharmacophore for binding. The target, the human serotonin transporter (hSERT), is a sodium neurotransmitter symporter (see Section 6.5.1). Prior to solution of the X-ray structure of (*S*)-citalopram bound to hSERT in 2016, homology models based on an X-ray structure of a related transporter, bacterial LeuT, were used to dock SSRIs and develop binding hypotheses. For example, one model of (*S*)-citalopram bound to hSERT based on the LeuT model is shown in **Figure 13.10**, along with an X-ray structure of (*S*)-citalopram bound to hSERT.

In this docking model, interactions in the binding pocket include a salt bridge between the amine on the drug to Asp98, π–π interactions of aryl rings with Phe335 and Tyr176, and hydrophobic interactions with Ala411, Tyr95, Ala173, and Ile172 (see **Figure 13.10a**). A pharmacophore model based on this binding pose includes a cationic site (P1) corresponding to the amine found in all SSRIs, as well as two aryl sites (R1, R2) and a hydrophobic area (H1). While the SSRIs in use were developed by use of traditional SAR to identify modifications providing selectivity, pharmacophore models may be used in virtual screening programs to identify novel SSRIs. A recent X-ray structure confirms that (*S*)-citalopram binds in the central cavity of SERT (see **Figure 13.10b**), with a slightly different orientation than the model in **Figure 13.10a**. Important interactions

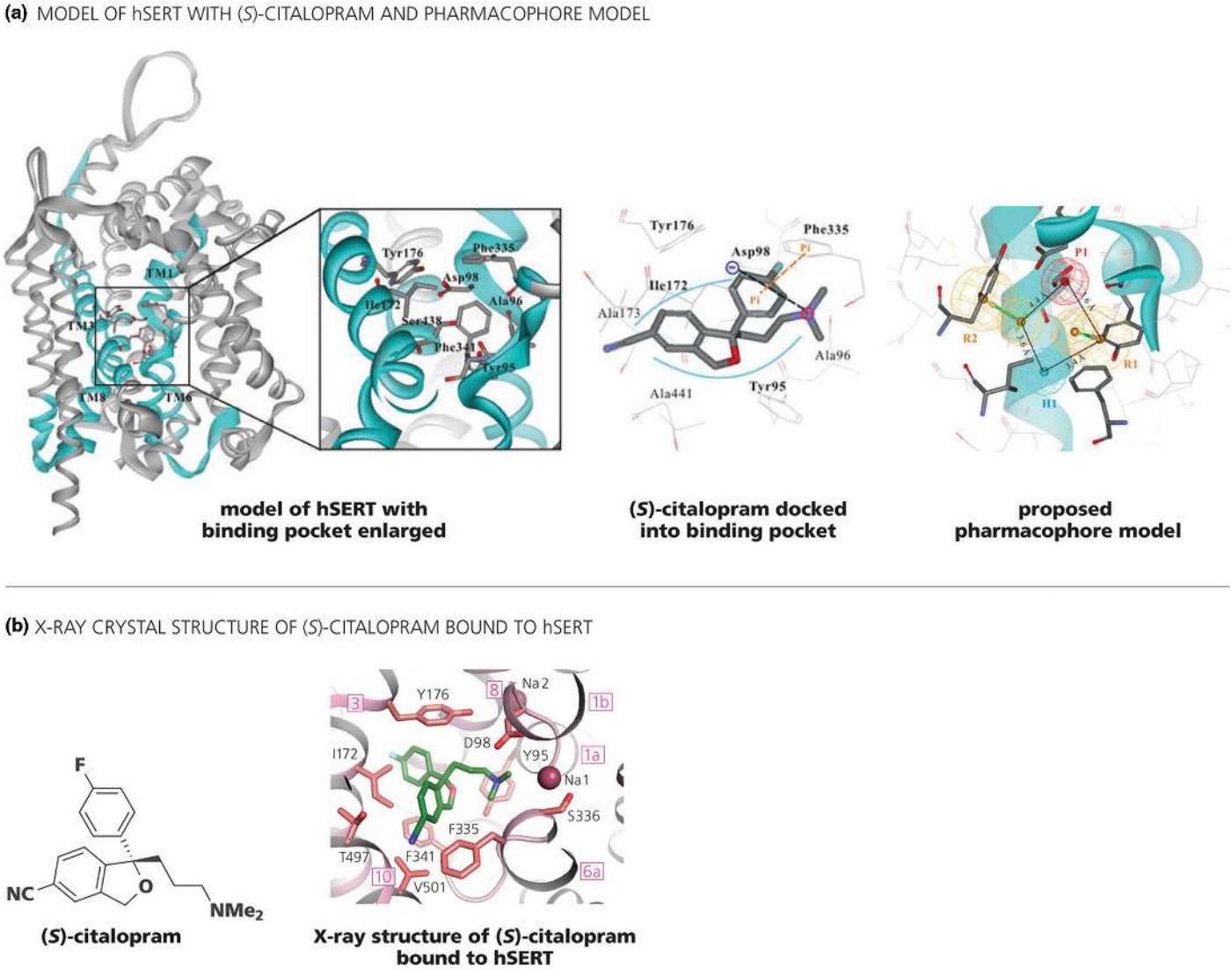


Figure 13.10 Homology and pharmacophore model of the human serotonin transporter and X-ray co-structure. (a) Models of hSERT with (S)-citalopram, and phamacophore model. *Left*, homology model of hSERT with binding site enlarged to show key side chains. *Middle*, (*S*)-citalopram docked to the homology model shows proposed interactions, including ion–ion interaction between the amine and Asp98 (black dotted line), π stacking of the aryl ring with Phe335 (gold dotted line), and a hydrophobic region (Ile172, Ala441, Tyr95, and Ala173; light blue lines) where the bicyclic portion binds. *Right*, the pharmacophore model is based on the binding model where R1, R2 (yellow) = aryl; H1(blue) = hydrophobic; and P1 (red) = cationic (model defined by distances between dots that are centered in pharmacophore grid points). (b) Solved X-ray structure of (*S*)-citalopram (green) bound to hSERT (red). Several proposed binding interactions are confirmed, including specific binding site and ionic interaction between citalopram nitrogen to Asp98. The aryl groups sit in a slightly different orientation than the model; important interactions observed include fluorophenyl to Tyr176, and bicyclic aryl to Phe335. Red balls = sodium ions. A second molecule of (*S*)-citalopram bound at an allosteric site is not shown.

Source: (a) Adapted from Zhou, Z.-L.; Liu, H.-L.; Wu, J. W.; et al. (2013). *Chem Biol Drug Des* 82:705–717. (b) From Coleman, J. A.; Green, E. M.; and Gouaux, E. (2016). *Nature* 532:334–339. With permission from Macmillan Publishers Ltd.

include nitrogen to Asp98 (as predicted), fluorophenyl to Tyr176, and bicyclic aryl to Phe335. The orientation of the aryl groups is opposite to what was predicted. In addition, the crystal structure shows a second molecule of citalopram binds at a nearby allosteric site, preventing the molecule in the central cavity from dissociating.

Reuptake inhibitors that target other neurotransmitters also have important clinical uses. Duloxetine and venlafaxine both have activity as combined serotonin and norepinephrine reuptake inhibitors, although they are more potent inhibitors of the serotonin transporter than the norepinephrine transporter. Both are used to treat both depression and anxiety. Reboxetine is a selective norepinephrine reuptake inhibitor

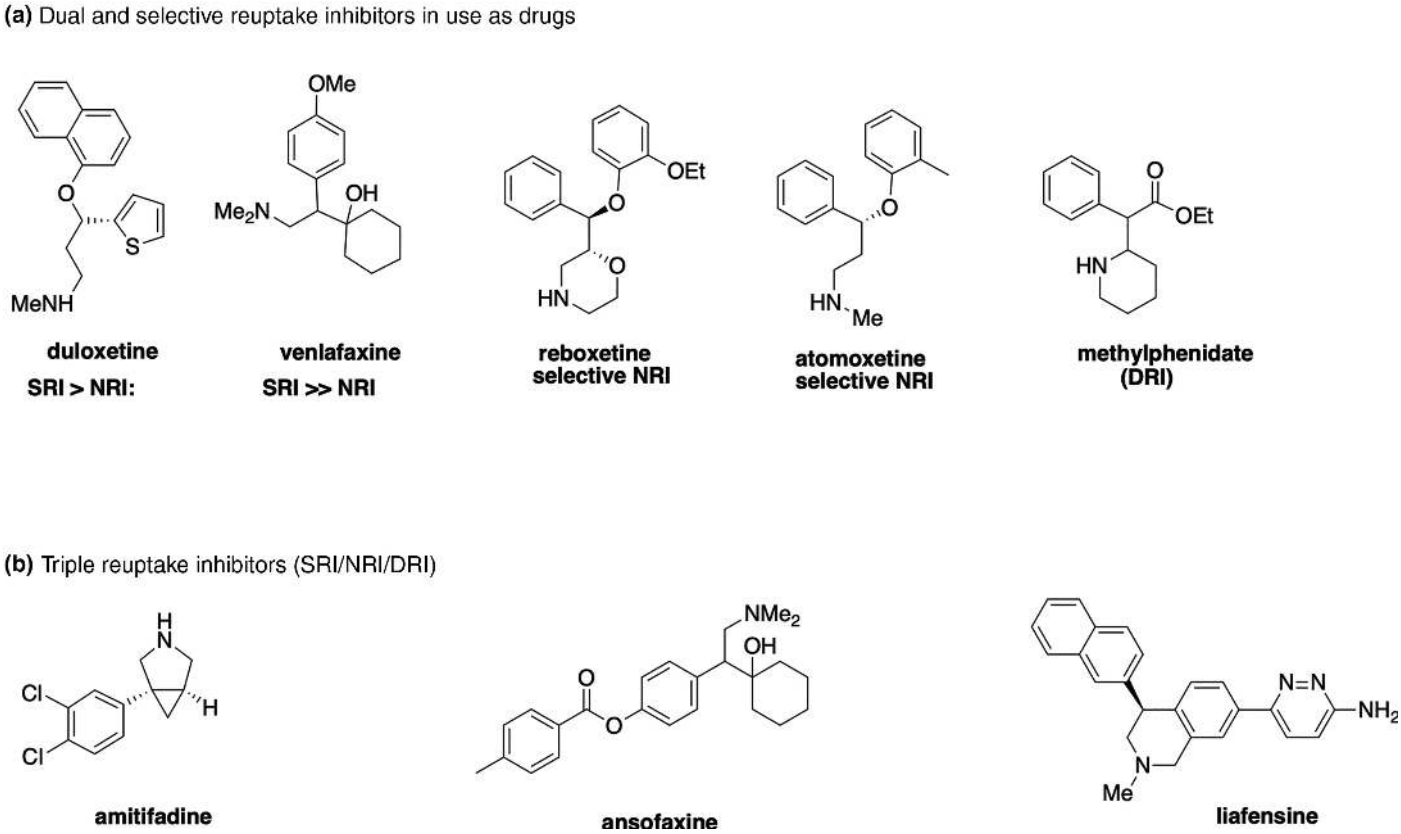


Figure 13.11 Structures of reuptake inhibitors targeting multiple neurotransmitters. (a) Dual and selective reuptake inhibitors in use as drugs. Duloxetine and venlafaxine are dual SRI/NRIs (predominant target indicated) used to treat depression and anxiety. Reboxetine (for depression) and atomoxetine (for ADHD) are selective NRIs. Methylphenidate, a dopamine reuptake inhibitor, is used to treat ADHD. (b) Triple reuptake inhibitors. Several triple reuptake inhibitors (SRI/NRI/DRI) have entered clinical trials to treat depression, including amitifadine (approved in China), ansofaxine, and liafensine (granted Fast Track status by the FDA in 2024).

(SNRI) developed in the 1990s and is used in Europe to treat depression. SNRIs, such as atomoxetine, and dopamine reuptake inhibitors (DRI), such as methylphenidate (Ritalin), can be used to treat attention deficit hyperactivity disorder (ADHD), although methylphenidate is not a selective inhibitor and has additional activity at the adrenergic receptor (Figure 13.11a).

Positive clinical results for use of SSRIs in combination with dopamine agonists have led to the idea of developing triple reuptake inhibitors (TRIs) that affect levels of serotonin, norepinephrine, and dopamine. Ansofaxine (toludexvenlafaxine) is an example of a TRI, which is approved for treatment of depression in China (Figure 13.11b). Others include amitifadine, which did not progress beyond Phase III trials, and liafensine, which received Fast Track designation from the FDA for treatment-resistant depression in 2024.

In addition to the neurotransmitter reuptake inhibitors, partial agonists at the 5-HT_{1A} receptor are also effective as antidepressants and anti-anxiety agents. **Buspirone** (Figure 13.12), an azapirone, is one such example. The 5-HT_{1A} receptor subtype is located

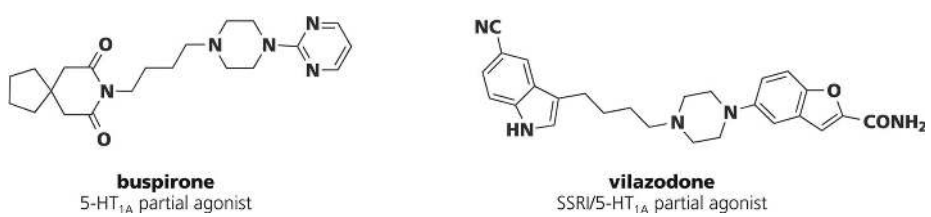


Figure 13.12 Serotonin partial agonists. Buspirone is a 5-HT_{1A} partial agonist, while vilazodone is both an SSRI and a 5-HT_{1A} partial agonist.

in the hippocampus and frontal cortex, which are areas of the brain involved in anxiety modulation. Buspirone is as effective as benzodiazepines, but with no addiction potential. Since buspirone does not target the GABA receptor, it is lacking in the sedative and anticonvulsant effects of benzodiazepines. It was originally prepared in 1969 during a search for tranquilizers devoid of the sedation side effects that characterized earlier drugs.

One shortcoming of SSRIs is a delay in onset of effectiveness, which may take 3–4 weeks. It has been suggested that this delay is because the increase in concentration in 5-HT in the synapse, resulting from inhibition of the transporter, leads to feedback inhibition of 5-HT_{1A} receptors, resulting in a decrease in release of 5-HT. After a few weeks of treatment, however, the 5-HT_{1A} receptors are desensitized, restoring release of serotonin. A recent approach has been the development of compounds with dual activity as 5-HT_{1A} partial agonists and serotonin reuptake inhibitors. This approach should circumvent the receptor sensitization that is believed to cause the delay in response. The first compound approved in this class is vilazodone (see **Figure 13.12**), which may have a shorter onset than other SSRIs and exhibits fewer side effects, such as sexual dysfunction (see Case Study 1 in this chapter).

As seen in these examples, a focus on neurotransmitter reuptake has provided many drugs that are effective in treating anxiety and depression. These have progressed from early nonselective drugs like the tricyclic antidepressants to SSRIs and SNRIs, as well as TRIs, which affect levels of serotonin, norepinephrine, and dopamine concurrently. Other indications for reuptake inhibitors include ADHD and insomnia. As these drugs have been used more widely, additional indications have also been approved, including use in sleep disorders, in the treatment of neuropathic pain, and in smoking cessation.

13.3 DRUGS TARGETING PSYCHOSIS

Antipsychotics are most widely used to treat schizophrenia, as well as other disorders characterized by delusions and hallucinations. *Schizophrenia* is a serious brain disorder characterized by an abnormal perception of reality that may include delusions, hallucinations, and disordered thinking and behavior. The disease typically manifests in people between the ages of 15 and 45, and it affects about 24 million people worldwide. Symptoms are characterized as positive and negative; positive symptoms include hallucinations and delusions, while negative symptoms include loss of drive, social withdrawal, and loss of cognitive function. Although a specific cause is unknown, schizophrenia is believed to be the consequence of a combination of genetic and environmental factors that result in an imbalance in certain neurotransmitter levels. There are multiple proposals on the physiological basis of schizophrenia. The dopamine hypothesis of schizophrenia is based on the observations that effective drugs are dopamine antagonists, as well as the fact that overuse of amphetamine (**Figure 13.13**),

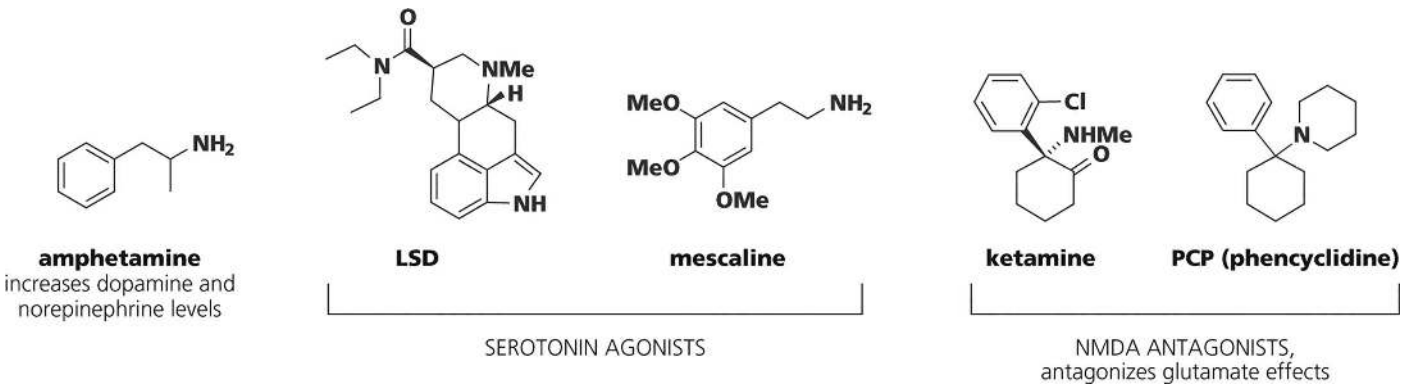


Figure 13.13 Compounds that can cause schizophrenia-like symptoms. Compounds with effects on dopamine, serotonin, and glutamate levels are known to cause schizophrenia-like symptoms. Amphetamine increases dopamine and norepinephrine levels, and overuse causes paranoid psychoses; LSD and mescaline are serotonin agonists and produce hallucinations; ketamine and PCP are NMDA antagonists, which antagonize glutamate effects and cause schizophrenia-like psychoses.

which causes dopamine and noradrenaline release, results in paranoid psychosis, similar to the positive symptoms exhibited by schizophrenics. In the brain, there are four major dopamine pathways: mesolimbic, mesocortical, nigrostriatal, and tuberoinfundibular. The dopamine hypothesis purports that overactive dopamine mesolimbic neurons in the striatal regions of the brain cause the positive symptoms of schizophrenia, while underactive mesocortical dopamine neurons in the prefrontal region cause the negative symptoms.

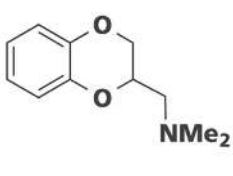
Another hypothesis, the serotonin hypothesis of schizophrenia, is based on evidence that serotonin agonists such as lysergic acid diethylamide (LSD) and mescaline produce hallucinations. Although these compounds are nonselective, agonism at the 5-HT_{2A} receptor is thought to be responsible for their hallucinogenic effects. The glutamate hypothesis proposes that insufficient glutamate levels contribute to disease symptoms. It has its basis in the observation of reduced levels of glutamate in the cerebrospinal fluid of patients, as well as higher levels of NMDA receptors in the brain of patients compared to unaffected individuals. In addition, the NMDA antagonists ketamine and phencyclidine (PCP), a drug of abuse, cause schizophrenia-like psychosis (see **Figure 13.13**). It has been suggested that neurodevelopmental abnormalities result in dysfunctional NMDA receptors and predispose those affected to the development of schizophrenia. NMDA receptors are involved in the regulation of dopaminergic neurons, providing an explanation for the observed changes in these neurons and a link to the dopamine hypothesis.

Treatment of schizophrenia prior to the 1950s involved convulsive therapy, partial lobotomy, and treatment with the plant alkaloid reserpine. Convulsive therapy included medical and electroconvulsive treatment as well as insulin injections to induce coma. Reserpine, from the Indian snakeroot plant, has been used in India for centuries to treat psychoses. However, it has serious side effects, including lowering blood pressure and decreasing heart rate. Beginning in the 1950s, several safer drugs were introduced to treat schizophrenia, and some have found use as treatments for other conditions, such as bipolar disorder.

13.3.1 Typical Antipsychotics Act as Dopamine Receptor Antagonists

Early drugs used to treat schizophrenia are termed **typical antipsychotics**; they act as antagonists at dopamine receptors. The five subtypes of dopamine receptors are all GPCRs and are divided into either the D_1-like family (including the D_1 and D_5 subtypes) or the D_2-like family (including D_2, D_3, and D_4 subtypes). Binding of dopamine at the D_1-like family of receptors activates inhibitory G proteins, while binding at the D_2-like subtypes activates stimulatory G proteins. Within the D_2-like receptor family, binding of dopamine at the D_2 receptor activates psychomotor pathways, and binding at the D_3 and D_4 receptors regulates emotion and cognition.

The first drugs used in treating psychosis associated with schizophrenia were the **phenothiazines**, characterized by the tricyclic ring system (red, in **Figure 13.14**). The prototype, chlorpromazine, was first synthesized in 1951 at Rhone-Poulenc in France. Its discovery can be traced to the benzodioxane antihistamines and the phenothiazine compound promethazine. Phenothiazine itself (not shown), first synthesized in 1883,



benzodioxane
(antihistamine)
structural basis for
chlorpromazine



promethazine
(antihistamine, sedative)
structural basis for
chlorpromazine



chlorpromazine
original typical
antipsychotic



fluphenazine (R = CF_3)
perphenazine (R = Cl)
more modern
typical antipsychotics

Figure 13.14 Structures of phenothiazines. Phenothiazines were synthesized as analogs of benzodioxane and promethazine, both antihistamines. The phenothiazine core is shown in red. Chlorpromazine is considered a typical antipsychotic; fluphenazine and perphenazine are more recent analogs. All three are used to treat positive symptoms of schizophrenia.

was used as an anti-helminthic drug but had no antipsychotic activity. However, its derivative, promethazine, had activity as both a sedative and antihistamine. Lengthening the alkyl chain, along with adding a chlorine atom to the aryl ring, led to chlorpromazine, which is both a D_2 and D_3 antagonist. Chlorpromazine was originally introduced early in 1952 to potentiate the effects of anesthesia, but due to its sedating effect, it was approved by the end of that year to control psychosis. As early as 1955, it was noted that its sedating effects were related to anti-serotonin effects, which are weak. However, in the 1960s, Arvid Carlsson demonstrated its anti-dopamine effects as the primary antipsychotic mechanism. He later won the Nobel Prize in Physiology or Medicine in 2000 for this discovery. The importance of dopamine effects led to searches for dopamine antagonists as new anti-schizophrenic agents. Chlorpromazine is effective on the positive symptoms of schizophrenia, but not the negative symptoms. Currently, there are several phenothiazines, with variations in the nitrogen and aryl substituents, that are used clinically, including fluphenazine and perphenazine (see Figure 13.14).

Although phenothiazines were the first effective antipsychotics to be approved, they are limited by their side effects. These include extrapyramidal symptoms (EPS), such as tremors, and muscle rigidity that are similar to symptoms of Parkinson's disease. Other side effects include sedation and weight gain. These side effects are so debilitating that they often limit patient compliance.

Haloperidol, approved in 1958 in Europe and in 1967 in the United States, is an example of the **butyrophenone** class of antipsychotics (Figure 13.15a). It was first synthesized at Janssen Pharmaceuticals in the 1950s as a derivative of the opiate meperidine (pethidine), as part of a program to identify analgesics (Figure 13.15b). Replacement of the *N*-methyl group of meperidine with a propiophenone moiety led to R-951, which exhibited improved analgesic activity. However, patent issues with the propiophenones led to synthesis of the butyrophenone homologs. Surprisingly, the homolog R-1187 was only a short-acting analgesic, but it had activity as a tranquilizer and cataleptic agent (causing rigidity), with an activity profile like that of chlorpromazine. Over 400 analogs of R-1187 were synthesized, including haloperidol.

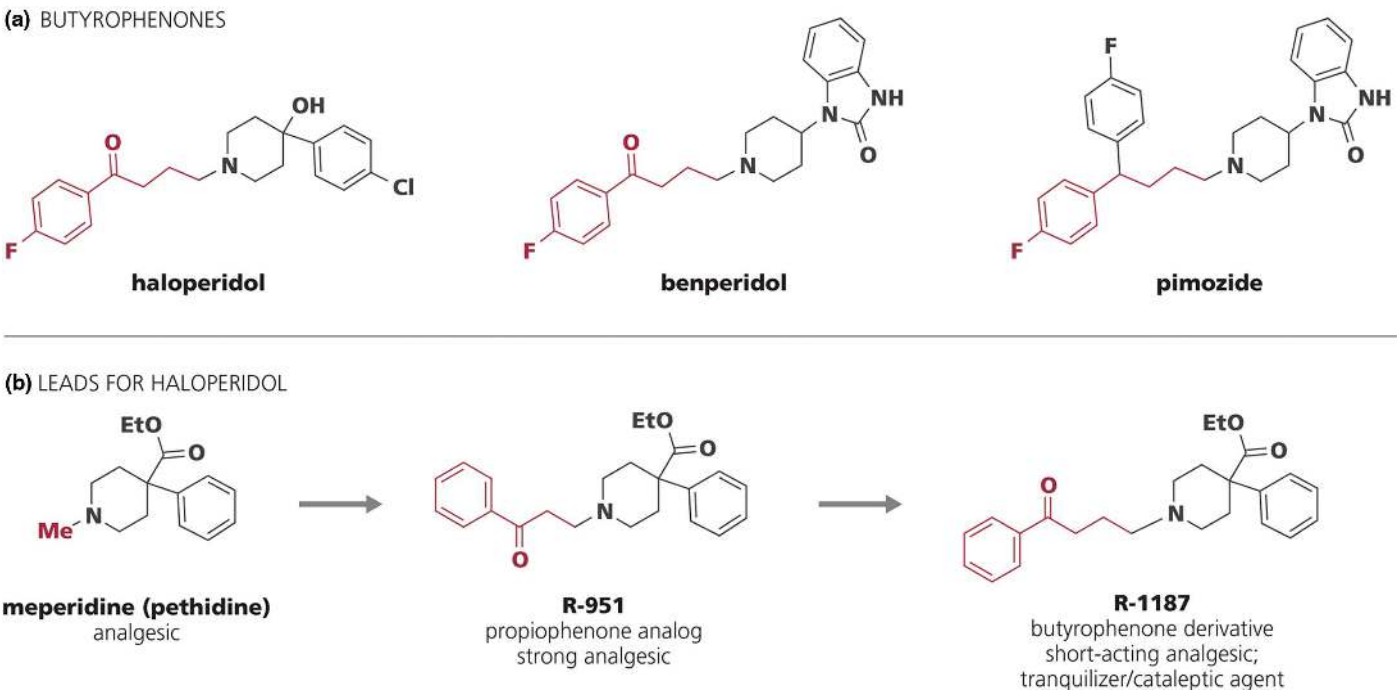


Figure 13.15 Development of the antipsychotic butyrophenones. (a) *Butyrophenones.* Members of the class of butyrophenones include haloperidol, benperidol, and pimozide (butyrophenone substructure/derivative in red). (b) *Leads for haloperidol.* The propiophenone R-951 was synthesized as an analog of meperidine (substitution shown in red). The butyrophenone homolog R-1187 was prepared to circumvent patent issues and was a tranquilizer and cataleptic agent and had a chlorpromazine-like profile rather than strong analgesic activity.

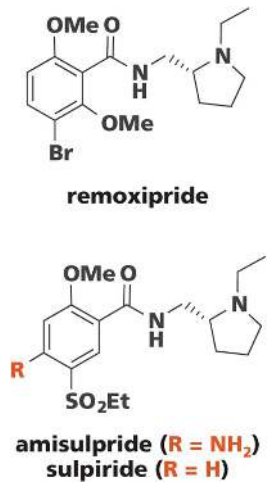


Figure 13.16 Structures of benzamide antipsychotics. Remoxipride (withdrawn), amisulpride, and sulpiride are D_2/D_3 antagonists and approved drugs.

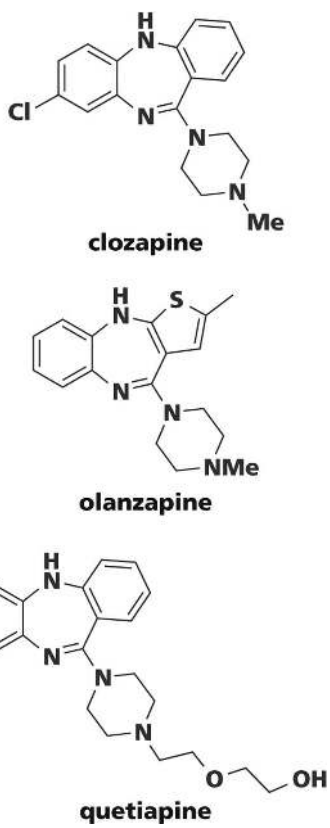


Figure 13.17 Structures of dibenzodiazepine atypical antipsychotics. Clozapine, olanzapine, and quetiapine are examples of atypical antipsychotics which exhibit a lower incidence of EPS than typical antipsychotics.

Although haloperidol is less sedating than chlorpromazine, patients still experience a high incidence of EPS. Despite this effect, several butyrophenones are used clinically, including haloperidol and benperidol. They act primarily as D_2 antagonists, with some activity at the 5-HT_2 receptor. Some other analogs, such as pimozide, have a fluorophenyl group in place of the carbonyl group. This compound, in addition to being used as an antipsychotic, is also used to treat severe symptoms of Tourette's syndrome.

Benzamides, such as remoxipride, act as D_2/D_3 antagonists and have a relatively low incidence of EPS due to some selectivity for the D_2 receptor. However, remoxipride was withdrawn several years after its introduction in 1990 due to occurrences of aplastic anemia (Figure 13.16). Other approved analogs include amisulpride and sulpiride.

The typical antipsychotics have largely been replaced as first-line therapy by the atypical antipsychotics, most of which were developed in the 1990s.

13.3.2 The Mechanism of Atypical Antipsychotics Involves Dopamine Receptor Antagonism Combined with Serotonin Receptor Antagonism

The **atypical antipsychotics** are characterized by a binding profile that includes stronger antagonism at D_4 than D_2 receptors, as well as 5-HT_{2A} antagonism. They are more effective than earlier antipsychotics against the negative symptoms of schizophrenia, as well as positive symptoms, and have a lower incidence of EPS.

Clozapine, a member of the class termed **dibenzodiazepines** developed at Sandoz in the 1990s, was the first atypical antipsychotic approved, in 1990 (Figure 13.17). It had originally been synthesized by Wander AG Pharmaceuticals and patented in 1960 as an antidepressant. One theory at that time was that the seven-membered central ring, which results in a bent rather than planar orientation of the two fused aryl rings, was necessary for antidepressant activity. This structural arrangement is present in tricyclic antidepressants, such as imipramine (see Figure 13.8). Clozapine was investigated by clinicians in Europe in the 1960s, who expected antidepressant activity. They found instead that it was effective as an antipsychotic and, surprisingly, had a low incidence of EPS. At that time, in 1966, it was believed that, in order to be effective, an antipsychotic needed to cause EPS, so interest in clozapine was low. In spite of this belief, clozapine entered clinical trials as an antipsychotic in the 1970s, and its efficacy was confirmed, but it was withdrawn due to serious side effects. Agranulocytosis was seen in about 1% of patients taking clozapine, and roughly half those patients died. It was reinvestigated and finally approved in 1990, but weekly monitoring of white blood cell count is required with its use. Later analogs, such as olanzapine (1996) and quetiapine (1997), do not exhibit the same side effect profile; quetiapine is currently the most widely prescribed drug for schizophrenia.

The binding profile of drugs related to clozapine, compared to the typical antipsychotics, is one explanation for the differences in their profiles. For example, clozapine binds more strongly to the D_4 receptor than to D_2, with K_i = 74 nM at the D_4 receptor and 254 nM at the D_2 receptor, whereas haloperidol has roughly equivalent binding at both (D_4, K_i = 7.3 nM; D_2, K_i = 4.8 nM). The atypical antipsychotics, such as clozapine, are also antagonists at the 5-HT_{2A} receptor.

Like the dibenzodiazepines, the **benzisoxazoles** are antagonists at both the D_2 and 5-HT_{2A} receptors (Figure 13.18). Risperidone was first synthesized at Janssen Pharmaceuticals in 1984 and was approved to treat both schizophrenia (in 1993) and bipolar disorder (in 2003). Its metabolite paliperidone is also used as an antipsychotic. Ziprasidone and lurasidone are in the same family; they contain a sulfur bioisostere in place of the oxygen of the benzisoxazole ring, as well as different bicyclic heterocycles.

Today, aripiprazole, brexpiprazole, and cariprazine are among the drugs most frequently used to treat schizophrenia. These phenylpiperidines are also considered atypical antipsychotics on the basis of their side effect profile, but they are D_2 and 5-HT_{1A} partial agonists as well as antagonists at the 5-HT_{2A} receptor.

As with the antidepressants that target neurotransmitter levels, antipsychotic drugs that have been developed to treat schizophrenia have also found other uses. For example, aripiprazole is used, in combination with other drugs, to treat depression, and many antipsychotics are useful in treating bipolar disorder. Aripiprazole is also used to treat autism, where it reduces some symptoms, such as stereotypy.

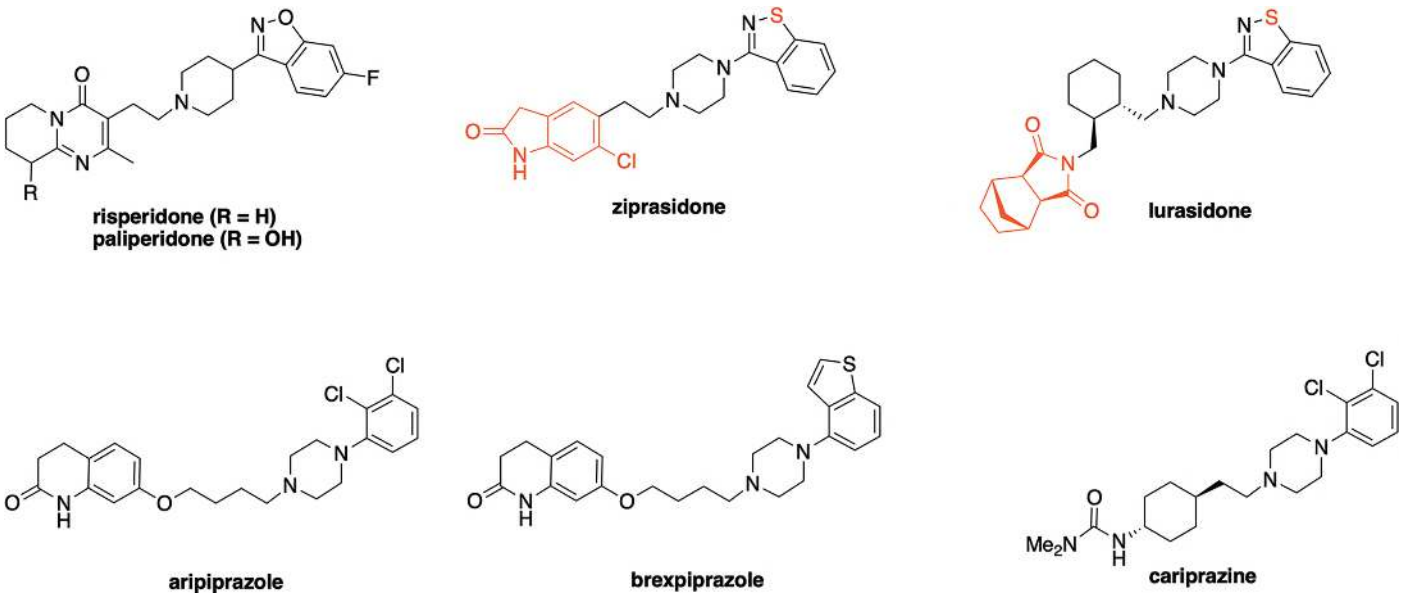


Figure 13.18 Benzisoxazole and phenylpiperidine atypical antipsychotics. The benzisoxazoles risperidone and its metabolite paliperidone, as well as analogs ziprasidone and lurasidone, are characterized by D_2 and 5-HT_{2A} receptor antagonism (orange indicates bioisosteres and modified heterocycles of ziprasidone and lurasidone). The phenylpiperidines aripiprazole, brexpiprazole, and cariprazine are D_2 and 5-HT_{1A} partial agonists as well as antagonists at the 5-HT_{2A} receptor.

13.3.3 Drugs Targeting Glutamate, Phosphodiesterase 10A, and the Muscarinic Acetylcholine Receptor Also Act as Antipsychotics

As mentioned previously, the molecular basis of schizophrenia is not completely understood, and there are multiple hypotheses that attempt to link dysregulation of levels of specific NTs to disease state. Current drugs affect serotonin and dopamine receptors; however, other hypotheses are also being pursued since ~30% of patients do not respond to current drugs, and few drugs have effects on the "negative" symptoms of the disease. The glutamate hypothesis is based on the idea that hypofunction in NMDA receptor-mediated neurotransmission is involved in schizophrenia. The NMDA receptor is a ligand-gated ion channel (see Section 6.3.1), with binding sites for both glutamate and glycine, both of which are required for opening of the channel. Glutamate also binds to the AMPA receptor, another ionotropic receptor. Several approaches have been investigated to increase glutamate levels, with the goal of treating schizophrenia, as well as other afflictions, such as obsessive–compulsive disorder (OCD). One approach has been to develop modulators of the metabotropic glutamate receptors (mGluR). These GPCRs are either postsynaptic ($mGlu_1$ and $mGlu_5$) and increase NMDA receptor activity or presynaptic ($mGlu_2$ and $mGlu_3$) and decrease NMDA receptor activity ($mGlu_{4,6,7,8}$ receptors have not been as widely targeted). **Figure 13.19a** shows several examples of modulators of mGlu receptors that have advanced to clinical trials, including the $mGlu_{2/3}$ agonist pomaglumetad, based on the idea of developing a conformationally constrained analog of glutamate (see Section 4.5.5). An amide prodrug, pomaglumetad methionil, developed at Eli Lilly, advanced to Phase III trials but was ineffective against the negative symptoms of schizophrenia. A similar drug still in trials is the $mGlu_2$ positive allosteric modulator MGS0274 (TS134), an ester prodrug of MGS0008. It is proposed that compounds targeting mGluRs would avoid the EPS seen with the marketed drugs that affect dopamine levels.

Another approach based on the glutamate hypothesis has led to the development of inhibitors of the glycine transporter GlyT1. This sodium transporter regulates glycine levels, which binds to and activates the NMDA receptor. High-throughput screening hits have led to several compounds that advanced to late clinical trials. These include bitopertin and iclepertin, which progressed to Phase III trials for both schizophrenia and OCD (see **Figure 13.19b**). While they were not sufficiently effective against the

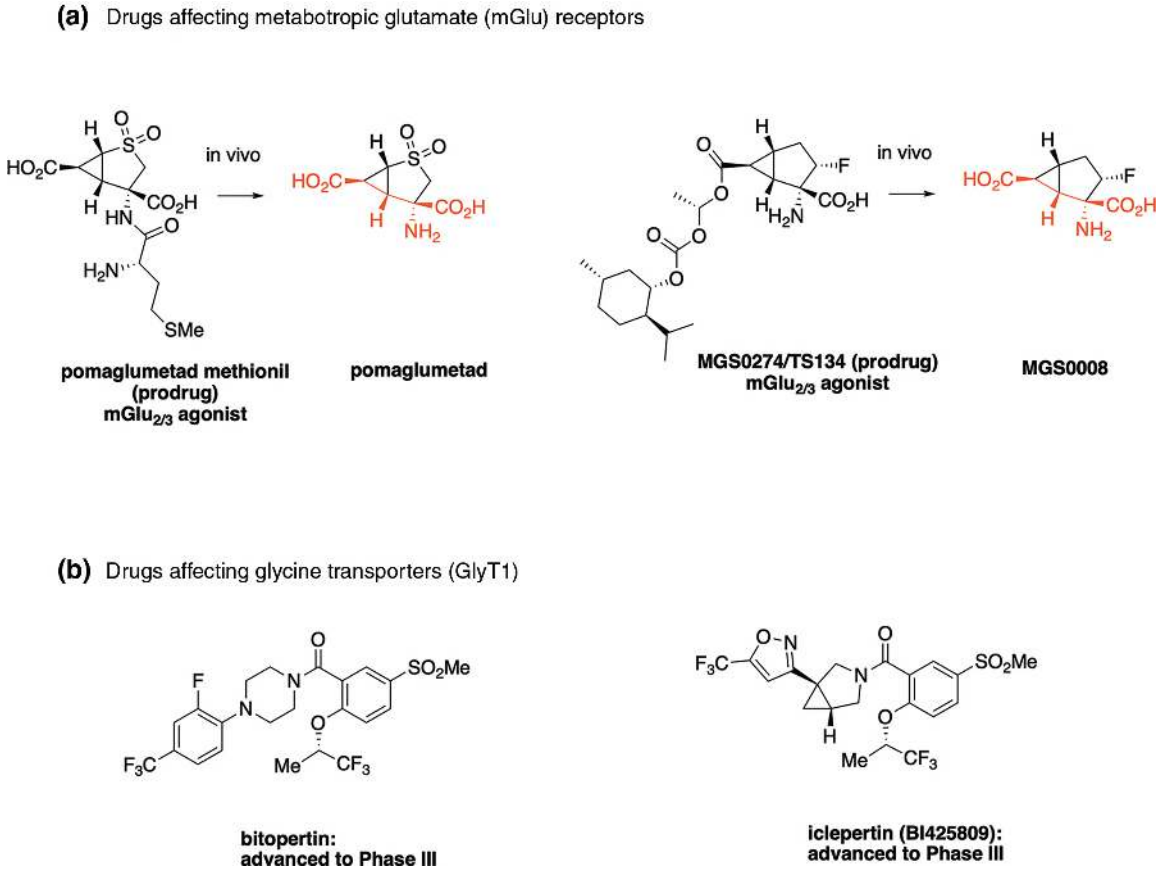
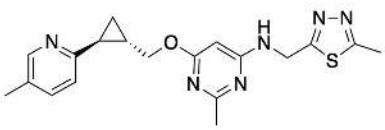


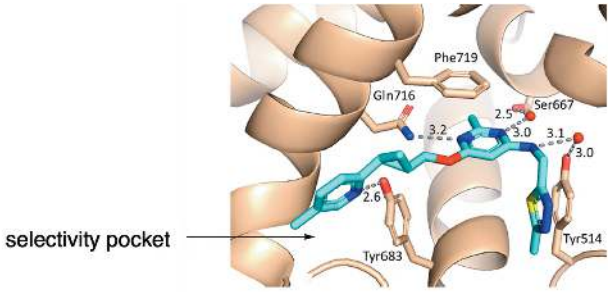
Figure 13.19 Examples of experimental compounds based on the glutamate hypothesis of schizophrenia. (a) Experimental compounds affecting the metabotropic glutamate (mGlu) receptors. Several compounds that act at mGlu receptors have advanced to clinical trials. Pomaglumetad, administered as a prodrug, whose design was based on conformational constraint of glutamate (shown in orange), is an $mGlu_{2/3}$ agonist, as is the similar prodrug MGS0274 (TS134). (b) Compounds affecting glycine receptors (GlyT1). Experimental drugs that affect glycine levels by inhibiting the GlyT1 transporter include bitopertin and iclepertin, which both advanced to Phase III trials.

negative symptoms of schizophrenia for further advancement, investigations into this approach continue.

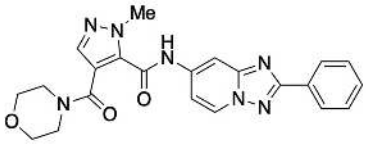
One novel strategy for treating schizophrenia that is distinct from directly affecting NT receptors is based on inhibiting phosphodiesterases. There are multiple forms of these enzymes, all of which hydrolyze phosphodiester bonds (see Chapter 7). Phosphodiesterase 10A (PDE10A) is primarily found in the striatal region of the brain, and abnormalities in this region are seen in patients with schizophrenia. PDE10A hydrolyzes cAMP and cGMP and regulates cAMP/PKC signaling and affects both glutamate and dopamine pathways. Since its discovery in 1999, many groups have been investigating selective inhibitors of PDE10A as a potential treatment for schizophrenia, as well as other disorders, such as Huntington's and Parkinson's diseases. Several advanced to clinical trials but were discontinued. However, as of 2025, MK-8189 looks promising as a treatment for schizophrenia and is in late Phase II clinical trials (**Figure 13.20**). It was based on a hit from fragment-based screening and has high selectivity for PDE10A. All phosphodiesterases have a conserved glutamine at the active site, which in most cases forms a hydrogen bond with known inhibitors. In PDE10A, the adjacent residue is a glycine, and its small size opens a selectivity pocket found only in PDE10A. Inhibitors with moieties that occupy this selectivity pocket form a hydrogen bond to a tyrosine (Tyr683 in the variant used for this X-ray), and bind to Gln716 by a cation–π interaction rather than an H-bond. In the case of MK-8189, the 5-methyl pyridine binds in the selectivity pocket, and its nitrogen acts as an H-bond acceptor to Tyr683. Other interactions include π stacking between Phe residues with the aryl rings Phe719 and Tyr514 (see **Figure 13.19**), and an H-bond between a pyrimidine nitrogen and a water molecule that bridges to the PDE10A backbone. Some PDE10A inhibitors show promise for other conditions, with CPL500036 in late Phase II trials to treat dyskinesia in Parkinson's patients, and NOE105 (gemlapodect) in trials to treat Tourette's syndrome.



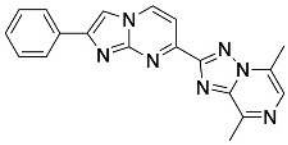
MK-8189



MK-8189 X-ray in PDE10A



NOE 105 (gemlapodect)



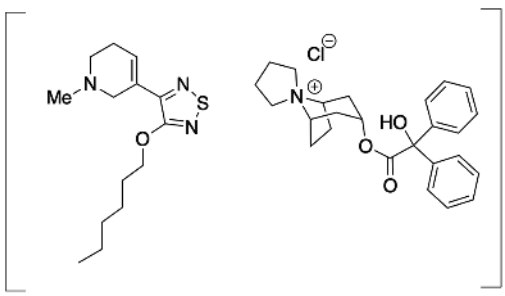
CPL 500036

The muscarinic acetylcholine receptors (mAChR) are also an emerging target, with agonists at these G protein–coupled receptors having activity against both the positive and the negative symptoms of schizophrenia, as well as benefits in Huntingdon's and Parkinson's diseases. There are five subtypes of muscarinic receptors (M1-M5), with agonists of M1, M4, and M5 showing efficacy. Xanomeline (Figure 13.21) is an agonist at M1 and M4; it had been tested in Alzheimer's patients and improved cognitive function, while also reducing hallucinations and agitation. This led to testing in schizophrenia patients, where it showed significant improvement, but also had muscarinic side effects, such as nausea, salivation, and sweating. Adding a broad acting muscarinic antagonist, trospium chloride, mitigated the side effects, and the combination of xanomeline and trospium chloride (KarXT or Cobenfy) is now approved for treatment of schizophrenia. Other muscarinic agonists in late development include the M4 selective emraclidine and NMI-1117568 (structure not available).

In summary, numerous advances in the treatment of schizophrenia have been made since the first drugs were introduced in the 1950s. The earliest drugs, the typical antipsychotics, are still in use but have mostly given way to the atypical antipsychotics, which have fewer side effects. These approved drugs all modulate levels of dopamine as well as serotonin. Newer approaches are focused in multiple areas and include experimental drugs to affect glutamate receptors, muscarinic acetyl choline receptors, as well as PDE10A.

Figure 13.20 Targeting phosphodiesterase 10A to treat schizophrenia and other diseases. Structures of PDE10A inhibitors MK-8189, NOE 105, and CPL 500036 that are in clinical trials as of 2025, along with an X-ray structure of MK-8189 (depicted in blue) bound to PDE10A, showing an H-bond between the pyridine nitrogen and Tyr683 in the selectivity pocket. MK-8189 is in trials to treat schizophrenia, NOE 105 to treat Tourette's, and CPL 500036 to treat dyskinesia in Parkinson's patients.

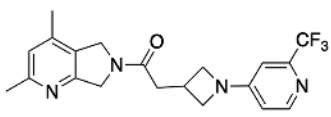
Source: X-ray from Layton, M. E.; Kern, J. C.; Hartingh, T. J.; et al. (2023). *J Med Chem* 66:1,157–1,171. With permission from the American Chemical Society.



xanomeline

trospium chloride

(KarXT)



emraclidine

Figure 13.21 Structures of mAChR agonists xanomeline and emraclidine, and muscarinic antagonist trospium chloride. The combination of xaomeline and trospium chloride (KarXT) is approved to treat schizophrenia.

13.4 DRUGS TARGETING PAIN

Various forms of pain affect 20% of the world population and lead to health-care and lost worktime costs of approximately $600 billion per year in the United States alone. Pain can be categorized as either **nociceptive**, resulting from injury to muscle or bone tissue, or **neuropathic**, resulting from injury to nerves. Nociceptive pain is typically acute and of shorter duration, whereas neuropathic pain is long-lasting and more resistant to treatment. Conditions that may cause neuropathic pain include osteoarthritis, fibromyalgia, diabetes, stroke, and certain chemotherapy drugs.

13.4.1 Drugs Used to Treat Pain Include Opiates and Nonsteroidal Anti-Inflammatory Drugs as Well as Antidepressants

The oldest treatments for pain are the opiates, such as morphine (see Section 1.3), which have been used for thousands of years. While effective against nociceptive pain, opiates are not as effective against neuropathic pain and exhibit dangerous side effects, including respiratory depression and addiction potential. Morphine and its semisynthetic analogs (differing mainly in oxidation states), such as oxycodone and hydrocodone, are agonists at the opiate receptors, which are GPCRs (**Figure 13.22**). Increasing steric bulk on the amine substituent leads to antagonists, such as naloxone, which is used to treat opiate overdose.

There are three different subtypes of the opioid receptor, μ, δ, and κ, with agonism at the μ receptor primarily responsible for reducing pain. It is believed that activity at the δ receptor mediates the addictive characteristics of opioids, and numerous efforts have been made to develop subtype-selective drugs. In the 1980s, the natural ligands for the opiate receptors were identified as the enkephalin peptides (see **Figure 13.22**). The structural feature essential for activating the receptor is the N-terminal Tyr-Gly-Gly-Phe sequence. Some of these can be mapped onto the opiates: Tyr1 corresponds to the aryl and amine moieties found in the opioids, as seen in red in **Figure 13.22**. Other analgesics that bind at the opiate receptors include the synthetic drug meperidine (or pethidine), synthesized in the late 1930s as an anti-cholinergic agent and later found

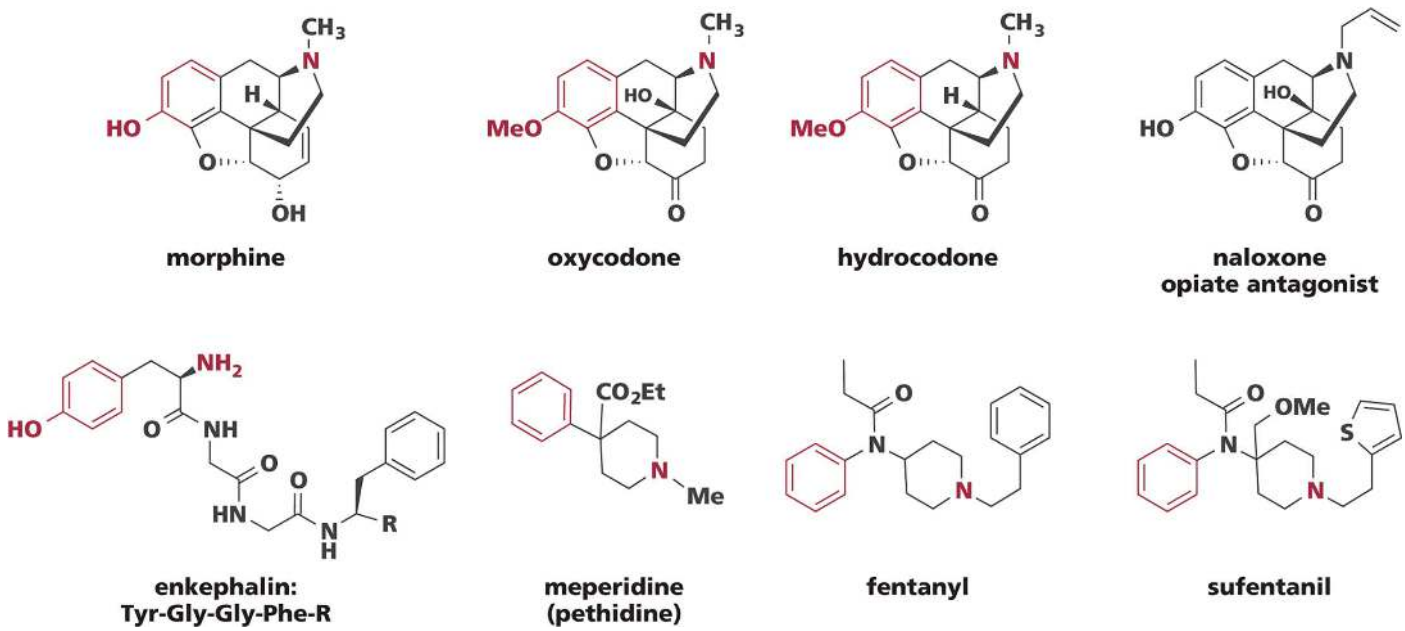


Figure 13.22 Structures of compounds that bind to opioid receptors. Morphine and the semisynthetic analogs oxycodone and hydrocodone are agonists at opiate receptors; the related analog, naloxone, is an antagonist used to treat opiate overdose. The Tyr-Gly-Gly-Phe fragment of the natural ligand enkephalin (R = Met or Leu) displays key pharmacophoric elements for activating the opioid receptor. Synthetic agents used to treat pain include meperidine, fentanyl, and sufentanil. The shared pharmacophore elements for agonists are shown in red.

to have analgesic activity. Fentanyl, synthesized by Paul Janssen in 1960 as an analog of meperidine, is 100× more potent than morphine, as is its analog, sufentanil. These analogs all maintain the aryl and nitrogen elements of the opioid pharmacophore.

Acetaminophen and the nonsteroidal anti-inflammatory drugs (NSAIDs), as well as muscle relaxants, are also used to treat nociceptive pain. The NSAIDs are inhibitors of cyclooxygenase and are discussed in Section 7.7. These include aspirin as well as the selective COX-2 inhibitors, such as celecoxib, which primarily act on COX-2 to reduce inflammation; these compounds work in the periphery rather than the brain.

One pain medication with a unique mechanism of action is ziconotide, which selectively targets the N-type calcium channel $Ca_v2.2$ and must be administered intrathecally (injected into spinal fluid). This 25-amino-acid peptide (ω-conotoxin) is produced by cone snails as one of a cocktail of peptides that are used to immobilize fish so that the snail can eat them. Investigation of these peptides by Baldomero Olivera at the University of Utah led to the discovery of the utility of ω-conotoxin, which is 1,000× more potent than morphine as an analgesic. The neuronal N-type calcium channel is a voltage-gated ion channel found on presynaptic terminals, and ziconotide blocks the entry of this channel. One model, shown in **Figure 13.23**, shows a molecule of ω-conotoxin (green) bound in the ion channel through an ionic interaction between an aspartate (D178) on the channel and a lysine (K2, blue) on the peptide toxin located just above the selectivity filter. An additional interaction between the N-terminal cysteine (C1, blue) and subunit IV glutamate (E177, red) is also evident. The large molecular weight and peptidic nature of this agent require it to be delivered intrathecally.

Neuropathic pain has been more difficult to treat than nociceptive pain. Drugs that have some utility include tricyclic antidepressants, such as amitriptyline (see **Figure 13.8**); monoamine reuptake inhibitors, such as duloxetine and venlafaxine (see **Figure 13.11**); and anticonvulsants, such as carbamazepine, lamotrigine, and phenytoin (**Figure 13.24**). These anticonvulsants inhibit pain by varying mechanisms. Both carbamazepine and lamotrigine act primarily by blockage of voltage-gated sodium channels, while phenytoin is a barbiturate, acting as a PAM at the $GABA_A$ chloride ion channel.

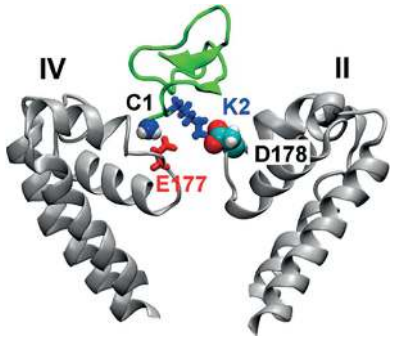


Figure 13.23 Model of ω-conotoxin (green and blue) blocking entry to the $Ca_v2.2$ ion channel (gray). The salt bridge between ω-conotoxin lysine (blue, K2) and subunit II aspartate (teal, D178) is shown. An additional interaction between the N-terminal cysteine (C1) and subunit IV glutamate (E177, red) is also shown.

Source: From Chen, R., and Chung, S.-H. (2013). *Biochemistry* 52:3,765–3,772. With permission from the American Chemical Society.



Figure 13.24 Structures of drugs targeting channels approved to treat pain. Carbamazepine and lamotrigine act primarily by blockage of voltage-gated sodium channels. Phenytoin is a barbiturate, acting as a PAM at the $GABA_A$ chloride ion channel.

13.4.2 Novel Approaches to Pain Treatment Include Selective Ion Channel Blockers

While pain accounts for 20% of all doctor visits and 10% of all drug sales, there have been few new pain medications in the last 20 years. This is despite the limitations of current drugs, including the abuse potential of the opioids and the relative lack of efficacy of drugs to treat neuropathic pain. Some newer areas of research have advanced compounds to clinical trials, but few novel compounds have received approval since 2004.

The voltage-gated sodium channels (Na_v) are the targets of the caine anesthetics, such as lidocaine and bupivacaine (**Figure 13.25**; see also Section 6.3.4). These are nonselective ligands that bind at a conserved local anesthetic binding site. There are nine subtypes of the Na_v channels, with $Na_v1.7$, 1.8, and 1.9 located primarily in the peripheral nervous system rather than the CNS. Drug discovery for entities selective for these targets has relied on screening and design based on known modulators. These efforts have led to the approval of first-in-class suzetrigine (Journavx) in early 2025 for the treatment of acute pain. Suzetrigine is selective for $Na_v1.8$ and shows

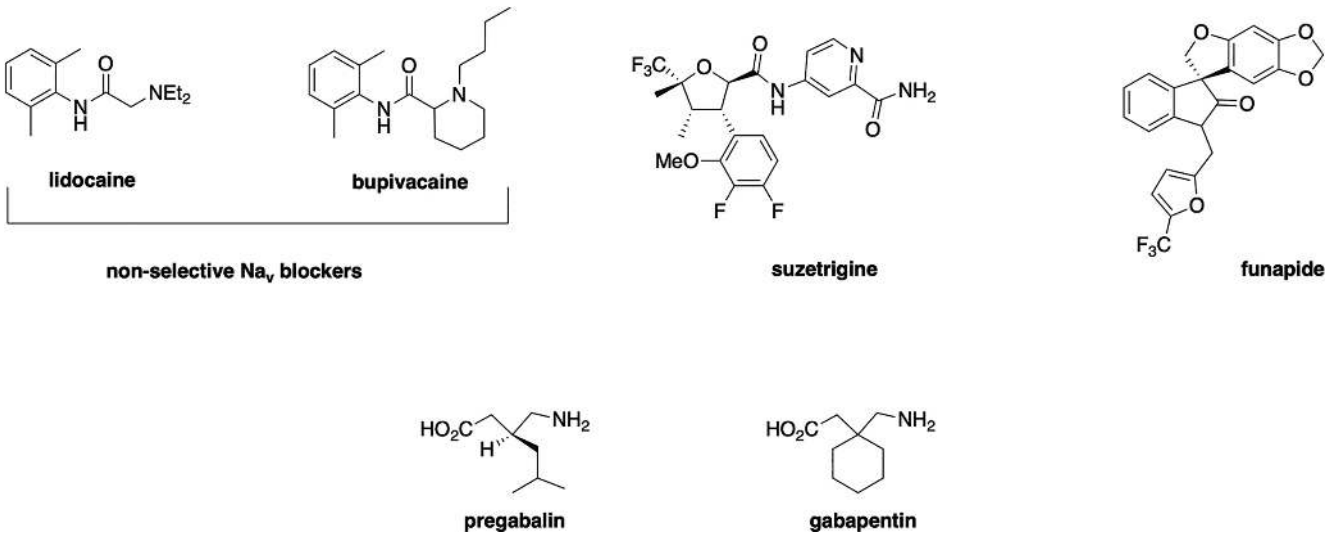


Figure 13.25 Structures of nonselective and selective sodium and calcium channel inhibitors used to treat pain. Lidocaine and bupivacaine bind indiscriminately to sodium channels. Suzetrigine is selective for the $Na_v1.8$ and is the first-in-class selective Na_v inhibitor approved. Funapide is an example of an experimental compound that has selectivity for the $Na_v1.7$ and $Na_v1.8$ channel. Pregabalin and gabapentin target the Ca_v channel.

no addiction potential, a vast improvement over opioids. Interest in subtype $Na_v1.7$ as a target for pain was based on the observation that individuals with no functional $Na_v1.7$ channel are insensitive to pain, while those with gain-of-function mutations in these channels suffer from severe chronic pain syndromes. There are several other compounds in clinical trials for neuropathic pain that are selective $Na_v1.7$ and $Na_v1.8$ blockers, such as funapide.

Pregabalin and gabapentin, are prescribed to treat pain, (and are also approved as anticonvulsants). They were originally developed as analogs of GABA but actually target the $\alpha_2\delta$ subunit of voltage-gated calcium channels, which are up-regulated in models of neuropathic pain. They bind to the Ca_v channel, which reduces Ca^{2+} influx into presynaptic terminals and thus indirectly blocks release of the excitatory NT glutamate as well as other NTs.

Treatments for pain have been available for thousands of years, with two major classes of drugs, opiates and NSAIDS, originally derived from natural sources. However, there are limitations to these drugs based on side effects and abuse potential; therefore, there is a continued effort to identify novel drugs, especially those to treat neuropathic pain and those acting through new mechanisms. Most of the older drugs have been discovered on the basis of effects found through *in vivo* assays. New approaches are based on genetic data that have identified new targets. With an understanding of the mechanisms responsible for pain, more selective and effective drugs may be developed.

13.4.3 Selective Serotonin Agonists Are Used to Treat Migraine

Migraines are a severe type of headache that can be triggered by hormonal changes, certain foods, stress, or environmental prompts, such as bright lights or loud noise. These headaches affect about 11% of adults in the United States, mostly women, and about 25% of cases are accompanied by visual auras. While the cause of migraines is not understood, they are proposed to result from dysfunction of neuronal circuits in the brainstem, related to the neuromodulator and vasoconstrictor calcitonin gene–related peptide (CGRP), which results in vasodilation in cranial vessels. While these headaches are sometimes alleviated by NSAIDs and other drugs, one line of therapy uses selective serotonin agonists that are believed to work via modulation of vasodilation and neurostabilizers, which reduce excitability of neuronal pathways. Recently, approaches targeting CGRP directly have been successful, with approval of antagonists such as ubrogepant (Chapter 5, Case Study 1) and atogepant and rimegepant.

The first drug approved to specifically treat migraine was the nonselective serotonin agonist ergotamine (a natural product); however, its side effects include vomiting, nausea, and increased blood pressure due to peripheral vasoconstriction (**Figure 13.26a**).

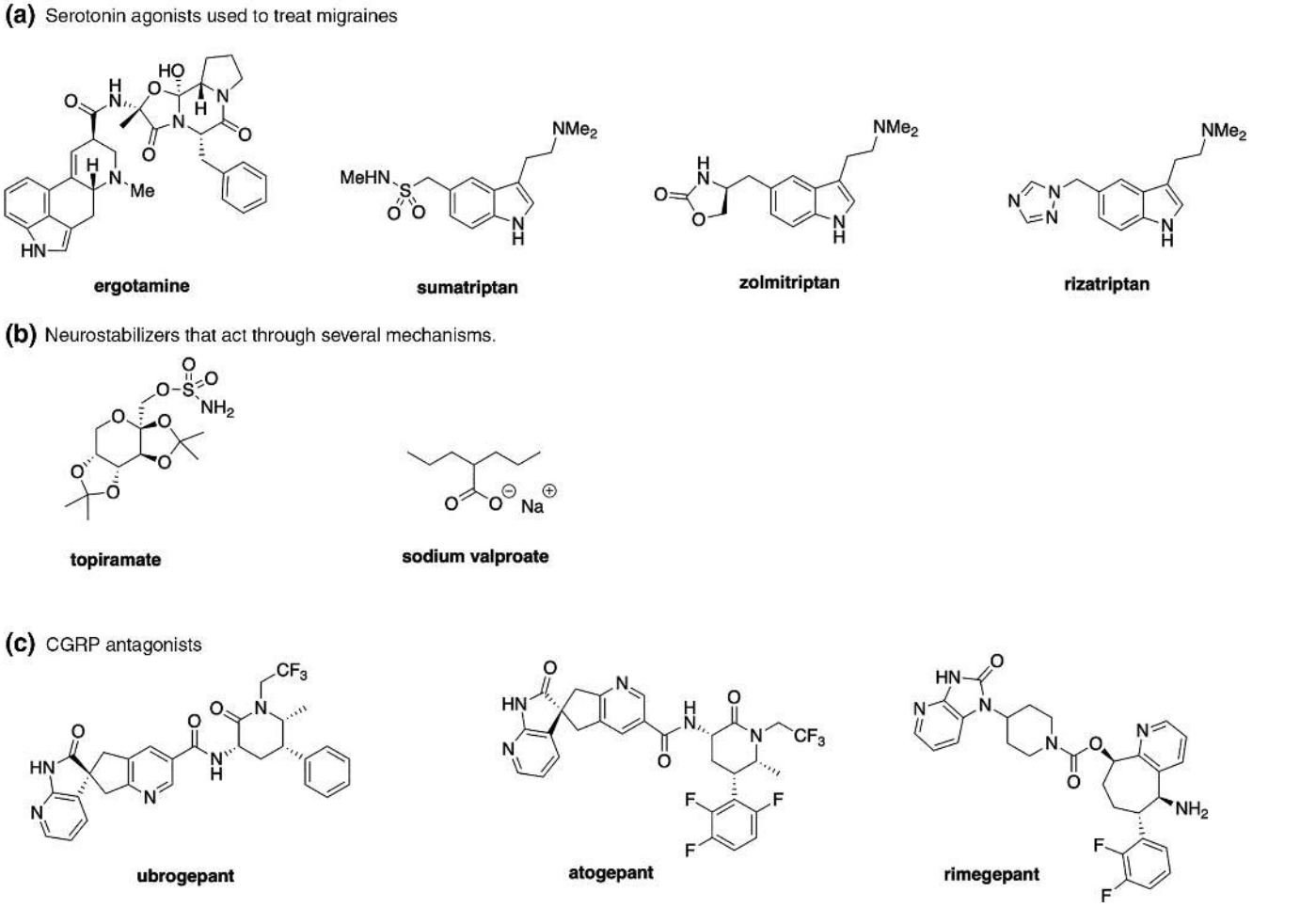
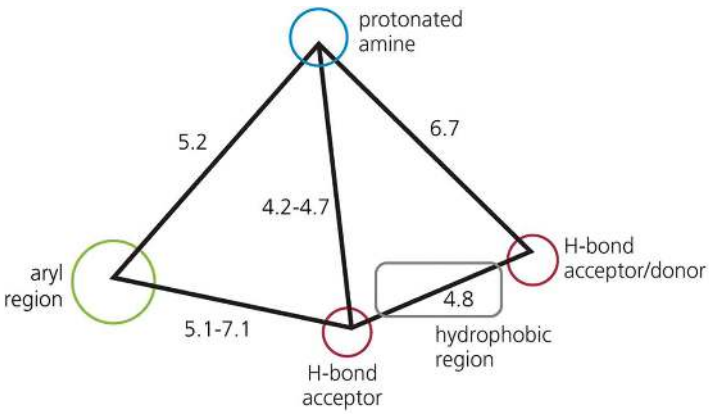


Figure 13.26 Structures of drugs used to treat migraines. (a) Serotonin agonists used to treat migraines. Ergotamine is a natural product that is nonselective. Sumatriptan is a selective 5-HT_{1B}/5-HT_{1D} agonist, as are zolmitriptan and rizatriptan. (b) Neurostabilizers that act through several mechanisms. Topiramate and valproate are considered neurostabilizers, lowering the excitability of neuronal pathways through multiple mechanisms. (c) CGRP receptor antagonists to treat migraine.

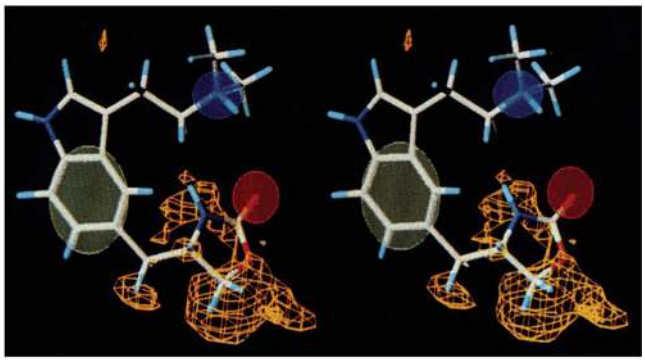
Sumatriptan is a selective 5-HT_{1B}/5-HT_{1D} agonist and acts by constricting blood vessels in the brain. While it had been known that migraines were associated with decreased levels of serotonin, the 5-HT_{1B} receptor subtype was not identified until the 1970s. This subtype is a GPCR located in the cranial vessels, and selective constriction of those vessels increases the rate of blood flow in the brain, providing migraine relief. Later analogs, such as zolmitriptan and rizatriptan, are more lipophilic and may have better BBB penetration.

Zolmitriptan was developed from a 5-HT_{1D} agonist pharmacophore model that was derived by the active analog approach (**Figure 13.27**). This model was based on data from a series of serotonin analogs that were evaluated for agonism at both 5-HT_{1D} (desired) and 5-HT_{2A} (undesired) receptors in order to generate a model that would predict binding to both subtypes. This pharmacophore exercise identified optimal regions for a protonated amine, an H-bond acceptor, an H-bond donor or acceptor and an aryl group. In addition, a hydrophobic region was identified between the H-bonding regions. Overlaying both selective and nonselective compounds identified a selectivity pocket in the hydrophobic region that was occupied by compounds with selectivity for 5-HT_{1D} subtype over the 5-HT_{2A} subtype. How zolmitriptan adheres to this pharmacophore is illustrated in **Figure 13.27**.

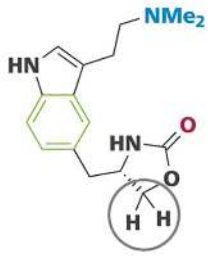
Several drugs that were originally marketed as anticonvulsants have been approved to prevent migraines. These include topiramate and valproate (see **Figure 13.26b**). Topiramate was originally synthesized in 1979 by Bruce Maryanoff at Johnson & Johnson as a chemical intermediate and became part of their legacy library. Testing in a mouse model showed potent anticonvulsant activity, and the compound entered preclinical testing in 1983. After it was approved as an anticonvulsant in 1996, subsequent clinical testing for other symptoms led to its additional approval to treat migraines. Topiramate has been found to operate through multiple mechanisms,



5–HT$_{1D}$ pharmacophore model
(numbers are distances in angstroms)



stereodiagram of zolmitriptan superimposed on model



zolmitriptan

Figure 13.27 Pharmacophore model for antagonists of the 5-HT$_{1D}$ receptor. The pharmacophore model on the left was generated from a series of serotonin analogs based on affinity at 5-HT$_{1D}$ (desired) and 5-HT$_{2A}$ (undesired) receptors. Specific elements with distance constraints are shown and include a protonated amine (blue), an aromatic region (green), H-bond donor/acceptor (red) and hydrophobic region (gray). Modeling of both selective and nonselective compounds onto the pharmacophore model identified the volume of a hydrophobic selectivity pocket that must be occupied for selectivity at 5-HT$_{1D}$ (orange, figure on right). This model was used to develop zolmitriptan, shown in a stereodiagram on the right, superimposed on part of the model. The amine (blue) fits the protonated amine region, the indole (green) fits the aryl region, and the carbonyl (red) fits the H-bond acceptor/donor region. The ring CH_2 fits into the selectivity pocket (yellow grid) in the hydrophobic region (gray in pharmacophore model).

Source: From Glen, R.; Martin, G. R.; Hill, A. P.; et al. (1995). *J Med Chem* 38:3,566–3,580. With permission from the American Chemical Society.

including blockage of the Na_v channel and acting as a PAM of the $GABA_A$ receptor at a site that is neither the benzodiazepine nor the barbiturate binding site, as well as several other mechanisms. Overall, it is classified as a "neurostabilizer," lowering excitability of several neuronal pathways. Another drug considered a neurostabilizer is valproate. Originally synthesized in 1881, it was marketed as an anticonvulsant in 1962 and is also effective in migraine prevention. Valproate has multiple mechanisms, effecting increased GABA synthesis as well as inhibiting glutamate function. As described in Chapter 5, Case Study 1, the GPCR antagonists of CGRP (e.g., ubrogepant and the related rimegepant) have recently been developed and represent a novel approach to treating migraine (**Figure 13.26c**). These orally bioavailable small molecules were derived from peptidic compounds, followed by extensive optimization.

13.5 DRUGS TARGETING NEURODEGENERATION

Neurodegenerative diseases are considered those resulting from the degeneration and death or loss of neurons. Many neurodegenerative diseases, such as Parkinson's and Alzheimer's diseases, are associated with aging. For example, 11% of Americans over 65 have been diagnosed with Alzheimer's disease, but that number rises to 45% for people over the age of 85. Other neurodegenerative diseases, like Huntington's disease, are due to congenital conditions or to autoimmune attack on the CNS, such as occurs in patients with multiple sclerosis (MS). These varied conditions, unfortunately, have few effective treatments.

13.5.1 Approved Small-Molecule Drugs to Treat Alzheimer's Disease Affect Acetylcholine or Glutamate Levels

Alzheimer's disease (AD) was first described by Alois Alzheimer in 1907 and is characterized by loss of memory and cognitive function, behavioral changes, and dementia. Loss of cholinergic neurons as well as the presence of amyloid plaques (aggregated β-amyloid protein) and neurofibrillary tangles (composed of phosphorylated tau protein in the brain) constitute a modern definition of the disease. A diagnosis of AD is based on a constellation of symptoms; however, until recently, ultimate diagnosis could not be made until autopsy. Current methods test for the presence of β-amyloid protein, with PET scans of the brain. Further complicating disease treatment is the fact that there are many forms of dementia, but not all forms result from AD.

While there is neither a cure nor a preventative for Alzheimer's disease, there are several small molecules that do not change disease progression but can improve cognition. Acetylcholinesterase inhibitors are one such class of approved drugs. The **cholinergic hypothesis** proposes that low levels of the neurotransmitter acetylcholine in AD patients contribute to cognitive decline. Levels of acetylcholine are regulated by its hydrolysis to choline, a reaction catalyzed by the enzyme acetylcholinesterase (AChE) (Figure 13.28). In this reaction, the catalytic Ser200, located at the bottom of a hydrophobic pocket, termed the *aromatic gorge*, reacts with the ester of acetylcholine to generate choline and the acylated enzyme. By inhibition of AChE, the concentration of acetylcholine is maintained, and cognitive function should improve.

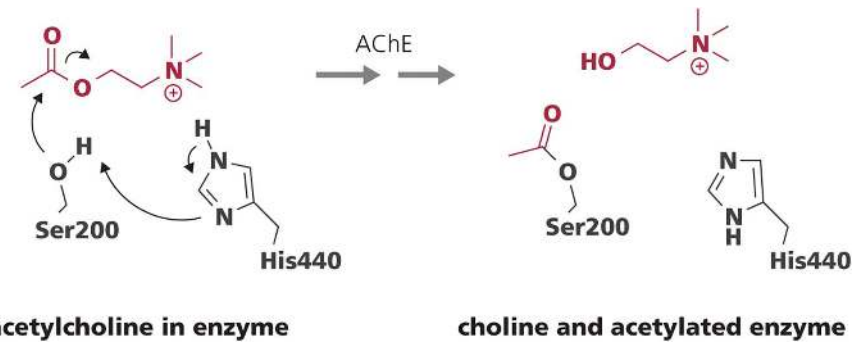


Figure 13.28 Abbreviated mechanism for hydrolysis of acetylcholine, catalyzed by acetylcholinesterase (AChE). Acid–base catalysis in the active site results in deprotonation of Ser200, which then reacts with the acetate carbonyl; cleavage of the ester bond generates choline and acetylated enzyme. For a more complete mechanism, see Figure 7.2.

The first drug approved to treat Alzheimer's disease was the AChE inhibitor tacrine (Figure 13.29). Adrian Albert originally discovered this aminoacridine in the 1940s during a search for antibacterial drugs. While tacrine was only weakly antibacterial, it was observed that it reversed anesthetic sleep and exhibited activity as an AChE inhibitor. It was used as an antidote for anti-cholinergic drug overdose from the 1950s through the 1970s but was not considered as a potential Alzheimer's disease treatment until 1981. A report by William K. Summers on its intravenous use in Alzheimer's patients described brief reversal of mild cognitive impairment. This observation ultimately led to its approval as an oral medication in 1993.

Tacrine is a noncompetitive reversible inhibitor of AChE. Due to limited oral availability and liver toxicity, it is no longer used; it has been replaced by donepezil, another noncompetitive reversible inhibitor. Donepezil was discovered by researchers at Eisai in the early 1980s while searching for an alternative to tacrine that had reduced toxicity. High-throughput screening identified an *N*-benzylpiperazine that exhibited *in vivo* activity and had AChE inhibitory activity in the rat brain. After synthesis of ~700 derivatives, donepezil emerged as the clinical candidate and received approval in 1997 (see Case Study 2 in this chapter).

The two other approved AChE inhibitors, galantamine and rivastigmine, are based on natural products and are competitive inhibitors of AChE. The lead for rivastigmine was physostigmine, an alkaloid found in the Calabar bean from western Africa. Ingestion of this poisonous plant was used by Indigenous tribes as a test for witchcraft, with death indicating the person was a witch. Physostigmine was tested as a treatment for Alzheimer's

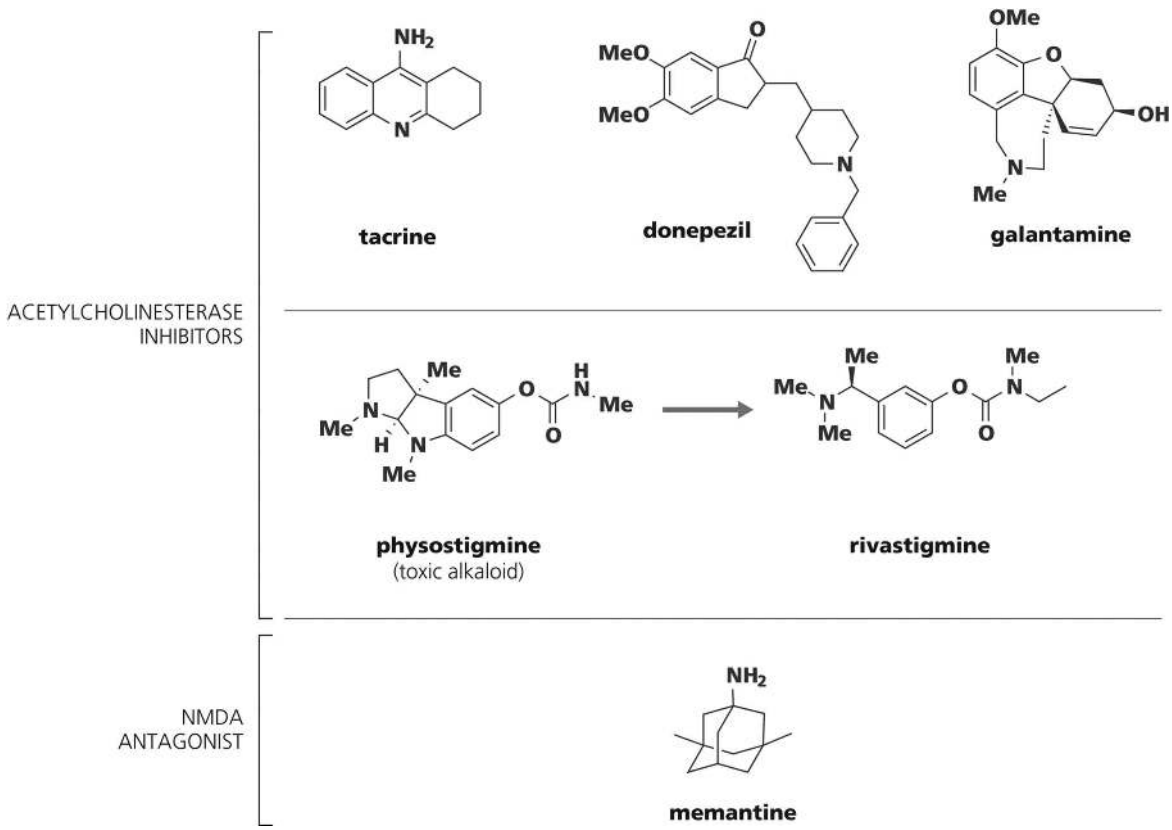


Figure 13.29 Structures of drugs used to treat Alzheimer's disease. The AChE inhibitors include the synthetic drugs tacrine and donepezil (noncompetitive), as well as the natural product galantamine and natural-product-derived rivastigmine, both competitive inhibitors. Memantine acts by a different mechanism, as an NMDA antagonist.

disease in the early 1980s, but its instability and toxicity limited its potential. The analog rivastigmine, developed at Novartis, maintains AChE inhibition and was approved in 1997. This compound is a competitive inhibitor of the enzyme. The carbamate carbonyl of rivastigmine reacts with the active site Ser200, releasing dimethylamino-ethylphenol (NAP). Figure 13.30 shows NAP and acylated enzyme bound in the active site.

The only other drug approved to treat Alzheimer's disease, memantine, does not work through the inhibition of AChE but is an NMDA receptor antagonist. It was originally synthesized in the 1960s as a possible anti-diabetic drug, and its activity as an NMDA antagonist was not discovered until the 1980s. The natural ligand for the NMDA receptor, glutamate, is excitatory, and excess activity leads to neurotoxicity. Memantine is a low-affinity noncompetitive antagonist with a fast off rate, and that combination leads to a low incidence of side effects.

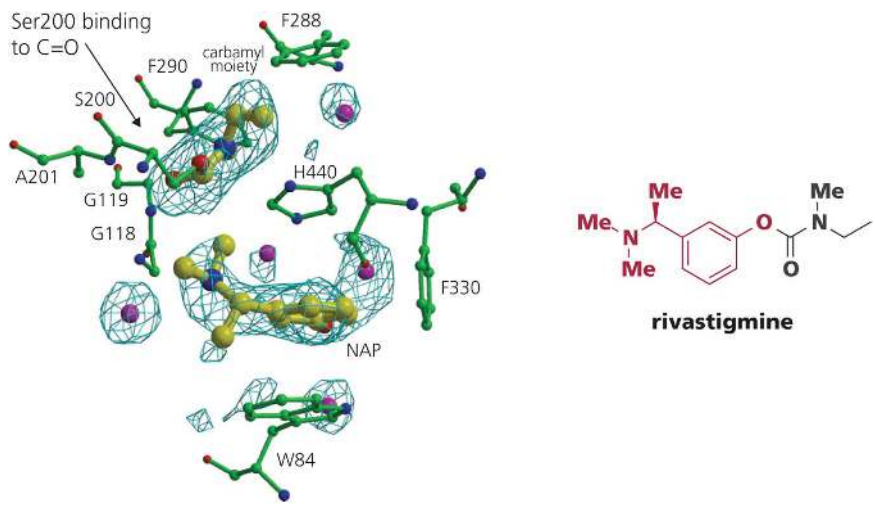


Figure 13.30 X-ray co-structure of AChE inhibitor rivastigmine bound to the enzyme AChE. Rivastigmine fragments are shown in yellow, and enzyme in green. The carbamate carbonyl of rivastigmine reacts with the active site Ser200, forming a covalent bond to the carbamyl moiety and releasing the dimethylamino-ethylphenol (NAP). The NAP residue shows π stacking with W84 and F330 (NAP portion colored red in rivastigmine structure on the right).

Source: From Bar-On, P.; Millard, C. B.; Harel, M.; et al. (2002). *Biochemistry* 41:3,555–3,564. With permission from the American Chemical Society.

13.5.2 Recent Approaches to Modifying the Course of Alzheimer's Disease with Small Molecules Have Focused on Decreasing Production of Amyloid Plaques

While AChE inhibitors and NMDA antagonists are approved to treat AD, they have no effect on changing the course of the disease. Therefore, research over the past 20 years has focused on alternative approaches. One is based on the **amyloid hypothesis**, which suggests that inhibiting production of the β-amyloid plaques found in the brains of Alzheimer's patients would modify the course of the disease. It is believed that these plaques precede the onset of dementia, and either they or their precursors are neurotoxic. Evidence for the importance of this pathway is based on the observation that some forms of early-onset AD occur in people with mutations in the genes that code for the protein (amyloid precursor protein, APP) that ultimately generates β-amyloid plaques. The plaques are formed by aggregation of amyloid β_{42} protein (**Figure 13.31**), which is produced by proteolytic cleavage of APP by sequential enzyme processing. Recently, antibody drugs (e.g., donanemab and lecanemab) that target β-amyloid have been approved for use at the earliest stages of disease, further supporting the validity of this approach.

APP is a membrane-bound substrate whose normal function is not well understood. The membrane-associated aspartyl protease β-secretase 1 (BACE1) cleaves the extracellular N-terminal portion of APP in the rate-limiting step to generate a soluble fragment (sAPPβ) and a membrane-bound fragment, C99. BACE1 was isolated in 1999 and is primarily found in neurons (another isoform, BACE2, is found in the periphery). The BACE1-catalyzed reaction is followed by cleavage of the intracellular portion catalyzed by a second protease, γ-secretase. This reaction generates small amyloid β peptides from 39 to 49 amino acids, including amyloid β_{42}, which is highly prone to aggregation and forms amyloid plaques. Approaches for small-molecule AD treatments have targeted both γ-secretase and BACE inhibition.

γ-Secretase, also a membrane-bound protease, consists of a multidomain protein that was difficult to isolate. Most approaches to identify inhibitors utilized high-throughput screening with phenotypic assays for hit identification and optimization. While a number of compounds advanced into clinical trials, most fell out for a variety of reasons, including lack of efficacy, toxicity, or poor pharmacokinetics.

Targeting BACE was facilitated by the vast knowledge acquired while developing other aspartyl protease inhibitors, most notably the HIV protease inhibitors used to treat AIDS. Several design approaches have been pursued in the development of BACE inhibitors, with a number of compounds progressing to clinical trials. However, while levels of amyloid β_{42} were lowered in some cases, clinical improvement in Alzheimer's symptoms using small-molecule drugs has not yet been realized. Most early examples that

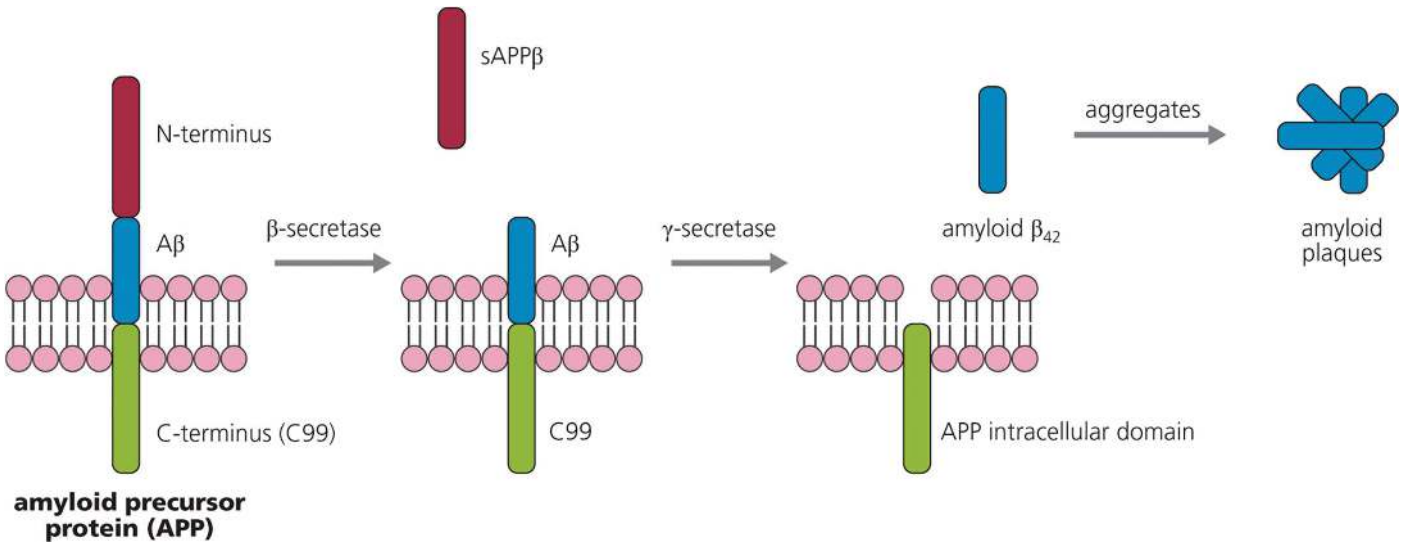


Figure 13.31 Sequential enzymatic cleavage of APP to provide amyloid β. The membrane-spanning amyloid precursor protein (APP) is cleaved on the extracellular N-terminal side by the enzyme β-secretase to release a soluble APPβ fragment (sAPPβ) (shown in red). Cleavage of the remaining C-terminal fragment (C99) by γ-secretase generates free amyloid β (blue), of 39–49 amino acid lengths, and the APP intracellular domain (green). The 42–amino acid amyloid β_{42} goes on to aggregate and form amyloid plaques.

advanced to clinical trials were peptidomimetics, such as CTS-21166 (Figure 13.32a), that were optimized by structure-based drug design. This compound, developed by Arun Ghosh, at Purdue University, inhibits BACE by binding at the active site, making key interactions with Asp32 and Asp228 (see Figure 13.32a). However, in clinical trials, while levels of amyloid β were reduced in plasma and the compound was safe, there was little efficacy based on cognition and daily function. In addition, the peptidomimetics had poor brain penetration due to their high surface area, flexibility, and polarity, and they were removed by efflux via P-glycoprotein (P-gp).

The sources of other BACE inhibitors were discovered through high-throughput and fragment-based screening programs and optimized by structure-based design. Examples that have advanced to clinical trials are shown in Figure 13.32b. These compounds have good brain penetration and result in lower levels of amyloid β in the cerebrospinal fluid. While several, such as LY2811376, LY2886721, and AZD3839, did not progress beyond Phase II trials, a few, such as verubecestat (formerly MK-8931) progressed to Phase III trials but unfortunately were dropped due to lack of efficacy in patients with advanced AD. Several are still under investigation in patients with early stages of AD. While no small-molecule drugs targeting β-amyloid have been approved, several monoclonal antibodies received FDA approval in 2024. Aducanumab, lecanemab, and donanemab all target removal of β-amyloid, and have shown approximately 30% reduction in cognitive decline for patients in early stages of Alzheimer's disease.

The complex nature of AD makes drug discovery particularly challenging. Novel approaches using small molecules beyond those directly involved in β-amyloid generation include those targeting tau, agonists of the nuclear receptor PPARγ, and agonists of sigma-1 receptors, among many others.

13.5.3 Drugs Developed to Treat Parkinson's Disease Are Based on Increasing Dopamine Levels

Parkinson's disease (PD) is characterized by the loss of dopaminergic neurons and reduced dopamine levels in the part of the brain called the substantia nigra. This damage results in movement disorders, including tremors and rigidity. An additional marker for

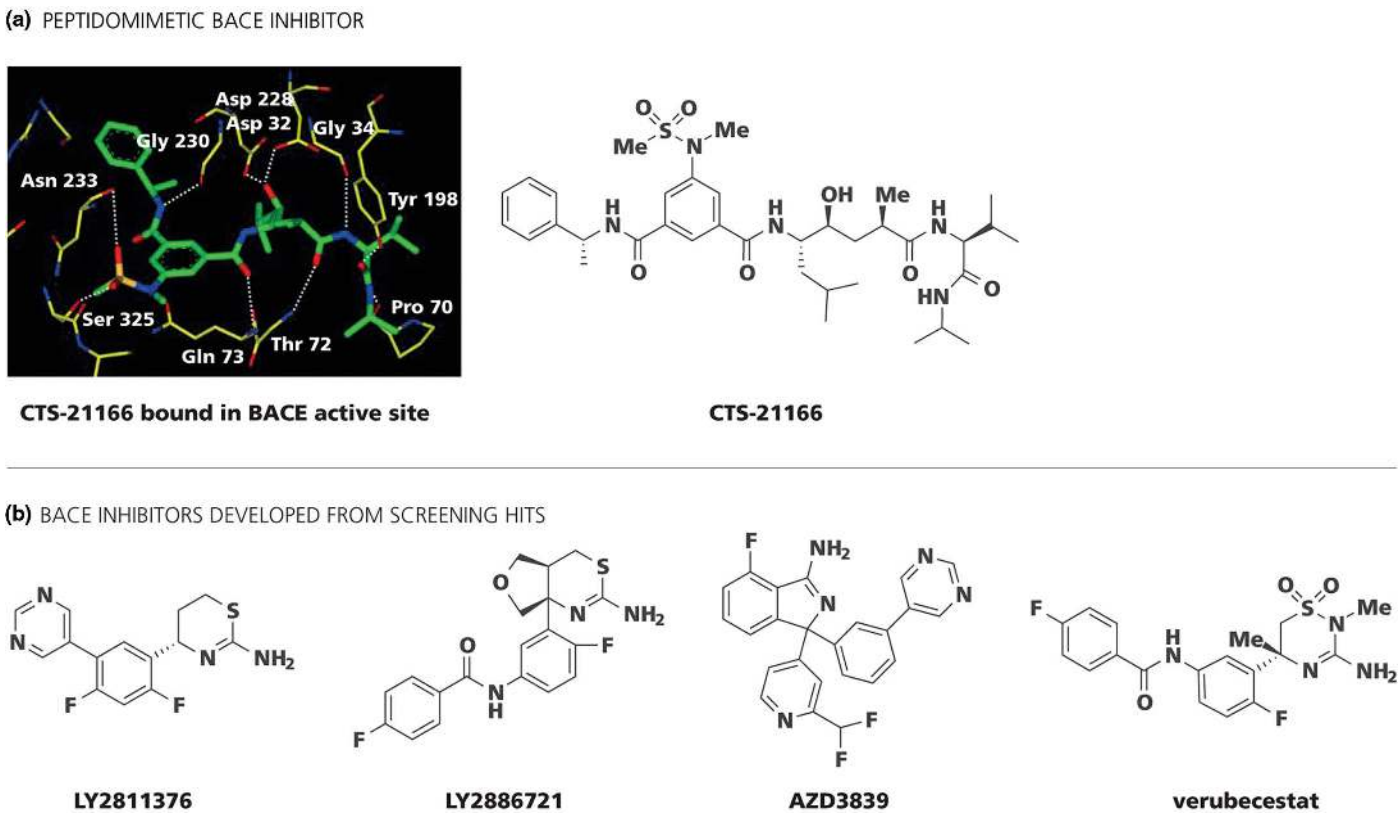


Figure 13.32 Structures of BACE inhibitors investigated to treat Alzheimer's disease. (a) Peptidomimetic BACE inhibitor. X-ray co-structure of an early peptidomimetic, CTS-21166, bound in the BACE active site. Binding interactions include H-bonding to the active site aspartates and additional H-bonds to the backbone. (Courtesy of Ghosh laboratories.) (b) BACE inhibitors developed from screening hits. Other BACE inhibitors, such as LY2811376, LY2886721, AZD3839, and verubecestat, originated from screening programs and were optimized by structure-based design. Several advanced to Phase II and III clinical trials.

Parkinson's disease, observed upon autopsy, is the presence of structures called Lewy bodies in the brain, which are composed of aggregates of the protein α-synuclein. While the cause of PD is unknown, it is believed to be due in small part to a genetic predisposition, possibly combined with prolonged exposure to environmental toxins and aging. Some support for the idea that external exposure can cause Parkinsonism came in the 1980s, with reports of young drug users being hospitalized with Parkinson-like symptoms. Investigation found that these patients had been taking a designer drug based on meperidine. Reversing the ester linkage of meperidine gave MPPP, a compound that had some of the analgesic effects of meperidine but was never marketed. The reverse ester readily underwent elimination to afford an alkene (MPTP), which was an impurity in the drug preparations (**Figure 13.33**). Metabolism of MPTP gives MPP^+, which is neurotoxic and destroys dopaminergic neurons in the substantia nigra. Prior to this discovery, there were no animal models for Parkinson's disease, and MPP^+ is now used to induce Parkinsonism in animal models for PD drug discovery.

While no drugs are currently available to cure Parkinson's disease, the few that are on the market alleviate symptoms by restoring levels and activity of dopamine. Since dopamine does not cross the blood–brain barrier, treatment with it would not be effective. Instead, the precursor L-dopa is administered and is taken up by amino acid transporters and delivered to the brain. Once in the brain, it is then converted to dopamine by the enzyme DOPA decarboxylase, as shown in **Figures 13.34a** and **c**.

The bioavailability of L-dopa is low due to first-pass metabolism, and about 95% of the drug is decarboxylated to dopamine before it can cross the BBB. Hence, it is given in combination with the DOPA decarboxylase inhibitor carbidopa (see **Figure 13.34b**) to improve safety and bioavailability. Another strategy for increasing dopamine levels is the use of deprenyl, a monoamine oxidase B (MAO-B) inhibitor that slows the degradation of dopamine to inactive 3,4-dihydroxyphenylacetic acid (DOPAC) (see **Figure 13.34a**). **Figure 13.34c** illustrates the interactions and effects of these various drugs.

Instead of increasing dopamine levels, another approach uses dopamine agonists to mimic the effects of dopamine. Several drugs have been developed by this approach. The earliest drug approved that works through this mechanism, apomorphine, a natural product, is also formed upon acid degradation of morphine (**Figure 13.35**).

Apomorphine occurs naturally in water lily and was used in both Mayan and Egyptian cultures as a hallucinogen and aphrodisiac. It is a nonselective D_2–D_5 agonist but lacks the opiate activity of morphine. Apomorphine was first investigated by Erick Harnack in the late 1880s for its emetic effects and effects on stereotypy (repetitive behaviors) in farm animals. While its use to treat PD was postulated in 1884, this was not put into practice until the 1970s.

Other dopamine agonists were developed on the basis of their structural similarity to apomorphine. These include pramipexole, ropinirole, and rotigotine. The common key structural element is based on the 2-aminotetralin core of apomorphine (shown in red in **Figure 13.34**) that is mimicked in the synthetic analogs by various isosteric aryl groups and an appropriately positioned substituted amine. The dopamine agonist pergolide is based on the ergot alkaloid structure. This compound was used briefly but was withdrawn due to cardiac side effects. A novel series of compounds was developed at Pfizer based on in a high-throughput screening campaign that screened more than 3 million compounds. The hit from this campaign led to the D1/D5 partial agonists tavapadon and razpipadon, which are being developed by Cerevel. Tavapadon has advanced to late Phase III clinical trials for Parkinson's disease, and razpipadon is in trials for dementia-related apathy.

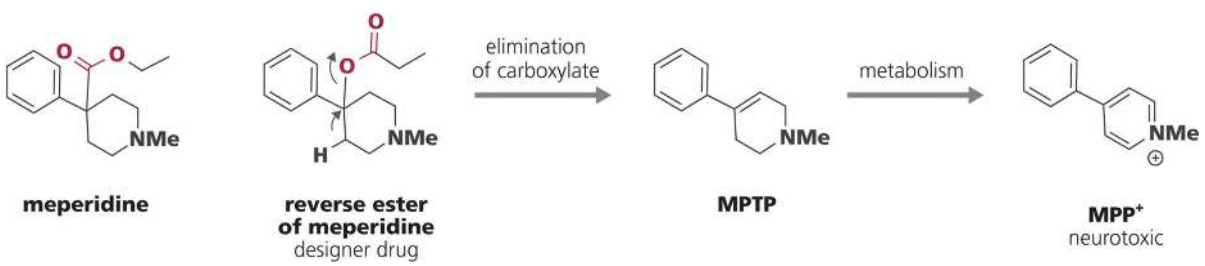


Figure 13.33 Elimination and metabolism of meperidine analog to give MPP^+ neurotoxin. The reverse ester of meperidine, MPPP, was made as a designer drug (change in red). This change makes the ester susceptible to elimination, and the product, MPTP, an impurity in the designer drug preparation, is metabolized to generate the neurotoxic MPP^+.

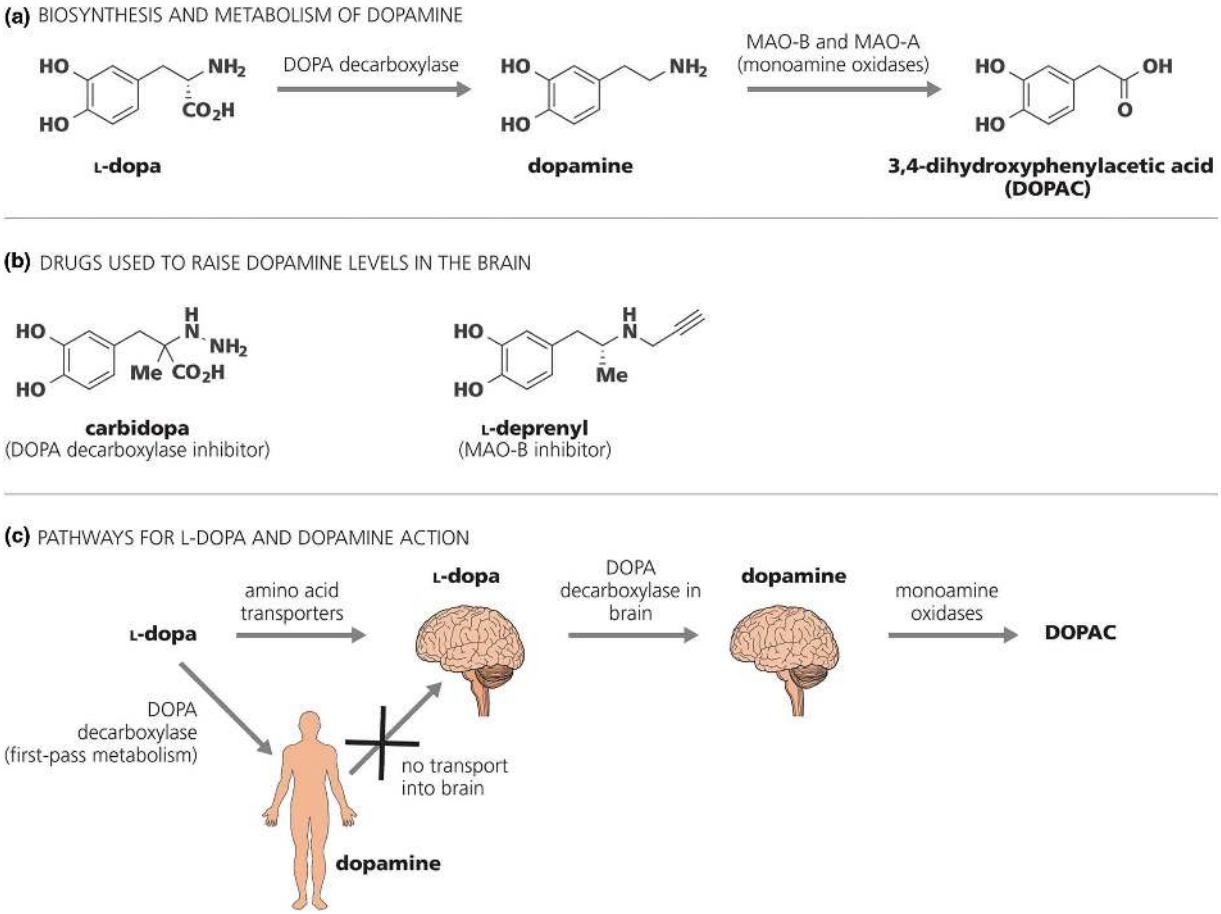


Figure 13.34 Targeting dopamine to treat Parkinson's disease. (a) Biosynthesis and metabolism of dopamine. L-dopa, delivered to the brain, is decarboxylated to generate dopamine, which is degraded by MAO-A and MAO-B to DOPAC. (b) Drugs used to raise dopamine levels. Carbidopa, administered in combination with L-dopa, blocks decarboxylation of L-dopa during first-pass metabolism, allowing it to accumulate in the brain; L-deprenyl inhibits enzymes responsible for metabolism of dopamine. (c) Pathways for L-dopa and dopamine action; L-dopa is taken into the brain by amino acid transporters, while dopamine is not. First-pass metabolism by DOPA decarboxylase lowers L-dopa levels, converting it into dopamine. Metabolism of L-dopa in the brain increases dopamine levels; these are lowered by subsequent metabolism by monoamine oxidases.

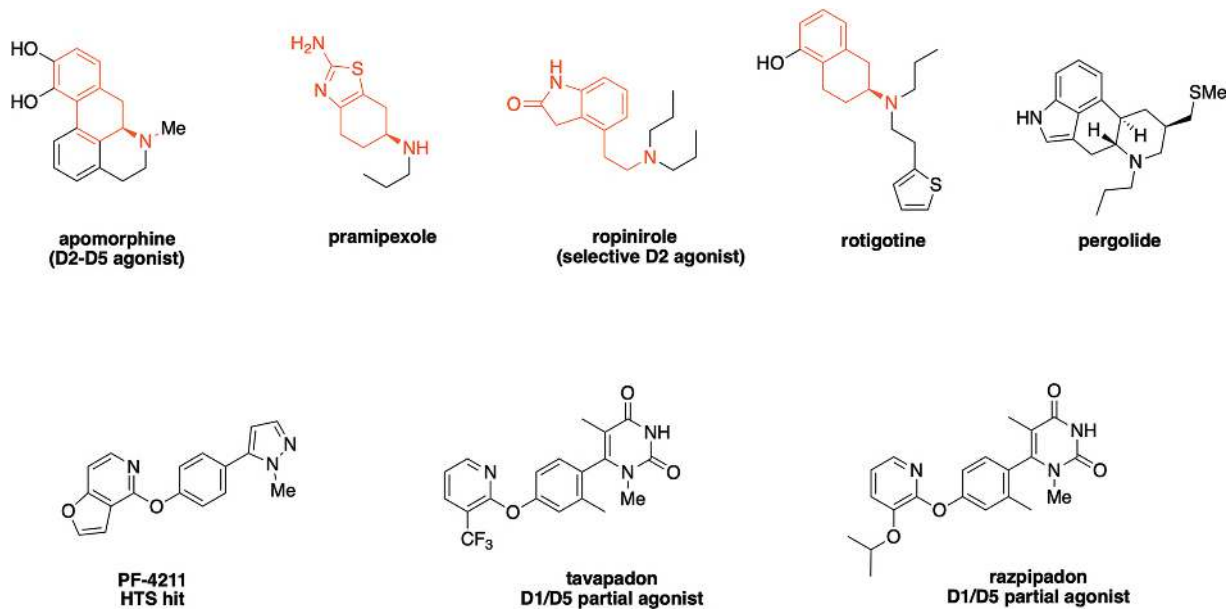


Figure 13.35 Structures of dopamine agonists used to treat Parkinson's disease. Apomorphine, a natural product, was the earliest dopamine agonist used to treat PD. Pramipexole, ropinirole, and rotigotine are based on the apomorphine scaffold, with all compounds containing an aromatic group and amine (similarities to apomorphine are shown in orange). Pergolide was developed on the basis of the structure and activity of ergot alkaloids but is no longer prescribed. Tavapadon and razpipadon were optimized based on the hit PF-4211 identified in a high-throughput screening program and are in late clinical trials.

Neurodegenerative diseases have proven to be some of the most difficult to treat, as seen by the two examples discussed here. While Alzheimer's and Parkinson's diseases are primarily diseases related to aging, there are a number of others that are related to genetic factors. Older drugs to treat AD and PD relied on increasing levels of NT levels in an effort to modify symptoms. Newer approaches are based on disease-modifying agents and include perhaps the first examples of structure-based drug design applied to CNS diseases. However, none of these newer approaches has yet successfully translated to drugs.

13.6 DRUGS OF ABUSE

Unfortunately, a number of the drugs mentioned in this chapter that have valid medicinal uses are also misused. Drug abuse and addiction are related to the self-reward system, which is initiated by dopamine release from dopaminergic neurons in certain parts of the brain. Any drug that increases dopamine transmission in the mesolimbic dopamine system (self-reward system), either directly or indirectly, potentiates self-reward and has a potential for abuse. Drug addiction is characterized by physical and/or psychological dependence. With drugs that cause physical dependence, a tolerance to the drug causes a need for larger doses of drug to achieve the desired effect. Physical withdrawal symptoms include nausea, sweating, difficulty breathing, tremors, and heart palpitations. Drugs that cause psychological dependence lead to uncontrolled compulsive use of the drug, and withdrawal symptoms include anxiety, depression, insomnia, and irritability.

13.6.1 Stimulants Include Both Legal and Illegal Drugs

Stimulants cause improvement in both mental and physical function, increasing alertness and endurance. They also can increase heart rate and blood pressure and may decrease food intake and act as appetite suppressants. Caffeine, the most widely used legal stimulant, is isolated from the beans of the *Coffea* plant but is also found in other plant sources (**Figure 13.36**). Its structure resembles the purine base of adenosine, and it acts as a competitive antagonist at adenosine A_{2A} receptors, which are involved in the sleep–wake cycle. Caffeine is nonselective and is also an agonist at the β-adrenergic receptors in the heart, increasing heart rate. Nicotine, found in the tobacco plant, is another widely used legal stimulant. This plant alkaloid acts as an agonist at the nicotinic acetylcholine receptor (see Section 6.3.1). Its action on receptors in the adrenal medulla indirectly increases levels of dopamine as well as adrenaline, raising both heart rate and blood pressure.

Other stimulants include the phenethylamine amphetamines, both legal and illegal, as well as cocaine. Amphetamines are derived from ephedrine and act as norepinephrine and dopamine reuptake inhibitors, causing increases in synaptic concentrations of both neurotransmitters. Ephedrine, found in *Ephedra sinica*, has been used for thousands of years as a stimulant and asthma treatment (see Section 1.2.1). Although still legal as a treatment for asthma and colds, its use as a dietary supplement (as a stimulant and for weight loss) was banned in the United States in 2004. Two similar compounds, cathinone and cathine, are found in the plant *Catha edulis*. The plant is found in East Africa, and its leaves, known as khat or qat, are chewed for stimulant and euphoric effects. Both are controlled substances in the United States.

While legal and beneficial drugs, such as methylphenidate, used for the treatment of ADHD, have been developed from amphetamines, a number of drugs of abuse also come from this family (**Figure 13.37**). Amphetamine was originally synthesized in the 1880s, and methamphetamine was synthesized from ephedrine in 1919. In the 1930s, amphetamine was marketed as Benzedrine, a treatment for asthma as well as for bronchial congestion. Its stimulant and anorectic effects led to it being prescribed to treat narcolepsy and as diet pills. Methamphetamine was widely used by soldiers in World War II and afterward as a stimulant.

By the 1960s, methamphetamine was becoming widely abused, with extended use causing amphetamine psychosis. It was made illegal in 1970, and as of 2025, methamphetamine is still among the top ten abused illegal drugs in the United States. Amphetamine is still a legal drug; it is marketed as Adderall to treat ADHD. Methylhexanamine

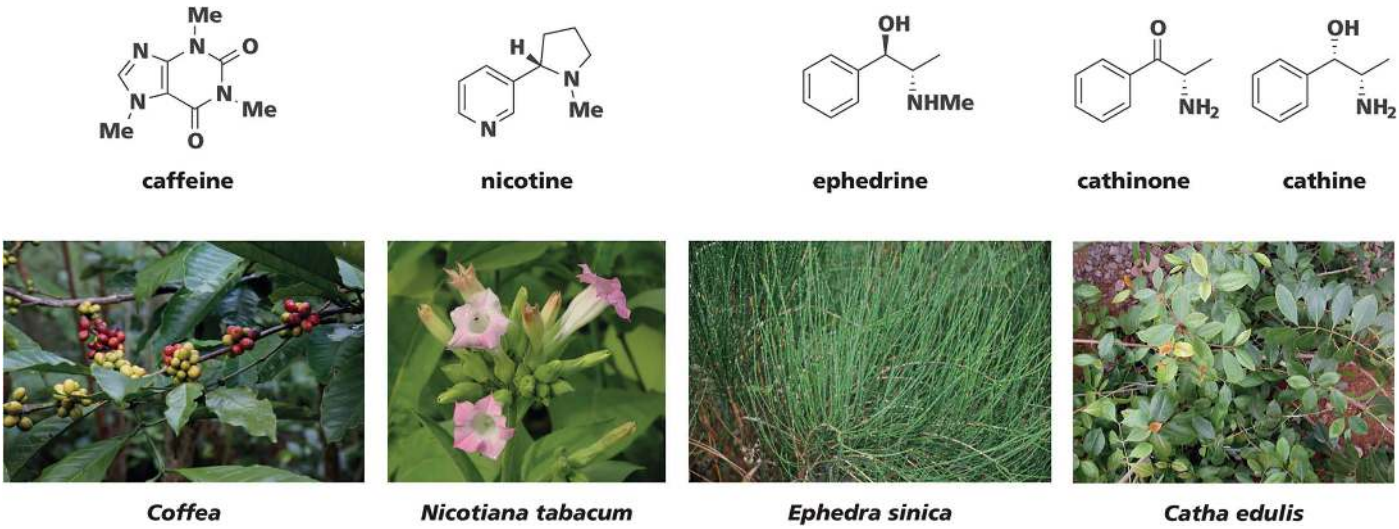


Figure 13.36 Naturally occurring plant-based stimulants and their sources. Caffeine and nicotine are legal stimulants, while ephedrine has been banned as a dietary supplement. The closely related stimulants cathinone and cathine are controlled substances in the United States.

Source: Coffee photo by Jean-Marie Hullot/ Wikimedia Commons, CC BY-SA 3.0. *Nicotiana tabacum* photo by Joachim Mullerchen/ Wikimedia Commons, CC BY–2.5. *Ephedra* photo by Lazare Gagnidze/Wikimedia Commons, CC BY-SA 4.0.

is commonly known as DMAA (1,3-dimethylamylamine) and was originally marketed in the 1960s as a nasal decongestant. Later, in 2006, it was marketed as a natural dietary supplement, in products such as OxyElite Pro and Jack3d, and used as a stimulant. Although DMAA has been touted as being found in geraniums, the FDA found no evidence that this is true and issued warnings in 2012 for its removal. This drug has been abused by athletes and is banned as a stimulant, with deaths of several athletes attributed to DMAA, which causes elevated heart rate and high blood pressure.

Cocaine is a psychostimulant that is one of the most widely abused drugs. This highly addictive compound acts as a dopamine reuptake inhibitor, among other mechanisms. While used in South American cultures for thousands of years, cocaine was brought to Europe in the 1800s. It was legally used at that time to treat dental pain and introduced as a stimulant and euphoric, but its addictive properties were recognized in the late 1800s and it has been banned since the early 1900s. There are currently over 1 million cocaine users in the United States alone.

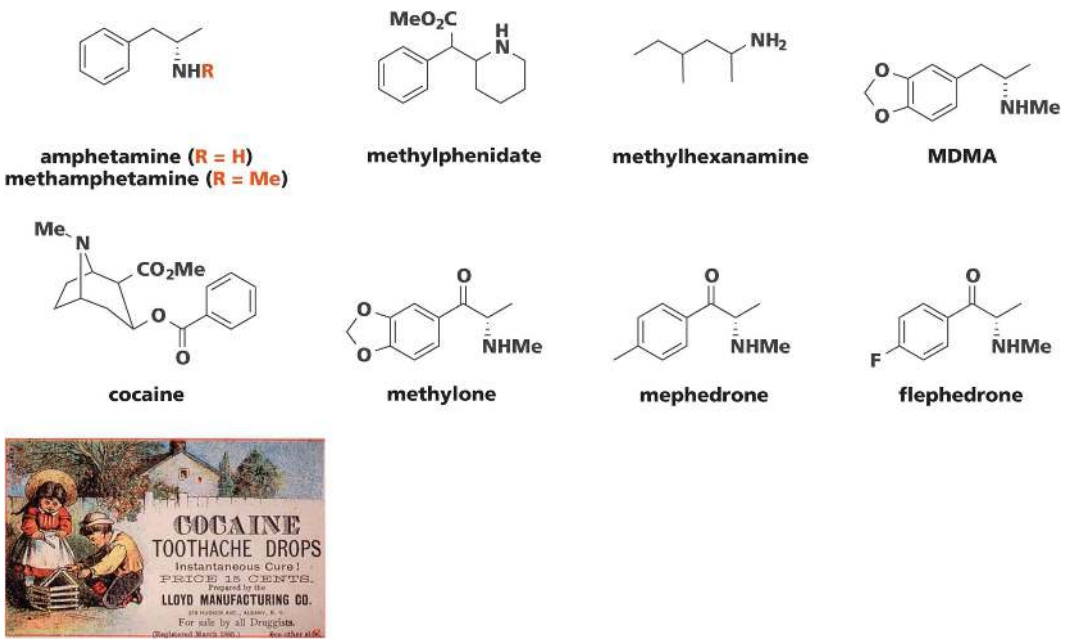


Figure 13.37 Structures of psychostimulants. Amphetamine and methamphetamine were originally used to treat asthma and as stimulants. Methylphenidate, prescribed to treat ADHD, is a derivative of amphetamine, as is methylhexanamine. Cocaine is a natural product originally used as a local anesthetic. MDMA, methylone, mephedrone, and flephedrone are all stimulants and empathogens.

Other phenylethylamines include MDMA, or ecstasy, which is an empathogen, increasing feelings of empathy and social connection. MDMA was originally synthesized in 1912 as a stimulant but was not marketed. Its psychotropic effects were noted in the 1970s, and it became a schedule I drug in 1985. In addition to acting at the dopamine and norepinephrine transporters, MDMA also increases levels of serotonin. It has weak agonist activity at the 5-HT_1 and 5-HT_2 receptors, resulting in an increase of oxytocin levels and its activity as an empathogen. A more recent trend is the use of cathinone derivatives, such as methylone, mephedrone, and flephedrone, which act as stimulants and empathogens. These chemicals are sprayed onto dried plant material and sold under names such as plant food and bath salts in convenience stores, labeled "not for human consumption."

13.6.2 Hallucinogens Include Serotonin Receptor Agonists and *N*-methyl-D-Aspartate Antagonists

The classical hallucinogens have the general structure Ar-C-C-N and are structurally related to serotonin. They bind at the 5-HT_{2A} receptor, acting as agonists, although serotonin itself is not hallucinogenic. The indole alkylamines, natural products or derivatives thereof that are found in a variety of sources, are one category. Psilocin is a hallucinogenic serotonin analog that is found in *Psilocybe* mushrooms (**Figure 13.38**). Bufotenine, the *N*,*N*-dimethyl analog of serotonin, is secreted by psychoactive toads, such as *Bufo alvarius*, that are found in the American Southwest. It has also been isolated from some mushrooms. The closely related *N*,*N*-dimethyltryptamine has been used in the Amazonian Indian cultures as a spiritual medicine and is found in the ayahuasca brew made of several plant preparations. A more complex structure in this family is LSD, lysergic acid diethylamide. It is a semisynthetic derivative of an ergot alkaloid found in the fungus ergot, which infects grasses. Ingestion of these fungus-infected plants led to a condition called St. Anthony's fire, with symptoms including hallucinations. Historically, these symptoms are believed to have led to accusations of witchcraft. While studying the ergot compounds, Albert Hofmann, at Sandoz Pharmaceuticals, synthesized the *N*,*N*-diethylamide LSD in 1938. Although his goal was

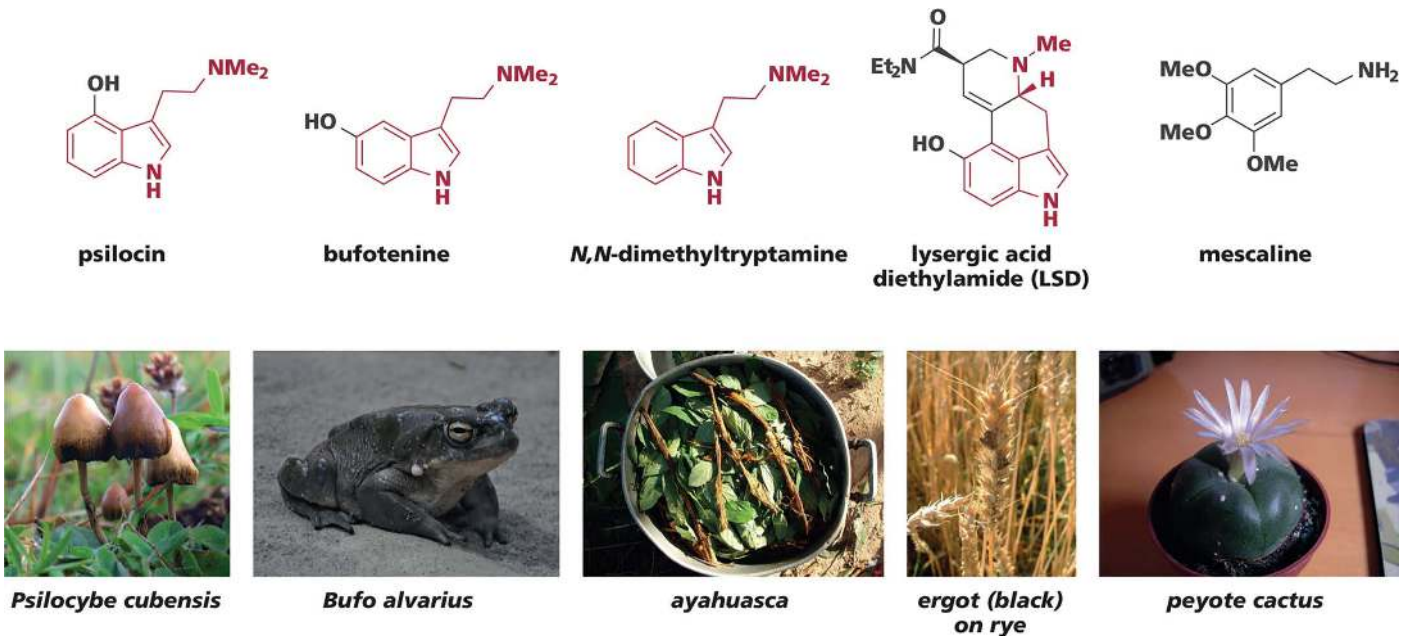


Figure 13.38 Structures of naturally occurring and derived hallucinogens and their sources. These include the indole alkylamines psilocin, bufotenine, *N*,*N*-dimethyltryptamine, and LSD and the phenylalkylamine mescaline. Most contain a common indole alkylamine substructure (shown in red). The natural sources of each compound or precursor are also shown.

Source: Psilocybe photo by the Mushroom Observer/Wikimedia Commons, CC BY-SA 3.0. Ayahuasca photo by Awkipuma/Wikimedia Commons, CC BY-SA 3.0. Ergot photo by JoJan/Wiki-media Commons, CC BY-SA 2.5.

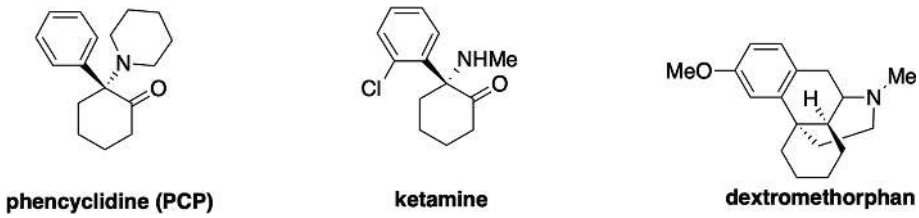


Figure 13.39 Structures of NMDA antagonists drugs of abuse. Both phencyclidine and ketamine were originally developed as anesthetics, with ketamine still used for that purpose. Dextromethorphan is a cough suppressant but is also a weak NMDA antagonist.

to make a circulatory stimulant, accidental ingestion of a small amount of this compound resulted in hallucinations and an altered state of consciousness. The drug was used in psychiatry in the 1950s and was popularized in the 1960s as a recreational drug, most notably by Timothy Leary, a Harvard psychologist.

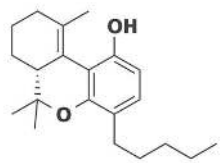
Currently, LSD is still widely used, with estimates of about 200,000 teens per year using it in the United States, and with reports in the popular press of microdoses being useful to treat certain neuropsychiatric disorders. Ayahuasca is legal in most South American countries and is actually promoted in healing tours to Peru. There are anecdotal reports of its use to treat posttraumatic stress disorder (PTSD) as well as drug addiction, and clinical trials are in progress to evaluate its potential use in depression and other neuropsychiatric disorders. Another category of hallucinogens is the phenylalkylamines, such as mescaline, found in peyote cactus. Like the indole alkaloids, mescaline is a serotonin agonist. Mescaline has been used by Native American tribes in the Southwest United States for thousands of years and gained popular attention in Aldous Huxley's book *The Doors of Perception*.

Phencyclidine (PCP), ketamine, and dextromethorphan act as noncompetitive NMDA receptor antagonists and, as such, have the indirect effect of increasing dopamine levels (**Figure 13.39**). They are dissociative drugs that distort perceptions of sight and sound and lead to a feeling of detachment. PCP was synthesized and used as an anesthetic in the 1950s and was originally believed to be relatively safe since it caused no cardiac or respiratory suppression. However, it was found to induce schizophrenia-like psychosis and was withdrawn from use in 1965.

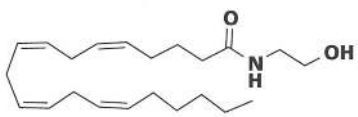
PCP binds at an allosteric site inside the NMDA receptor, which is an ion channel (Na^+ and Ca^{2+}). Binding of *N*-methyl-D-aspartate and glycine opens the channel, and PCP prevents this and inhibits depolarization. Both ketamine and dextromethorphan bind at the same site, blocking the channel. Ketamine is used legally as an anesthetic but is also a drug of abuse. Dextromethorphan is the enantiomer of an opiate and is used as a cough suppressant. However, in high doses, it can cause hallucinations due to its activity as an NMDA antagonist. As mentioned in Section 13.2.1, both of these drugs have now been approved to treat depression, in low doses.



Cannabis sativa



Δ^9-THC (tetrahydrocannabinol)



anandamide

Figure 13.40 Structures of natural cannabinoid Δ^9-THC, found in *Cannabis sativa*, and anandamide, the cannabinoid receptor's natural ligand. Δ^9-tetrahydrocannabinol is the main psychoactive ingredient in *Cannabis sativa*. Anandamide, the ethanolamine amide of arachidonic acid, is an endogenous ligand of CB receptors.

13.6.3 Cannabinoids Have Potential Medical Uses

The cannabinoids are found in the *Cannabis sativa* plant, which has been used for medicinal purposes since about 6,000 BC. Their effects include euphoria and enhanced sensory perception, as well as alleviation of pain; less-desirable effects consist of difficulty in concentration and memory impairment. The component primarily responsible for the psychoactive effects among various related analogs is Δ^9-THC (tetrahydrocannabinol) (**Figure 13.40**). In the 1990s, the receptors at which THC acts were discovered: CB_1 in the brain and CB_2 in the periphery. Both are GPCRs; the CB_1 receptor is responsible for the euphoric and anticonvulsant effects of cannabinoids, while the CB_2 receptor may be involved in immune effects. As a class, cannabinoids indirectly increase activity of dopaminergic neurons.

Anandamide was identified as the CB receptors' natural ligand in the early 1990s (see **Figure 13.40**). It is hydrolyzed by the enzyme fatty acid amide hydrolase (FAAH), and inhibitors of this enzyme have been investigated as potential antidepressants. There are few actual clinical trials of THC, due to the listing of marijuana as a schedule I drug, which requires DEA approval for studies. In spite of this, synthetic Δ^9-THC was approved by the FDA in the 1990s to treat nausea associated with cancer

chemotherapy, and as of 2025, recreational marijuana is legal in many states, and use of it for medical purposes available in many more.

CB_1 agonists have the potential to be anti-nausea agents and appetite stimulants, and hundreds of synthetic cannabinoids have been investigated in academic and industrial laboratories. Agonists at the CB_2 receptor have been investigated to treat neuropathic pain, inflammation, and seizures. Antagonists have been investigated as appetite suppressants and memory enhancers. One CB_1 antagonist, rimonabant (**Figure 13.41**), was approved in Europe in 2006 to treat obesity but was later withdrawn due to psychiatric side effects, including severe depression.

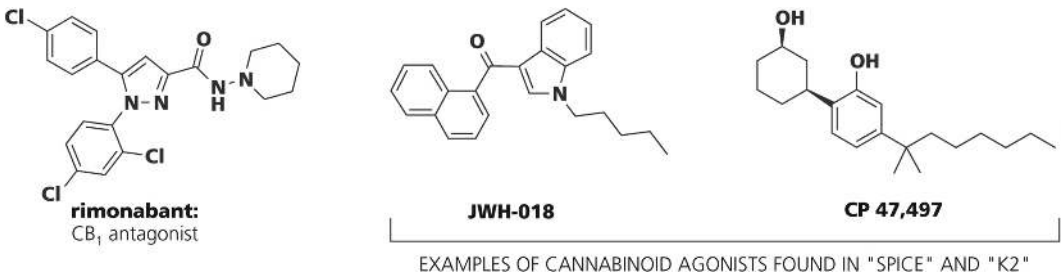


Figure 13.41 Examples of synthetic cannabinoids. Rimonabant, a CB_1 antagonist, was approved to treat obesity but was later withdrawn. JWH-018 and CB 47,497 are CB agonists and have been found in illegal drug substances.

Some of these synthetic analogs, such as JWH-018 and CP 47,497, have been used as designer drugs or synthetic cannabis (see **Figure 13.41**) and have become dangerous drugs of abuse. Spice and K2 are preparations of cheap herbs and a mixture of synthetic cannabinoids sprayed onto the plant material. Since the components were, in most cases, never tested on humans, serious side effects are common. These range from seizures, psychosis, and agitation to death from heart attack, sometimes even after one dose. Although JWH-018 and other *N*-alkyl analogs, as well as CP 47,497, are now illegal in many countries, there are many analogs being sold that are still not classified as illegal.

Many drugs of abuse are based on natural products that have been known for centuries and exert dangerous psychotropic effects. However, the study of these agents and their mechanism of action has served to help the development of legitimate and potentially valuable drugs. For example, the study of cannabis led to identification of the cannabinoid receptors and potential new drug targets CB_1 and CB_2.

13.7 SUMMARY

For thousands of years, people have used plant preparations that have effects on the CNS for relief of pain, as stimulants, and for euphoric effects. However, until the 20th century, there were no known medications for mental health conditions, such as schizophrenia, anxiety, and depression. However, today there are now many efficacious drugs that alleviate neuropsychiatric conditions. Many of the therapies that originated in the mid-20th century were based on clinical observations, with knowledge of their molecular target not found until later. However, today newer treatments for neuropsychiatric disorders were derived from target-driven approaches and based on modern understanding of how neurotransmitters are regulated. The complexity of pathways in the brain and maintaining a balance of neurotransmitters makes drug discovery in this area very challenging. Furthermore, patients may respond differently, or not at all, even when given the same drug.

In contrast, there are far fewer drugs that successfully treat neurodegenerative diseases. These are particularly challenging, and with an aging population, there is a more urgent need for effective drugs. With a better understanding of the causes of diseases such as Alzheimer's and Parkinson's, there have been numerous efforts put into developing drugs. While a number have advanced to clinical trials, as yet, there are no drugs that affect the course of these diseases.

CASE STUDY 1 DEVELOPMENT OF VILAZODONE

Vilazodone was added to the arsenal of antidepressants in 2011; it exhibits combined SSRI activity and partial agonism at the 5-HT_{1A} receptor. This combination helps avoid the typical slow onset of antidepressant effects of the SSRIs and also exemplifies the complexity of targeting neurotransmission. The slow onset of SSRIs is believed to be due to a negative feedback loop: SSRIs increase levels of serotonin in the synapse, which leads to the activation of presynaptic 5-HT_{1A} autoreceptors and the inhibition of firing of serotonergic neurons; this feedback loop limits the rise in extracellular serotonin that SSRIs cause. Chronic use of SSRIs over a period of weeks desensitizes the 5-HT_{1A} autoreceptors, allowing neurons to resume normal firing. The observation, made in the mid-1990s, that coadministration of the 5-HT_{1A} antagonist pindolol and SSRIs led to faster onset of antidepressant effects in humans was the basis of the idea of finding a dual SSRI/5-HT_{1A} antagonist. However, antagonism of postsynaptic 5-HT receptors would be undesirable since the overall goal is to increase serotonin activity. Alternatively, use of a 5-HT_{1A} agonist (rather than antagonist) along with an SSRI could speed up the desensitization, also leading to a faster overall increase in extracellular serotonin.

The discovery of vilazodone began in the 1990s with Merck's investigation of the indole alkylamine roxindole as a potential antipsychotic. Roxindole is similar in structure to the butyrophenone antipsychotics, such as haloperidol, but it lacks the ketone moiety. While roxindole had high affinity for dopamine receptors D_2 and D_3, in functional assays, it is an agonist, and in animal models, it was found to have activity predictive of anxiolytic and antidepressant action. In addition to its dopamine activity, it is also a potent selective serotonin agonist, with selectivity for the presynaptic 5-HT_{1A} receptor subtype, and it inhibits the serotonin transporter (**Figure 13.42**).

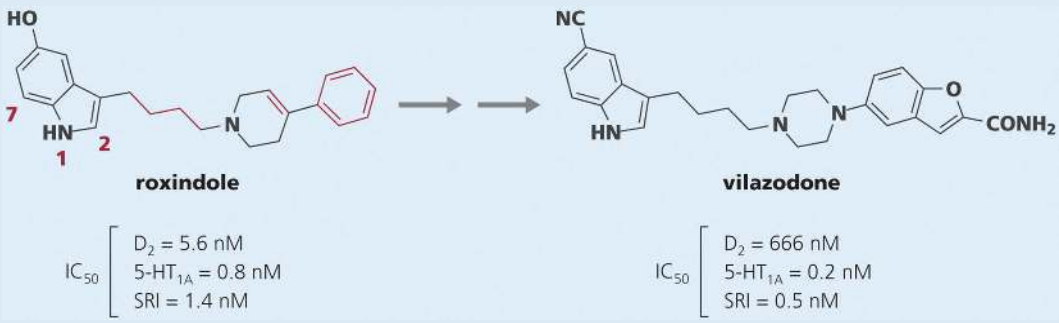
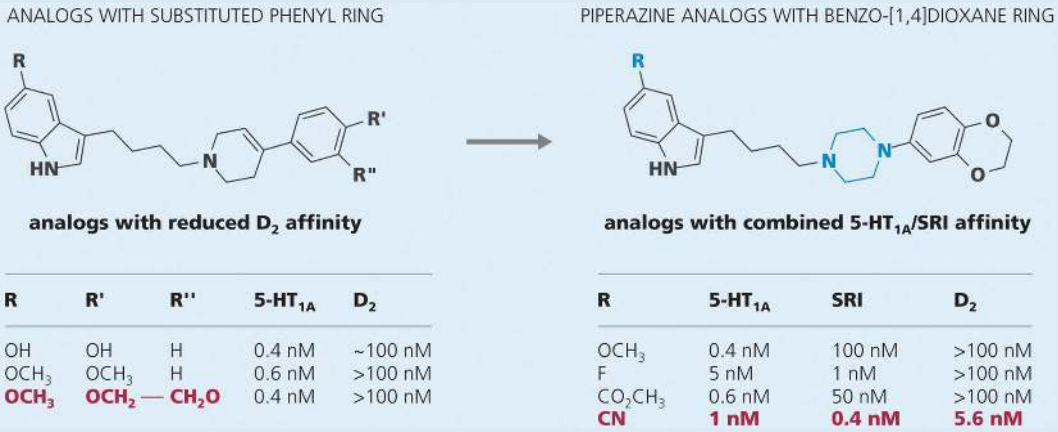


Figure 13.42 Structures of roxindole and vilazodone. The lead, roxindole, is a mixed agonist (requirements for D_2 agonism are shown in red). The final drug, vilazodone, is a selective 5HT_{1A} agonist and SSRI.

During the course of studies to differentiate dopamine agonist activity from 5-HT_{1A} receptor binding, SAR for each activity emerged. For binding to dopamine (D_2) receptors, requirements are the unsubstituted phenyl ring appended to the 4-phenyltetrahydropyridine, unsaturation in the heterocyclic piperidine ring, a saturated butyl linker between the indole and basic nitrogen, and an indole with no substituents at positions 1, 2, or 7. Addition of substituents on the phenyl ring (in particular, hydroxyls and ethers) of roxindole led to selective 5-HT_{1A} agonists, and most eliminated dopamine binding while maintaining serotonin reuptake inhibition, as seen in **Figure 13.43**.

While the 5-hydroxyindole piperidine analog exhibited the desired binding profile, the alkoxy analogs were preferred due to greater metabolic stability. Changes to the unsaturated piperidine system were made by replacement with saturated piperidines as well as piperazines. In particular, the piperazines (see **Figure 13.43**) were investigated further.

In this series, substitution of an electron-withdrawing group at the 5-position on the indole maintained serotonin reuptake inhibition (SRI), and the 5-F and 5-CN analogs were chosen for further study. Although the SRI and 5-HT_{1A} binding profile was optimal, the 5-cyano derivative still had substantial binding to the D_2 receptor. A series of compounds was prepared in which the benzo[1,4]dioxane was replaced with other



R	R'	R''	5-HT$_{1A}$	D$_2$
OH	OH	H	0.4 nM	~100 nM
OCH_3	OCH_3	H	0.6 nM	>100 nM
OCH_3	**OCH_2 — CH_2O**		0.4 nM	>100 nM

R	5-HT$_{1A}$	SRI	D$_2$
OCH_3	0.4 nM	100 nM	>100 nM
F	5 nM	1 nM	>100 nM
CO_2CH_3	0.6 nM	50 nM	>100 nM
CN	**1 nM**	**0.4 nM**	**5.6 nM**

Figure 13.43 Examples of series investigated to optimize 5-HT$_{1A}$ and serotonin transporter binding profile and eliminate dopamine effects. Introduction of substitution on the phenyl ring of roxindole with hydroxyl and ether groups eliminated D$_2$ affinity. Substitution of piperazine (blue) for the unsaturated piperidine and incorporation of a substituent on the indole ring at position 5 (blue) decreased dopamine activity but maintained 5-HT$_{1A}$ potency. The benzo[1,4]dioxane series had the desired 5-HT$_{1A}$/SRI profile with either fluoro or cyano substituents on the indole ring, although the cyano analog (red in table) had undesirable D$_2$ affinity.

heterocycles, such as benzofuran, while the 5-fluoro and 5-cyano indole substituents were maintained. Of these, the carboxamide-substituted benzofuran known as vilazodone exhibited the optimal binding profile, maintaining transporter and 5-HT$_{1A}$ binding but decreasing D$_2$ binding from 5.6 to 666 nM.

In addition to binding assays, compounds were tested *in vivo* (in rats) in assays for ultrasonic vocalization as well as elevation of core temperature, which are reflections of activation of presynaptic 5-HT$_{1A}$ autoreceptors. Inhibition of apomorphine-induced climbing in mice was used as a measure of D$_2$ antagonism. In these animal models, vilazodone showed an *in vivo* profile of 5-HT$_{1A}$ agonism with no evidence of D$_2$ antagonism.

Vilazodone entered clinical trials in the early 2000s and showed efficacy against major depressive disorder, with side effects primarily being diarrhea and nausea. Advantages include decreased sexual dysfunction as compared to other SSRIs, as well as anti-anxiety activity. Although there has not been a direct comparison in time of onset, one clinical trial showed improvement in symptoms after 1 week, while other SSRIs typically take 3–4 weeks to be effective. Vilazodone was licensed to Forest Pharmaceuticals and received FDA approval in 2011.

CASE STUDY 2 DISCOVERY OF DONEPEZIL

The first drug approved to treat Alzheimer's dementia was tacrine, in 1993. Although there was some improvement in the memory of treated patients, shortly after its introduction, concerns developed over hepatic toxicity. At Eisai in Japan, library screening for acetylcholinesterase inhibitors identified *N*-benzylpiperazine **1**, which had originally been synthesized as a potential treatment for arterial sclerosis (**Figure 13.44**). Compound **1** had an IC$_{50}$ of 12,600 nM against rat brain AChE, and a program was initiated to improve its potency. A large number of analogs were prepared, including those with substitutions on and for the heterocyclic ring, which led to *N*-benzylpiperidine **2** with binding of 340 nM. Replacement of the ether linkage with an amide further improved potency. Investigation of aryl substituents led to a promising candidate, sulfonate **4**, with 21,000× better binding to AChE than the lead compound **1**. However, while novel and selective, **4** had a short duration of action.

One of the goals of the program was to develop inhibitors that were selective for acetylcholinesterase (AChE) over butyrylcholinesterase (BuChE). Both enzymes hydrolyze esters of choline. However, AChE is found primarily in the brain while BuChE exists primarily in the periphery, specifically in the heart and intestines. It

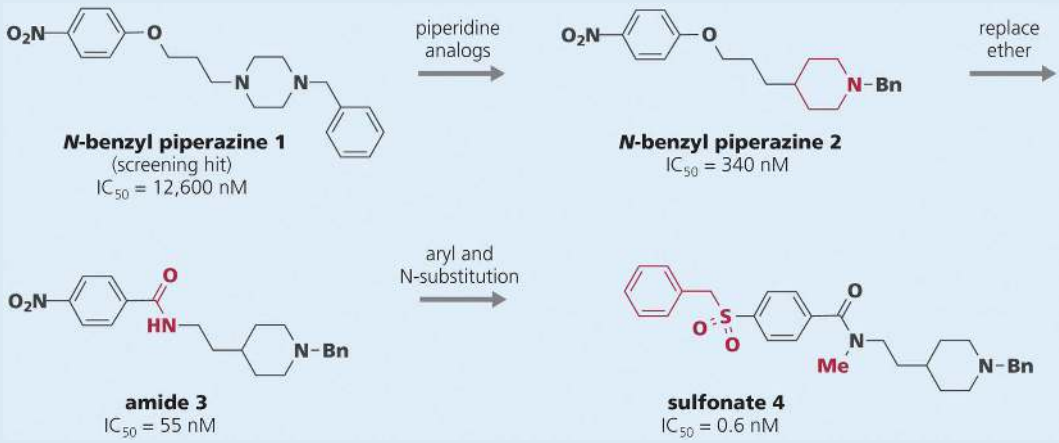


Figure 13.44 Evolution of a screening hit during development of donepezil. The screening hit, *N*-benzylpiperazine **1**, was modified to give the more potent *N*-benzylpiperidine **2**. Replacement of the ether linkage led to the amide **3**, and then N-methylation and substitution for the nitro group gave sulfonate **4**, which exhibited potent AChE activity but poor pharmaceutical properties.

was thought that selectivity would help prevent some of the toxicity of tacrine, which inhibits both enzymes equally. The sulfonate **4**, while not ideal, showed 18,000-fold selectivity for AChE over BuChE and was a reversible inhibitor.

The next advances in structure modification came through cyclization of the amide to the aryl ring (**5**) and then replacement of the labile amide with a ketone moiety (**6**). The constrained analog **5** was twice as potent as its parent acyclic *N*-methylamide, and the ketone (indanone) **6** maintained selective enzyme inhibition and had a longer duration of action (Figure 13.45). Further optimization of linker length and indanone substitution led to donepezil, a reversible noncompetitive inhibitor of AChE with high enzyme affinity and 1,100× selectivity for AChE over BuChE.



Figure 13.45 Late lead development of donepezil. The constrained amide **5** was modified to the indanone **6**, which exhibited a longer half-life. Substitution on the aryl ring and shortening of the linker led to donepezil.

Prior to solution of the X-ray structure of acetylcholinesterase, donepezil was proposed to bind as shown in Figure 13.46: with both aryl groups binding in hydrophobic pockets, the required nitrogen of the piperidine binding to an anionic site, and the ketone accepting a hydrogen bond from a serine OH. Some of these interactions are seen in the actual X-ray structure of donepezil co-crystallized with acetylcholinesterase, which also gives some understanding of its selectivity over BChE.

The active site of acetylcholinesterase sits at the base of a long cleft called the aromatic gorge. While tacrine binds near the active site, the solved X-ray structure showed that donepezil extends through the gorge. Key interactions include binding of the benzyl group through π–π interactions to Trp84 near the active site, binding of the nitrogen to Phe330 through a cation–π interaction, and binding of the aromatic ring of the indanone to Trp279 at the opposite end of the gorge through a π–π interaction. Although there is no H-bond between the serine OH and the indanone carbonyl observed (as proposed), a bridging bond to the backbone through a water molecule is evident. Since BChE does not have the Phe330 and Trp279 residues, the X-ray provides an explanation for selectivity.

In animal models, donepezil exhibited inhibition of AChE in rats at an oral dose of 2.6 mg/kg, as compared to 9.5 mg/kg for tacrine. It also showed good brain permeability, and selectivity was evidenced by 14× better inhibition of brain AChE than plasma

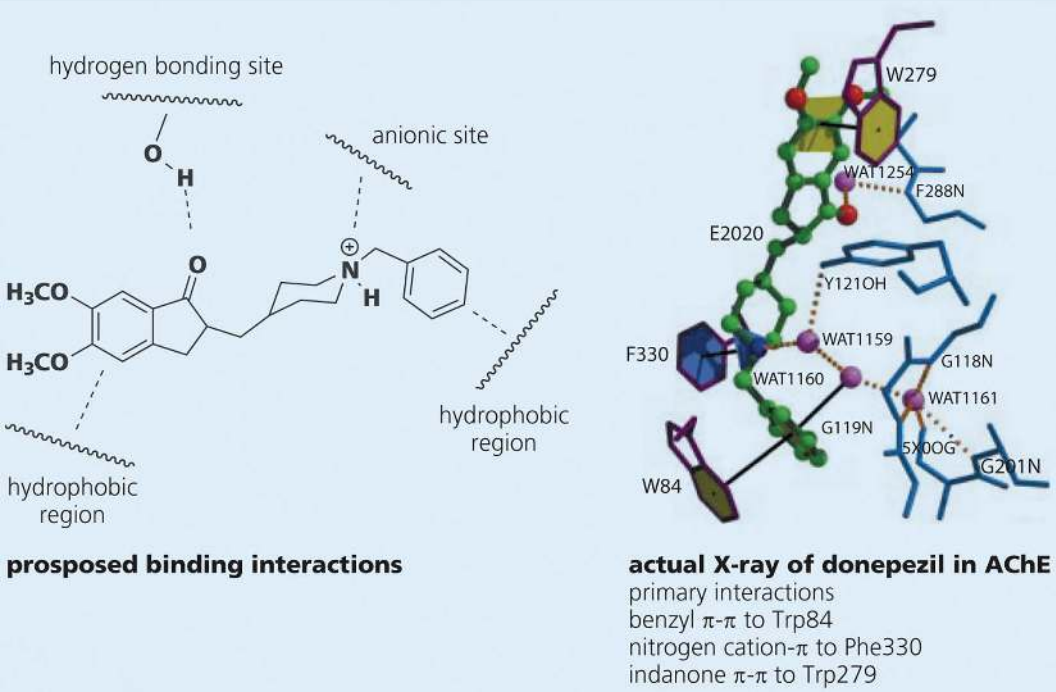


Figure 13.46 **Proposed and observed binding of donepezil (E2020) in AChE.** *Left*, the originally proposed binding had both aryl groups in hydrophobic pockets, the piperidine nitrogen in an anionic site, and the ketone as an H-bond acceptor. *Right*, X-ray co-crystal structure shows donepezil (green) with π–π interactions of the benzyl to Trp84 and indanone to Trp279; the piperidine nitrogen engages in a cation–pi interaction to Phe330.

Source: From Kryger, G.; Silman, I.; and Sussman, J. L. (1999). *Structure* 7:297–307. With permission from Elsevier.

AChE. Donepezil was shown to increase ACh levels in rat brain and to be effective in several learning and memory models in rats and monkeys. Donepezil entered clinical trials in 1989, and while early trials did not show promising efficacy, Phase III trials showed improvement in memory for patients with mild to moderate Alzheimer's disease and only mild side effects. Donepezil was approved by the FDA in 1996, with once-daily oral dosing. While not a cure or a disease-modifying agent for Alzheimer's disease, it can slow progression of memory loss.

13.8 REVIEW QUESTIONS

1. What is the blood–brain barrier, and why is it an important consideration for drug discovery for CNS agents? What is another factor that prevents drugs from having good brain penetration?
2. How do the physical properties of CNS-active drugs compare to those of drugs that act at other organ systems?
3. Identify three classes of drugs used as antidepressants, each with a different mechanism of action. Draw the structure of one example of each.
4. Name five neurotransmitters, and draw their structures.
5. Name five different natural products that have significant CNS effects and are either themselves a drug or the basis of a drug. For each product, what is the plant that they were isolated from, what is the effect, and what is the drug?
6. Describe three classes of pain medications that act by different molecular mechanisms.
7. Haloperidol and olanzapine are both used as antipsychotics. Explain how they differ.
8. Why is targeting specific neurotransmitters as an approach to CNS drug discovery particularly challenging?
9. Give examples of two different classes of drugs of abuse and an example of each.
10. Name two different drugs that are approved to treat Alzheimer's disease. Choose one and explain how it works. Describe recent approaches to develop drugs to treat the underlying cause of the disease.
11. Based on the examples in this chapter, identify one drug that is a positive allosteric modulator, one that is a partial agonist, and one that is an antagonist. Draw the structure of each, and name the receptor at which it acts.

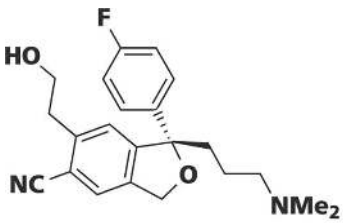
13.9 APPLICATION QUESTIONS

1. Listed in the table are some binding values for reuptake inhibition for (*R*)- and (*S*)- citalopram. Referring back to Chapter 3 and **Figure 3.4**, what is the eudismic ratio for (*S*)-citalopram vs. (*R*)-citalopram for each symporter?

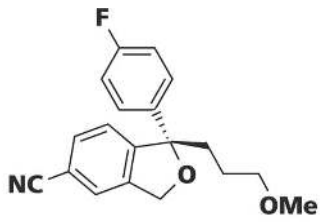
	IC50 (nM)		
Compound	**SRI[a]**	**NRI**	**DRI**
(*S*)-citalopram	2.1	2,500	65,000
(*R*)-citalopram	275	6,900	54,000

[a]SRI = serotonin; NRI = norepinephrine; DRI = dopamine.*Source:* Data from Hyttel, J. (1994). *Int Clin Psychopharmacol* 9 (Suppl 1):19–26. In *Textbook of Drug Design and Discovery*, 3rd ed, p 320. CRC Press.

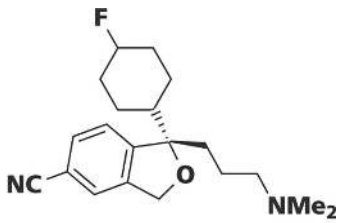
2. Referring to the model of hSERT shown in **Figure 13.10** and the binding values in question 1, explain why the *S*-enantiomer is more potent than the *R*-enantiomer.
3. Three analogs of citalopram are shown. For each one, on the basis of **Figure 13.10** and the binding values in question 1, would you predict it to be more or less active than citalopram? Explain what binding interactions may be gained or lost.



1

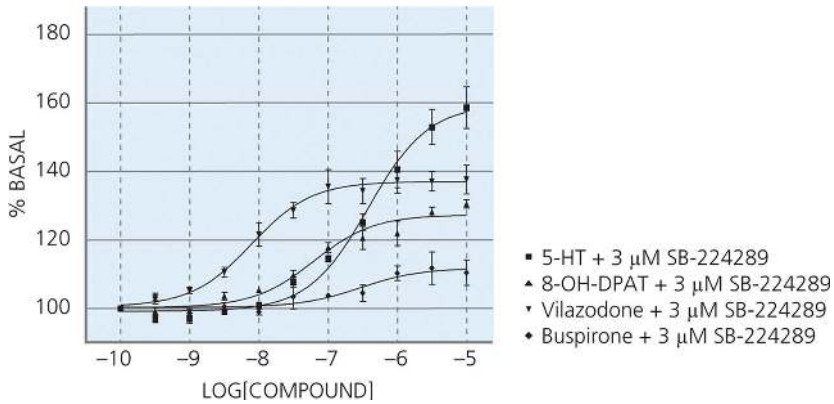


2



3

4. During testing of vilazodone, an assay for the effect on binding of radiolabeled GTPγS to cells expressing 5-HT_{1A} receptors gave the curves shown. This is a functional assay measuring potency and intrinsic activity at the 5-HT_{1A} receptor. SB-224289 is a selective 5-HT_{1B} antagonist and was coadministered so that activity was due only to 5-HT_{1A}. Comparison was done with serotonin, 8-OH-DPAT, and buspirone.



Reference: Hughes, Z. A.; Starr, K. R.; Langmead, C. J.; et al. (2005). "Neurochemical evaluation of the novel 5-HT_{1A} receptor partial agonist/serotonin reuptake inhibitor, vilazodone." *Eur J Pharmacol* 510:49–57.

Referring back to earlier sections in this Chapter, answer the following questions:

a. What are the structures of 8-OH-DPAT, serotonin, and buspirone?

b. Based on the graph, identify whether each of these three compounds is a full agonist, a partial agonist, or an antagonist at 5-HT_{1A}.

c. Which is the most potent?

d. Which has the highest intrinsic activity?

5. Memantine (used for Alzheimer's) and PCP (drug of abuse) are both NMDA antagonists. Research both compounds, and explain why memantine does not have the psychosis-inducing effects of PCP.

6. Using examples in this chapter, identify four instances where conformational constraint was used as a medicinal chemistry strategy.

7. Using examples in this chapter, identify two compounds where an isosteric replacement was used. Compare the isostere to the original compound.

8. Several drugs of abuse that are cannabinoid agonists are described in this chapter. Propose a pharmacophore that is consistent with three of the structures.

9. Compound CTS-21166 is a BACE1 inhibitor that was not successfully advanced. Research the structure of this compound. Does this compound adhere to Lipinski's rules? Would you expect this compound to be brain-penetrant, given what you know about the physical properties required for BBB penetration?

10. **Figure 13.30** shows the binding of rivastigmine to AChE. Given what you know about designing serine proteases, design two new analogs of rivastagmine that contain different warheads.

11. Identify five drugs in this chapter that contain a privileged structure. At least three of them should be different privileged structures.

FURTHER READING

General

Pajouhesh H & Lenz GR (2005) Medicinal chemical properties of successful central nervous system drugs. *NeuroRx* 2:541–553.

Ornstein P (2015) Some thoughts on the past 50 years of central nervous system therapeutic research. In *2015 Medicinal Chemistry Reviews* (Desai MC ed), Vol 50, pp 31–61. Medicinal Chemistry Division, American Chemical Society.

Anti-Anxiety/Antidepressant

Calcaterra NE & Barrow JC (2014) Classics in chemical neuroscience: diazepam (Valium). *ACS Chem Neurosci* 5:253–260.

Dawson LA & Watson JM (2009) Vilazodone: a 5-HT1A receptor agonist/serotonin transporter inhibitor for the treatment of affective disorders. *CNS Neurosci Ther* 15:107–117.

Depoortere H, Zivkovic B, Lloyd KG et al. (1986) Zolpidem, a novel nonbenzodiazepine hypnotic. I. Neuropharmacological and behavioral effects. *J Pharmacol Exp Ther* 237:649–658.

Heinrich T, Bottcher H, Gericke R et al. (2004) Synthesis and structure–activity relationship in a class of indolebutylpiperazines as dual 5-HT_{1A} receptor agonists and serotonin reuptake inhibitors. *J Med Chem* 47:4684–4692.

Lopez-Munoz F, Ucha-Udabe R & Alamo C (2005) The history of barbiturates a century after their clinical introduction. *Neuropsychiatr Dis Treat* 1:329–343.

New JS (1990) The discovery and development of buspirone: a new approach to the treatment of anxiety. *Med Res Rev* 10:283–326.

Sieghart W (1992) $GABA_A$ receptors: ligand gated Cl^- channels modulated by multiple drug-binding sites. *Trends Pharmacol Sci* 13:446–450.

Sternbach L (1979) The benzodiazepine story. *J Med Chem* 22:1–7.

Antipsychotics

Bizo L, Carli M, Scarselli M & Longoni B (2025) Overview of novel antipsychotic drugs: state of the art, new mechanisms, and clinical aspects of promising compounds. *Biomedicines* 13:85.

Layton ME, Kern JC, Hartingh TJ et al. (2023) Discovery of MK-8189, a highly potent and selective PDE10A inhibitor for the treatment of schizophrenia. *J Med Chem* 66:1157–1171.

Menniti FS, Chappie TA & Schmidt CJ (2021) PDE10A inhibitors—clinical failure or window into antipsychotic drug action? *Front Neurosci* 14:600178.

Moghaddam B & Javitt D (2012) From revolution to evolution: the glutamate hypothesis of schizophrenia and its implication for treatment. *Neuropsychopharmacology* 37:4–15.

Rowley M, Bristow LJ & Hutson PH (2001) Current and novel approaches to the drug treatment of schizophrenia. *J Med Chem* 44:477–501.

Pain and Migraine

Bagal SK, Chapman ML, Marron BE et al. (2014) Recent progress in sodium channel modulators for pain. *Bioorg Med Chem Lett* 24:3690–3699.

Chen R & Chung S-H (2013) Complex structures between the N-type calcium channel (Cav2.2) and ω-conotoxin GVIA predicted by molecular dynamics. *Biochemistry* 52:3765–3772.

de Lera Ruiz M & Kraus RL (2015) Voltage-gated sodium channels: structure, function, pharmacology, and clinical indications. *J Med Chem* 58:7093–7118.

Han S, Thatte J, Buzard DJ & Jones RM (2013) Therapeutic utility of cannabinoid receptor type 2 (CB2) selective agonists. *J Med Chem* 56:8224–8256.

Hansen-Tfelt P, De Vries P & Saxena PR (2000) Triptans in migraine: a comparative review of pharmacology, pharmacokinetics and efficacy. *Drugs* 60:1259–1287.

Maryanoff BE (2016) Phenotypic assessment and the discovery of topiramate. *ACS Med Chem Lett* 7:662–665.

Max MB & Stewart WF (2008) The molecular epidemiology of pain: a new discipline for drug discovery. *Nat Rev Drug Discov* 7:647–658.

Neurodegenerative

Atta-ur-Rahman & Choudhary MI (eds) (2014) Enzyme inhibitors involved in the treatment of Alzheimer's disease. In *Drug Design and Discovery in Alzheimer's Disease*, pp 142–199. Bentham Science.

Connolly BS & Lang AE (2014) Pharmacological treatment of Parkinson disease: a review. *J Am Med Assoc* 311:1670–1683.

Oehlrich D, Prokopcova H & Gijsen HJM (2014) The evolution of amidine-based brain penetrant BACE1 inhibitors. *Bioorg Med Chem Lett* 24:2033–2045.

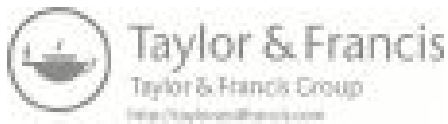
Case Studies

Dawson LA & Watson JM (2009) Vilazodone: a 5-HT1A receptor agonist/serotonin transporter inhibitor for the treatment of affective disorders. *CNS Neurosci Ther* 15:107–117.

Heinrich T, Böttcher H, Gericke R et al. (2004) Synthesis and structure–activity relationship in a class of indolebutylpiperazines as dual 5-HT_{1A} receptor agonist/serotonin reuptake inhibitors. *J Med Chem* 47:4684–4692.

Sugimoto H, Iimura Y, Yamanishi Y & Yamatsu K (1995) Synthesis and structure-activity relationships of acetylcholinesterase inhibitors: 1-benzyl-4-[(5,6-dimethoxy-1-oxoindan-2-yl)methyl]piperidine-hydrochloride and related compounds. *J Med Chem* 38:4821–4829.

Sugimoto H, Ogura H, Arai Y et al. (2002) Research and development of donepezil hydrochloride, a new type of acetylcholinesterase inhibitor. *Jpn J Pharmacol* 89:7–20.


Taylor & Francis
Taylor & Francis Group

Index

Note: Page numbers in *italic* indicate a figure and page numbers in **bold** indicate a table on the corresponding page.

A

absorption, 123–125
acetylcholine, 443–445
allosteric modulators, positive, 423–426
allosteric site
  HIV drugs and, 350–351
  polymerase inhibitors and, 217–218
allylamines, 365–367
Alzheimer's disease, 443–447
aminoglycosides, 394–399
amyloid plaques, 445–447
analogs
  advances in total synthesis and, 10–12
  combinatorial libraries and, 32–35
angiogenesis, 320–321
animals, (*see* higher animals)
antibacterial drugs, 380–381, *381*
  bacterial cell membrane permeability, 404–405, *404*
  bacterial cell wall synthesis inhibitors, 384–393, *384–393*
  bacterial DNA replication, 399–404, *400–403*
  bacterial protein synthesis inhibitors, 394–399, *394–399*
  discovery, 381–384, *382*, **383**
  introduction, 379
anti-cancer drugs, 206–211, 299–301, *300–301*
  DNA and, 278–284
    drugs targeting DNA replication and mitosis, 301–313, *301–302*, *304–310*, *312*
  drugs targeting hormone-dependent tumors, 327–329, *328*
  drugs targeting epigenetic processes, 324–327, *325–327*
  drugs targeting oncogenes and signaling pathways, 313–320, *314*, *316–319*
  drugs targeting ubiquitin–proteasome pathway, 321–324, *321–324*
  inhibition of angiogenesis, 320–321, *320*
anticonvulsants, 423–426
antidepressants, 426–432, 439–440
antifungal agents, 362–364, *363*, **364**
  drug discovery, 364
  drugs targeting fungi, 364–370, *364–369*
  introduction, 337
anti-infective agents, 258–260
  bacterial ribosomal RNA and, 284–285
anti-inflammatory agents, 206–211
  cyclooxygenase inhibitors as, 226–228
  nonsteroidal, 439–440
antiparasitic drugs, 405
  discovery, 405–412, **406**, *407–411*
  introduction, 379
antipsychotics, 433–438
antisense therapies, 285–288
anti-tumor compounds, 311–313
antiviral agents, 338–340, *338–339*
  drug discovery, 340–341
  introduction, 337
  targeting viral life cycle, 341–362, *342–346*, **342**, *348–362*
anxiety, 422–432, *423–426*, *428–432*
aqueous media, 89–90
artemisinin
  discovery of, 16, *16*
artificial intelligence (AI), 39–42
assays, (*see also* screening)
  hit and lead identification process, 23–24
  in vitro, 133–134
  in vivo, 134–135
atogepant, 140
azoles, 365–367

B

bacterial cell wall membrane permeability, 404–405
bacterial cell wall synthesis inhibitors, 384–393
bacterial DNA replication, 399–404
bacterial infections, 214–216
bacterial ribosomal RNA, 284–285
bifunctional degraders, 249–252
binding, 55–56
  conformational constraints and, 103–105
  DNA and, 276–277, 281–284
  histone deacetylase and, 326–327
  kinase inhibitors and, 206–211
  polymerase inhibitors and, 217–218
  RNA and, 288–289
  specific interactions and, 56–58
  steric bulk and, 93–94
  strategy to prevent binding between two proteins, 242–248
  therapeutic effects, 253–258
bioactivity
  hits and leads and, 24–30
bioisosteres/biososterism, 94–100
  scaffold hopping and, 100–101
biosynthesis, ergosterol, 365–367
biological activity
  lead optimization process and, 82–83
biological properties
  physical and pharmaceutical properties optimized in parallel, 83–86
biosynthesis
  DNA building blocks, of, 306–311
  folic acid, 402–404
β-lactams, 385–393

C

cancer, (*see* anti-cancer drugs)
cannabinoids, 453–454
cardiovascular disease, 214–216
  phosphodiesterase inhibitors and, 219–223
cell division, 311
cell replication, 367
central nervous system (CNS)
  abuse, 450–454, *450–454*
  anxiety and depression, 422–432, *423–426*, *428–432*
  donepezil, 456–458, *457–458*
  introduction, 419–422, *420–421*, **422**
  neurodegeneration, 443–450, *444–449*
  pain, 439–443, *439–443*
  psychosis, 433–438, *433–438*
  vilazodone, 455–456, *455–456*
chemically induced proximity (CIP), 248–253
cholesterol levels
  HMG-CoA reductase and, 225–226
clinical observations
  unexpected activities of known bioactive compounds and, 29–30
clinical trials in humans, 137–138
combinatorial libraries, 32–35
complex structures
  identification of, 8–10
compound libraries
  methods for screening, 35–42
computational algorithms, 39–42
computational methods, modern
  QSAR, for, 62–63
conformational constraints, 103–105
COVID-19, 109–111, *109–111*, **109**
cyclooxygenases, 224–228

D

daclatasvir, 371–372, *371–372*
defined groups
  quantifying structure-activity relationships, 60–62
delivery systems, 311–313
depression, 422–432, *423–426*, *428–432*
disease
  molecular causes of, 13–14
distribution, 125–126
diversity libraries, 31–32
DNA
  as drug target, 273–284, **274**, *275*, *277–284*, **278**
  drugs targeting bacterial DNA replication, 399–404
  introduction, 271–273, *272*
donepezil, 456–458, *457–458*
dopamine levels, 447–450
dopamine receptor antagonists, 433–436
doxorubicin, 290–292
drug development
  clinical trials in humans, 137–138
  drug toxicity determination, 132–136, **135–136**

drug development (*Continued*)
introduction, 117–118, *118*
patenting, 119–120
pharmacokinetic properties of leads, 122–132, *124–129*, **130**, *131–132*
process chemistry research, 120–122, *121–122*
drug discovery
antibacterial, 381–384
antifungal, 364
antiparasitic, 405–412
daclatasvir, 371–372, *371–372*
hit and lead identification, 23–24
compounds with known bioactivity, 24–30, *25–26*, *28–29*
hit triage, 42–44, *43–44*
introduction, 21
overview of process, 21–23, *22*
screening compound libraries, 35–42, *36–39*, *41–42*
synthetic libraries, 30–35, *31–32*, *34*
linezolid, 413–415, *413–414*
mipomersen, 292–293
overview
introduction, 3
medicinal chemistry, historical perspective, 3–8, *5*, *7–8*
medicinal chemistry, modern era, 8–14, *9–14*
summary, 14–15
patenting and, 119–120
rilpivirine, 372–374, *372–374*
toxicity and, 133–134
vepdegestrant, 261–263, *261–263*
vorinostat, 330, *330*
drugs of abuse, 450–454
drug–target interactions
introduction, 53–54, *54*
optimization of, 54–58, **56**, *58*
structure–activity relationships, 59–70, *59–60*, *62*, *64–67*, *69–70*
summary, 70–71
drug targets, 151–152, (*see also* drug–target interactions)
absorption and, 123–125
bacterial DNA replication, 399–404
conformational constraints and, 103–105
DNA and RNA as
DNA, 273–284, **274**, *275*, *277–284*, **278**
doxorubicin and analogs, 290–292, *290–292*
introduction, 271–273, *272*
mipomersen, 292–293
RNA, 284–289, *285–286*, *288–289*
fungi, 364–370
enzymes as
esterases, phosphodiesterases, and inhibitors, 219–223, *219–223*
grazoprevir, 230–234, *231–233*
introduction, 195–198, *196–198*
kinases and kinase inhibitor drugs, 203–212, *204–205*, *207–209*, **208**, *211–212*
oxidoreductases, cyclooxygenases, and inhibitors, 224–228, *224–227*
polymerases and polymerase inhibitors, 216–218, *217–218*
proteases and protease inhibitors, 212–216, *212–216*
small-molecule modulators, 198–203, *199–202*, **203**
tyrosine kinase inhibitor imatinib, 229–230, *229–230*
G protein-coupled receptors, 152–168, *152–153*, **154**, *155–156*, *158–167*
ion channels and, 168–177, *169–177*, **171**, **175**
lipophilicity and, 88
nuclear receptors and, 177–183, *178–183*, **179**
prodrugs and, 90–93
protein–protein interactions, 239–241, **240**
modulating, 241–248, *243–257*
ramelteon, 188–189, *188–189*
rosiglitazone, 189–190, *189–190*
structure-based drug design and, 68–70
transporters and, 183–187, *184–187*, **185**
drug toxicity, 132–136

E

echinocandins, 367–370
elimination, 126–132
endonuclease, 359–362
enzyme inhibitors, 101–103
enzymes
drug targets, as,
esterases, phosphodiesterases, and inhibitors, 219–223, *219–223*
grazoprevir, 230–234, *231–233*
introduction, 195–198, *196–198*
kinases and kinase inhibitor drugs, 203–212, *204–205*, *207–209*, **208**, *211–212*
oxidoreductases, cyclooxygenases, and inhibitors, 224–228, *224–227*
polymerases and polymerase inhibitors, 216–218, *217–218*
proteases and protease inhibitors, 212–216, *212–216*
small-molecule modulators, 198–203, *199–202*, **203**
tyrosine kinase inhibitor imatinib, 229–230, *229–230*
epigenetic processes, 324–327
erectile dysfunction
phosphodiesterase inhibitors and, 219–223
ergosterol biosynthesis, 365–367
erythromycins, 394–399
esterases, 219–223

F

false positives, 43–44
focused library, 47–48, *47–48*
folic acid biosynthesis, 402–404
formulation, 136
fragment-based screening, 37–39, 46, *46*
free radicals, 281–284
fungi, drugs targeting, 364–370

G

GABA receptor, 423–426
genetic diseases and infections, 285–288
glutamate, 436–438, 443–445
glycopeptide antibacterial agents, 392–393
grazoprevir, 230–234

H

hallucinogens, 452–453
helminthic diseases, 411–412
hepatitis C virus (HCV)
NS5A replication complex inhibitor, 357–358
NS3/4A protease inhibitor, 230–234
high-throughput screening (HTS), 35–37
higher animals
sources for hits, leads, and drugs, 26–27
histone deacetylase inhibitors, 326–327
historical perspective
introduction, 3
medicinal chemistry, practice of, 3–14, *5*, *7–14*
quantifying structure-activity relationships, 60–62
summary, 14–15
hit and lead identification, 23–24
compounds with known bioactivity, 24–30, *25–26*, *28–29*
hit triage, 42–44, *43–44*
introduction, 21
overview of drug discovery process, 21–23, *22*
screening compound libraries, 35–42, *36–39*, *41–42*
synthetic libraries, 30–35, *31–32*, *34*
hit triage, 42–44
HIV, 214–216, 350–351
HMG-CoA reductase, 225–226
hormone-dependent tumors, 327–329
humans, clinical trials in, 137–138

I

identification, (*see also* hit and lead identification)
imatinib, 229–230
influenza, 359–362
intercalation, 274–276
in vitro assays, 133–134
in vivo assays, 134–135
ion channels, 168–177, *169–177*, **171**, **175**
introduction, 151–152
irreversible mechanisms, 278–281

K

kinase and kinase inhibitor drugs, 203–212, *204–205*, *207–209*, **208**, *211–212*
target mutant or overexpressed kinases found in cancer cells, 313–318

L

lead modification
saruparib, 332–334
lead optimization
clinical trials in humans, 137–138
of drug–target interactions, 54–58, **56**, *58*
drug toxicity determination, 132–136, **135–136**
introduction, 53–54, *54*, 81, 117–118, *118*
lead properties, 86–88, *87*, **88**
lead structures, deficiencies in, 88–93, *89*, *91–93*
losartan, 107–108
medicinal chemistry strategies applied during, 93–106, *94*, **95**, *96–105*
patenting, 119–120
pharmacokinetic properties of leads, 122–132, *124–129*, **130**, *131–132*
process chemistry research, 120–122, *121–122*
properties evaluated during, 81–86, **85–86**
structure–activity relationships, 59–70, *59–60*, *62*, *64–67*, *69–70*
summary, 70–71
legacy libraries, 30–31
library screening
sources of hits and leads in drug discovery, 30–35
ligand-based drug design
knowledge of the pharmacophore and, 63–64
pharmacophore models widely used in, 64–68
linezolid, 413–415, *413–414*

lipids
introduction, 239–240, **240**
lipids of cell membranes as drug targets, 258–260, *259–260*
lipophilicity, 88
steric bulk and, 93–94
losartan, 107–108

M
malaria, 406–409
maraviroc, 263–266, *264–266*
medicinal chemistry
historical perspective of, 3–8
medicines isolated from plants, 4–6
semisynthetic drugs prepared from natural products, 6–8
modern era, 8–14
advances in total synthesis, 10–12
identification of complex structures, 8–10
molecular causes of disease, 12–14
strategies
applied during lead optimization, 93–106, *94*, **95**, *96–105*
introduction, 81
lead properties, optimization of, 86–88, *87*, **88**
lead structures, deficiencies in, 88–93, *89*, *91–93*
properties evaluated during lead optimization, 81–86, **85–86**
metabolic stability, 93–94
metabolism, 126–132
metrics
lead optimization process and, 86–88
microbes
sources for hits, leads, and drugs, 26–27
migraine, 441–443
mipomersen, 292–293
mitosis, 301–313
modulators, (*see* small-molecule modulators)
molecular causes of disease
medicinal chemistry based on, 13–14
molecular glues, 252–253
molecular interactions, 55–56
monoamine neurotransmitter levels, 426–432
multi-drug-resistant tuberculosis, 415–416, *415–416*
muscarinic acetylcholine receptor, 436–438

N
natural ligands
hits and leads based on, 27
natural products
advances in total synthesis and, 10–12
malaria and, 406–409
protein synthesis inhibitors based on, 394–399
semisynthetic drugs prepared from, 6–8
neurodegeneration, 443–450
neurotransmitters, 426–432
nonclassical bioisosteres, 97–100
noncovalent interactions, 276–277
nonsteroidal anti-inflammatory drugs, 439–440
nucleosides, 325–326
prodrugs and, 348–350
nuisance compounds, 43–44

O
olcegepant, 139–140
oncogenes, 313–320
opiates, 439–440
optimized lead, (*see* lead optimization)
oxazolidinones, 399
oxidoreductases, 224–228

P
pain, 439–443
parallel synthesis, 32–35
Parkinson's disease, 447–450
PARP1 inhibitor, 331–334, *331–334*
patenting, 119–120
pharmaceutical properties
optimized in parallel with biological properties, 83–86
pharmacokinetic properties
leads, of, 122–123
absorption, 123–125
distribution, 125–126
metabolism and elimination, 126–132
pharmacophore
introduction, 53–54, *54*
optimization of drug–target interactions, 54–58, **56**, *58*
structure–activity relationships, 59–70, *59–60*, *62*, *64–67*, *69–70*
summary, 70–71
virtual screening, 72–73, *72–73*
phosphodiesterases, 219–223, 436–438
physical properties
optimized in parallel with biological properties, 83–86
physicochemical parameters, 60–62
plants
medicines originally isolated from, 4–6, 24–26
polyene drugs, 364–365
polymerase, 301–306
antiviral drugs and, 344–348
positive allosteric modulators, 423–426
privileged structures, 105–106
process chemistry research, 120–122
prodrugs, 90–93, 348–350
properties optimized
introduction, 81
lead properties, optimization of, 86–88, *87*, **88**
lead structures, deficiencies in, 88–93, *89*, *91–93*
medicinal chemistry strategies applied during lead optimization, 93–106, *94*, **95**, *96–105*
properties evaluated during lead optimization, 81–86, **85–86**
proteases and protease inhibitors, 212–216
structure-based drug design and, 351–357
protein–protein interactions, 240–241
introduction, 239–240, **240**
modulating, 241–248, *243–257*
psychosis, 432–438

Q
quantitative structure-activity relationship (QSAR), 60–62
ligand-based drug design, 63–68
modern computational methods, 62–63
structure-based drug design, 63–64, 68–70
quinolone antibacterial agents, 400–402

R
RAS inhibitors, 319–320
receptors
G protein-coupled receptors, 152–168, *152–153*, **154**, *155–156*, *158–167*
introduction, 151–152
nuclear receptors, 177–183, *178–183*, **179**
remdesivir, 109–111, *109–111*, **109**
respiratory disease
phosphodiesterase inhibitors and, 219–223
retroviruses, 358–359
reverse transcriptase, 350–351
reversible noncovalent interactions, 276–277
rilpivirine, 372–374, *372–374*
rivaroxaban (Xarelto), 142–145, *142–145*
RNA
as drug target, 284–289, *285–286*, *288–289*
introduction, 271–273, *272*
rules
lead optimization process and, 86–88

S
saruparib, 331–334, *331–334*
scaffold hopping, 100–101
screening, 35–42, (*see also specific types*)
hit triage and, 42–44
sedatives, 423–426
selective ion channel blockers, 440–441
selective PAR1 inhibitor, 331–334, *331–334*
selective serotonin agonists, 441–443
semisynthetic drugs
prepared from natural products, 6–8
serotonin receptor agonists, 452–453
serotonin receptor antagonism, 435–436
signaling pathways, 313–320
small molecules
intercalation and, 274–276
RNA and, 288–289
small-molecule drugs to treat Alzheimer's disease, 443–447
small-molecule modulators, 198–203
solubility in aqueous media, 89–90
splicing, 288–289
split-pool synthesis, 32–35
SQ109, 415–416, *415–416*
steric bulk, 93–94
stimulants, 450–452
structural proteins involved in cell division, 311
structure–activity relationships (SARs), 59–60
quantitative (QSAR), 60–62
ligand-based drug design, 63–68
modern computational methods, 62–63
structure-based drug design, 63–64, 68–70
structure-based drug design
example, 74–76, *74–76*
knowledge of the pharmacophore and, 63–64
knowledge of the structure of the drug target and, 68–70
viral protease inhibitors and, 351–357
sulfonamide antibacterial compounds, 402–404
synthesis, (*see also* total synthesis)
bacterial cell wall synthesis inhibitors, 384–393
bacterial protein synthesis inhibitors, 394–399
bacterial ribosomal RNA and, 284–285
cell wall, 367–370, 392–393
grazoprevir, of, 233–234
parallel synthesis, 32–35
protein synthesis inhibitors, 394–399
split-pool synthesis, 32–35
ubrogepant, 140–142
vepdegestrant, 262

synthetic libraries
  sources of hits and leads in drug discovery, 30–35

T

Taxol
  discovery of, 16–17, *17*
telcegepant, 139–140
tetracyclines, 394–399
topoisomerase, 301–306
total synthesis
  access to natural products and analogs, 10–12
toxicity, (*see* drug toxicity)
transition state mimetics, 101–103
transporters, 183–187, *184–187*, **185**
  introduction, 151–152
trypanosomal diseases, 409–411
tuberculosis, 415–416, *415–416*
tyrosine kinase inhibitor imatinib, 229–230

U

ubiquitin–proteasome pathway, 321–324
ubrogepant (Ubrelvy), 139–142, *139–141*

V

vepdegestrant, 261–263, *261–263*
veterinary medicine, 411–412
vilazodone, 455–456, *455–456*
viral life cycle, 341–362
  active site of polymerases, 344–348
  allosteric sites, 350–351
  inhibitors of hepatitis C virus NS5A replication complex, 357–358
  prodrugs, 348–350
  structure-based drug design, 351–357
  uncoating of viral particles, 343–344
  viral attachment or entry, 341–343
  viral budding and endonuclease, 359–362
  viral integrase, 358–359
virtual screening (VS), 39–42
  pharmacophore model in, 72–73, *72–73*
vorinostat, 330, *330*

Z

zinc, 326–327