

(<https://educate.elsevier.com/>)

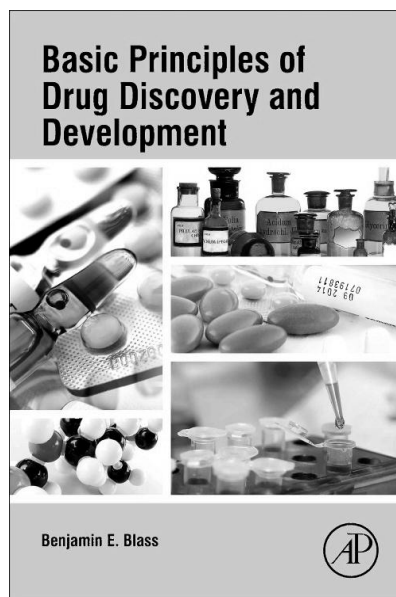
Search Elsevier Educate

Read now

Share

Update librarian

More



Purchase textbook

(<https://www.elsevier.com/books/catalog/isbn/9780124115088>)

8)

Basic Principles of Drug Discovery and Development, Edition 1

By Benjamin E. Blass

Publication Date: 27 Apr 2015

Student ancillaries

0 Reviews

Description **Table of Contents** Title details Student Resources Audience

- Dedication
- Foreword
- Chapter 1. Drug Discovery and Development: An Overview of Modern Methods and Principles

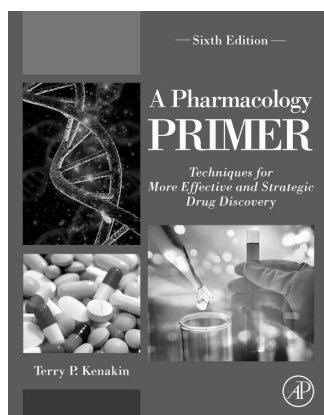
- Drug Discovery and Development from 20,000 Feet (Figure 1.7)
- Target Selection: The First Step Forward
- Hit Identification: Finding a Starting Point
- Identify a Clinical Candidate: Juggling the Properties
- Questions
- Chapter 2. The Drug Discovery Process: From Ancient Times to the Present Day
 - The Age of Botanicals: Preindustrial Drug Discovery
 - Paul Ehrlich: The Father of Modern Drug Discovery
 - Milestones in Drug Discovery
 - The Rise of Biologics and Macromolecular Therapeutics
 - Societal and Governmental Impacts
 - Regulatory Milestones
 - Future Developments in Drug Discovery
 - Questions
- Chapter 3. Classical Targets in Drug Discovery
 - Protein Structure
 - Enzymes
 - Inhibition of Enzymes
 - G-Protein-Coupled Receptors (GPCRs)
 - Ion Channels
 - Membrane Transport Proteins (Transporters)
 - Emerging Targets
 - Questions
- Chapter 4. *In vitro* Screening Systems
 - The Language of Screening: Basic Terms
 - Streptavidin and Biotin
 - Biochemical versus Cellular Assays
 - Assay Systems and Methods of Detection
 - Radioligand Assay Systems
 - Enzyme-Linked Immunosorbent Assay (ELISA)
 - Fluorescence-Based Assay Systems
 - Reporter Gene Assays
 - Kinetic Fluorescent Measurement Systems
 - Label-Free Assay Systems
 - Electrophysiological Patch Clamp
 - General Consideration for All Screening Methods
 - Questions
- Chapter 5. Medicinal Chemistry
 - Structure–Activity Relationships and Structure–Property Relationships
 - The Role of Chirality

- Push and Pull in structure–activity relationships
- Quantitative Structure–Activity Relationships
- The Pharmacophore
- Developing an SAR Data Set
- The Structure–Activity Relationship Cycle
- Bioisosterism
- Structure–Activity Relationship, Selectivity And Physicochemical Properties
- “Druglike” Guidelines
- Questions
- Chapter 6. *In vitro* ADME and *In vivo* Pharmacokinetics
 - Absorption
 - Distribution
 - Elimination Pathways
 - *In vitro* ADME Screening Methods
 - *In Vivo* Pharmacokinetics
 - Species selection
 - Questions
- Chapter 7. Animal Models of Disease States
 - Sources of Animal Models
 - Validity of Animal Models
 - Species Selection
 - Number of Animals
 - Exemplary Animal Models by Disease Category
 - Animal Models of Infectious Disease
 - Animal Models of Oncology
 - Questions
- Chapter 8. Safety and Toxicology
 - Sources of Toxicity
 - Acute versus Chronic Toxicity
 - Cytotoxicity
 - Carcinogenicity, Genotoxicity, and Mutagenicity
 - Drug–Drug Interactions
 - Cardiovascular Safety and Toxicology Studies
 - Central Nervous System Safety and Toxicology Studies
 - Immune System Mediated Safety Issues
 - Teratogenicity
 - *In Vivo* Toxicity and Safety Studies
 - Questions
- Chapter 9. Basics of Clinical Trials
 - Before the Clinic

- Drug Supply
- Delivery Methods
- Formulation
- Investigational New Drug Application
- Phase I Clinical Trials
- Phase II Clinical Trials
- Phase III Clinical Trials
- Phase IV Clinical Trials
- Adaptive Clinical Trial Design
- Questions
- Chapter 10. Translational Medicine and Biomarkers
 - Definition of a Biomarker and Their Classification
 - Characteristics and Impact of Biomarkers
 - Biomarkers versus Surrogate End Points
 - Imaging Technologies
 - The Practical Application of Biomarkers
 - Questions
- Chapter 11. Organizational Considerations and Trends in the Pharmaceutical Industry
 - Organizational Structures of Pharmaceutical Companies
 - Business Divisions Interactions
 - The Discovery Project Team Evolutionary Cycle
 - The Business Climate
 - Mergers and Acquisitions
 - Contract Research Organizations
 - Academic Drug Discovery
 - Funding Issues
 - Questions
- Chapter 12. Intellectual Property and Patents in Drug Discovery
 - Patentable Subject Matter
 - Inherent Properties and Patentability
 - Novelty and the Prior Art
 - Obviousness and the Prior Art
 - Inventorship
 - Assignment and Ownership
 - Classification of Patents and Patent Applications
 - Impact of Overlapping Patents
 - Patent Applications and their Contents
 - Contents of a Patent Application
 - Questions
- Chapter 13. Case Studies in Drug Discovery

- Tamiflu: From Mechanism of Action to Marketed Drug
- Histone Deacetylase Inhibitors: Physicochemical Optimization via Structural Change
- HIV Protease Inhibitors: Chemically Complex Miracle Drugs
- Nitrofurantoin: A Surprisingly Successful Drug
- Seldane® (Terfenadine) versus Allegra® (Fexofenadine): Metabolism Matters: Safety
- Claritin® (Loratadine) versus Clarinex® (Desloratadine): Metabolism Matters: Pharmacokinetics
- MPTP: Parkinson's Disease in a Bottle
- Bupropion and Methylphenidate: Improving Performance via Formulation Changes
- Selective Inhibition of Cyclooxygenase-2: The Impact of an Inadequate Written Description
- Questions
- Answers to Questions in Textbook by Chapter
- Subject Index
- Drug Index

Related Titles

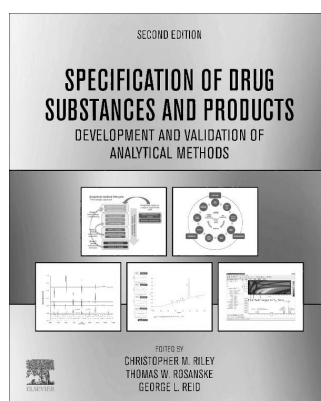


(<https://educate.elsevier.com/book/details/9780323992893>)

A Pharmacology
Primer
(<https://educate.elsevier.com/book/details/9780323992893>)

Kenakin

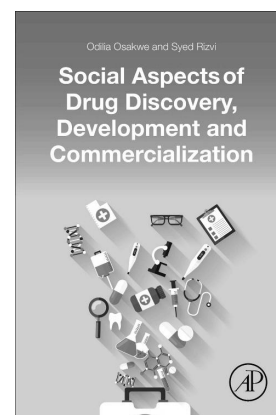
New Edition



(<https://educate.elsevier.com/book/details/9780081028247>)

Specification of Drug
Substances and
Products
(<https://educate.elsevier.com/book/details/9780081028247>)

Riley, Rosanske & Reid



(<https://educate.elsevier.com/book/details/9780280222207>)

Social Aspects of Drug
Discovery,
Development and
Commercialization
(<https://educate.elsevier.com/book/details/9780280222207>)

Osakwe & Rizvi

Basic Principles of Drug Discovery and Development



Benjamin E. Blass



Blass, Benjamin E.. Basic Principles of Drug Discovery and Development, Elsevier Science & Technology, 2015. ProQuest
Ebook Central, <http://ebookcentral.proquest.com/lib/upenn-ebooks/detail.action?docID=2036170>.
Created from upenn-ebooks on 2024-10-23 14:34:14.

BASIC PRINCIPLES OF DRUG DISCOVERY AND DEVELOPMENT

Copyright © 2015, Elsevier Science & Technology. All rights reserved.

Blass, Benjamin E.. Basic Principles of Drug Discovery and Development, Elsevier Science & Technology, 2015. ProQuest Ebook Central, <http://ebookcentral.proquest.com/lib/upenn-ebooks/detail.action?docID=2036170>.
Created from upenn-ebooks on 2024-10-23 14:34:14.

This page intentionally left blank

Copyright © 2015. Elsevier Science & Technology. All rights reserved.

Blass, Benjamin E.. Basic Principles of Drug Discovery and Development, Elsevier Science & Technology, 2015. ProQuest Ebook Central, <http://ebookcentral.proquest.com/lib/upenn-ebooks/detail.action?docID=2036170>.
Created from upenn-ebooks on 2024-10-23 14:34:14.

BASIC PRINCIPLES OF DRUG DISCOVERY AND DEVELOPMENT

BENJAMIN E. BLASS

Temple University School of Pharmacy

Moulder Center for Drug Discovery Research

North Broad Street

Philadelphia, PA, USA



AMSTERDAM • BOSTON • HEIDELBERG • LONDON
NEWYORK • OXFORD • PARIS • SAN DIEGO
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO

Academic Press is an imprint of Elsevier



Academic Press is an imprint of Elsevier
32 Jamestown Road, London NW1 7BY, UK
525 B Street, Suite 1800, San Diego, CA 92101-4495, USA
225 Wyman Street, Waltham, MA 02451, USA
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK

Copyright © 2015 Elsevier Inc. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

ISBN: 978-0-12-411508-8

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

For information on all Academic Press publications
visit our website at <http://store.elsevier.com/>

Printed and bound in the USA



Working together
to grow libraries in
developing countries

www.elsevier.com • www.bookaid.org

Dedication

Sir Isaac Newton, one of the greatest scientists of his time, wrote “If I have seen further it is by standing on ye shoulders of Giants.” Although he was almost certainly referring to his scientific achievements, the underling concept of learning from our forbearer is true in any endeavor. Indeed, this concept can be further extended to include those who are there in the present day, supporting the activities of an individual as he or she attempts to accomplish that which they view as important. With this thought in mind, I have dedicated this book to the scientists who came before me, those who mentored me, and those who work with me on a daily basis. In addition, and perhaps more importantly, this text is dedicated to the loving and supportive family that has helped me become the person that I am today. Special thanks are offered to my mother, father, sister, brother, my three children, and of course, my wife Kathleen. These are the giants on whose shoulders I have stood upon.

This page intentionally left blank

Copyright © 2015. Elsevier Science & Technology. All rights reserved.

Blass, Benjamin E.. Basic Principles of Drug Discovery and Development, Elsevier Science & Technology, 2015. ProQuest Ebook Central, <http://ebookcentral.proquest.com/lib/upenn-ebooks/detail.action?docID=2036170>.
Created from upenn-ebooks on 2024-10-23 14:35:13.

Contents

Foreword	xiii
1. Drug Discovery and Development: An Overview of Modern Methods and Principles	1
Drug Discovery and Development from 20,000 Feet	8
Target Selection: The First Step Forward	11
Hit Identification: Finding a Starting Point	15
Identify a Clinical Candidate: Juggling the Properties	21
Questions	26
References	27
2. The Drug Discovery Process: From Ancient Times to the Present Day	35
The Age of Botanicals: Preindustrial Drug Discovery	36
Paul Ehrlich: The Father of Modern Drug Discovery	39
Milestones in Drug Discovery	41
Milestones in Animal Models: Breeding a Better Model	42
The Wistar Rat	42
Immunocompromised Mice	43
Transgenic Animal Models	44
Knockout Animal Models	45
Milestones in Molecular Science	47
X-ray Crystallography	48
Molecular Modeling and Computational Chemistry	50
High Throughput Technology: Chemical Synthesis and Screening Science	53
Milestones in Biotechnology	59
Recombinant DNA and Transfection Technology	60
Polymerase Chain Reaction (PCR) Technology	62
Monoclonal Antibody and Hybridoma Technology	64
The Rise of Biologics and Macromolecular Therapeutics	66
Societal and Governmental Impacts	67
The Pure Food and Drug Act of 1906	68
The Elixir of Sulfanilamide Disaster of 1937	69
The Thalidomide Story	71

Regulatory Milestones	73
Durham–Humphrey Amendment of 1951	74
Kefauver–Harris Amendment of 1962	74
Hatch–Waxman Act of 1984	75
Biologics Price Competition and Innovation Act of 2009	77
Future Developments in Drug Discovery	78
Questions	78
References	79
 3. Classical Targets in Drug Discovery	 87
Protein Structure	89
Enzymes	95
Inhibition of Enzymes	101
G-Protein-Coupled Receptors (GPCRs)	105
G-Protein-Dependent Signaling Pathways	109
cAMP Signaling	110
IP ₃ Signaling	111
Modulating GPCR Activity	112
Ion Channels	116
Gating Mechanisms	120
Ligand-Gated Channels	120
Voltage-Gated Channels	123
Other Gating Mechanisms	126
Membrane Transport Proteins (Transporters)	126
Emerging Targets	134
Questions	135
References	136
 4. <i>In vitro</i> Screening Systems	 143
The Language of Screening: Basic Terms	144
Concentration Response Curves and IC ₅₀ s	145
Dissociation Constants (K _d) and Inhibition Constants (K _i)	146
Efficacy versus Binding: EC ₅₀ s	148
Agonist, Partial Agonist, Antagonist, Allosteric Modulators, and Inverse Agonists	149
Agonists and Partial Agonists	150
Antagonists	151
Basal Activity and Inverse Agonists	151
Allosteric Modulation	151
Receptor Reserve	152
Streptavidin and Biotin	153
Biochemical versus Cellular Assays	154
Assay Systems and Methods of Detection	156
Radioligand Assay Systems	157
Scintillation Proximity Assay (SPA)	159
Enzyme-Linked Immunosorbent Assay (ELISA)	162
Fluorescence-Based Assay Systems	164

Fluorescence Polarization (FP)	165
Fluorescence Resonance Energy Transfer (FRET)	168
Time-Resolved Fluorescence Resonance Energy Transfer (TRFRET)	172
Amplified Luminescent Proximity Homogeneous Assay (AlphaScreen™)	175
Fluorescent Detection of Calcium Flux	177
Reporter Gene Assays	180
Chloramphenicol Acetyltransferase (CAT)	181
β-Lactamase Reporter Assays	182
Luciferase Reporter Assays	183
Kinetic Fluorescent Measurement Systems	184
Label-Free Assay Systems	185
Cellular Dielectric Spectroscopy	186
Optical Biosensors	187
Surface Plasmon Resonance Technology	190
Electrophysiological Patch Clamp	191
General Consideration for All Screening Methods	194
Questions	196
References	197
 5. Medicinal Chemistry	 203
Structure–Activity Relationships and Structure–Property Relationships	204
The Role of Chirality	209
Push and Pull in Structure–Activity Relationships	212
Quantitative Structure–Activity Relationships	213
The Pharmacophore	218
Developing an SAR Data Set	220
The Structure–Activity Relationship Cycle	231
Bioisosterism	232
Structure–Activity Relationship, Selectivity and Physicochemical Properties	236
“Druglike” Guidelines	237
Questions	239
References	240
 6. <i>In vitro</i> ADME and <i>In vivo</i> Pharmacokinetics	 245
Absorption	249
Solubility	250
Permeability	254
Distribution	262
Permeability	264
Transporters	266
Plasma Protein Binding	268
Elimination Pathways	270
Metabolism	271
Excretion	283
<i>In vitro</i> ADME Screening Methods	287

<i>In vivo</i> Pharmacokinetics	289
Volume of Distribution	291
Clearance	293
Half-life	294
Bioavailability	296
Species Selection	299
Questions	300
References	301
7. Animal Models of Disease States	307
Sources of Animal Models	309
Validity of Animal Models	312
Species Selection	313
Number of Animals	314
Exemplary Animal Models by Disease Category	314
Animal Models in Neuroscience	315
The Forced Swimming Test: A Model of Depression	315
The Elevated Plus Maze: A Measure of Anxiety	315
The Novel Object Recognition Test: A Model of Memory and Cognition	316
Contextual Fear Conditioning Model: A Model of Contextual Learning	317
The Morris Water Maze: A Model of Spatial Learning and Memory	318
Animal Models of Neurodegeneration	320
The SOD1-G93A Mouse of Amyotrophic Lateral Sclerosis	320
The MPTP Model of Parkinson's Disease	321
Animal Models of Cardiovascular Disease	323
Models of Hypertension	323
Models of Hyperlipidemia and High Cholesterol	325
Models of Atrial Fibrillation	327
Models of Heart Failure	329
Animal Models of Infectious Disease	332
Murine Thigh Infection Model	332
Murine Model of Systemic Infection	333
The Mouse Model of Influenza Virus Infection	334
Limitations of Animal Models of Infection	335
Animal Models of Oncology	335
Mouse Xenograft Tumor Model	336
Mouse Allograft Tumor Model	337
Genetically Engineered Mouse Models of Cancer	338
Questions	339
References	340
8. Safety and Toxicology	345
Sources of Toxicity	347
Acute versus Chronic Toxicity	353
Cytotoxicity	354

Carcinogenicity, Genotoxicity, and Mutagenicity	356
Drug–Drug Interactions	360
Cardiovascular Safety and Toxicology Studies	362
Central Nervous System Safety and Toxicology Studies	369
Immune System Mediated Safety Issues	371
Teratogenicity	373
<i>In vivo</i> Toxicity and Safety Studies	375
Questions	376
References	377
9. Basics of Clinical Trials	383
Before the Clinic	386
Drug Supply	386
Delivery Methods	389
Formulation	391
Investigational New Drug Application	395
Phase I Clinical Trials	397
Phase II Clinical Trials	399
Phase III Clinical Trials	402
Phase IV Clinical Trials	407
Adaptive Clinical Trial Design	409
Questions	410
References	410
10. Translational Medicine and Biomarkers	415
Definition of a Biomarker and Their Classification	418
Characteristics and Impact of Biomarkers	421
Biomarkers versus Surrogate End Points	423
Imaging Technologies	424
The Practical Application of Biomarkers	432
DPP-IV Inhibitors (Januvia®)	432
Physiological Measurements as Biomarkers: Orexin Antagonists	434
FDG PET Imaging Agent	436
The Neurokinin 1 (NK ₁) Receptor, Depression, and PET Imaging: The Aprepitant Story	438
Questions	440
References	440
11. Organizational Considerations and Trends in the Pharmaceutical Industry	447
Organizational Structures of Pharmaceutical Companies	448
Business Divisions Interactions	448
The Discovery Project Team Evolutionary Cycle	449
The Business Climate	452
Mergers and Acquisitions	453
Contract Research Organizations	457

Academic Drug Discovery	460
Funding Issues	465
Questions	467
References	468
12. Intellectual Property and Patents in Drug Discovery	471
Patentable Subject Matter	473
Inherent Properties and Patentability	477
Novelty and the Prior Art	479
Obviousness and the Prior Art	480
Inventorship	483
Assignment and Ownership	485
Classification of Patents and Patent Applications	486
Impact of Overlapping Patents	487
Patent Applications and Their Contents	488
Contents of a Patent Application	492
Questions	497
References	497
13. Case Studies in Drug Discovery	499
Tamiflu: From Mechanism of Action to Marketed Drug	499
Histone Deacetylase Inhibitors: Physicochemical Optimization via Structural Change	504
HIV Protease Inhibitors: Chemically Complex Miracle Drugs	506
Nitrofurantoin: A Surprisingly Successful Drug	511
Seldane® (Terfenadine) versus Allegra® (Fexofenadine): Metabolism Matters: Safety	513
Claritin® (Loratadine) versus Clarinex® (Desloratadine): Metabolism Matters: Pharmacokinetics	516
MPTP: Parkinson's Disease in a Bottle	517
Bupropion and Methylphenidate: Improving Performance via Formulation Changes	521
Selective Inhibition of Cyclooxygenase-2: The Impact of an Inadequate Written Description	524
Questions	526
References	526
Appendix 1	531
Answers to Questions in Textbook by Chapter	539
Subject Index	565
Drug Index	579

Companion site for this book can be accessed at

<http://booksite.elsevier.com/9780124115088>

Foreword

The last three decades have witnessed a revolution in the drug discovery and development process. Medicinal chemistry and in vitro screening that were once major bottlenecks in the process of identifying novel therapeutics have been dramatically accelerated through the incorporation of automation and the development of enabling technologies such as recombinant DNA and transfection technology. High-throughput screening (HTS), parallel synthesis, and combinatorial chemistry have facilitated the synthesis and biological evaluation of large numbers of potentially useful compounds. These activities, in turn, have generated vast amounts of data that can be analyzed to develop structure–activity relationships and structure–property relationships useful for the optimization of lead compounds. At the same time, new techniques, technological advances, and a greater understanding of the importance of pharmacokinetics, animal models, and safety studies have dramatically altered how new molecules are selected for clinical study. Design strategies of clinical trials, biomarkers, translational medicine, the regulatory landscape, intellectual property rights, and the business environment have also changed dramatically over the course of the last 30 years.

The complexities of the drug discovery and development process cannot be overstated, nor can the wide range of expertise required for the successful development of new, marketable therapeutics. In order to thrive in this every changing landscape, individuals interested in a career in the pharmaceutical industry or related fields must be more than simply experts in their chosen field of study. They must also have an understanding of the numerous, overlapping fields of their colleagues. *Basic Principles in Drug Discovery and Development* has captured the critical information on the disparate processes, technologies, and expertise required for modern drug discovery and development and presents it in a logical and concise manner for students, practicing scientists, and nonscientists with an interest in the pharmaceutical industry. Dr Benjamin E. Blass, an experienced educator and scientist with foundational knowledge in medicinal chemistry, drug design, biological targets, and over 20 years of experience in industrial and academic drug discovery and development, provides a comprehensive account of the many functions involved in drug discovery and development, from initial medicinal chemistry conceptualization and in vitro biological evaluation to clinical trials and beyond.

There are many aspects of this book that will help practicing scientists, graduate students, and future drug discovery researchers to develop a strong foundation in the concepts that govern the multidisciplinary process of drug discovery. Through this unique text, they will acquire an understanding of key aspects of drug discovery and development. The organization of the subject material was chosen to allow the readers to incrementally increase their knowledge in the wide range of disciplines required to identify new, marketable therapeutic agents. The book is thoroughly written and includes 13 chapters with over 300 figures and 900 references. Throughout the text, the reader will become familiar with more than 100 drugs and clinical candidates that exemplify important theories and practices.

Each chapter contains examples of drugs pertaining to the material in the chapter. The opening chapter provides an overview of drug discovery and development. This serves as the foundation for the following 10 chapters which describe the various functions involved in drug discovery and development. The early phases of drug discovery are described in detail through discussions of important topics such as target identification, target validation, lead identification, multidimensional lead optimization, pharmacokinetics, preclinical pharmacodynamics, and early toxicology. This is followed by discussions of preclinical activities, clinical trial design, biomarkers, and translational medicine. Each chapter builds on the previous chapters and this approach provides the readers with an integrated view of the various, multidisciplinary functions required for the drug discovery and development process.

Chapters 11 and 12 describe two important topics essential for running an effective pharmaceutical R&D business: organizational structure and patent protection. These chapters give the reader a true understanding of the organizational structure required for the successful management of research and development organizations and the importance of protecting intellectual property to ensure a good return on investment. Patent protection is the life blood of the pharmaceutical and biotech industries, and at the same time a source of innovation for new discoveries. Patents ensure the sharing of discoveries and innovations that might otherwise be kept as trade secrets. In the final chapter, case studies demonstrating the practical application of the concepts and principles described in the previous chapters are provided. These vignettes also describe important lesson learned in each case, some of which changed the way modern drug discovery research and development programs are executed.

Although there are numerous textbooks that discuss various aspects of the drug discovery and development process, none of them provides a comprehensive view of the process. *Basic Principles in Drug Discovery and Development* is unique in its comprehensive approach to this

complex endeavor. In writing this textbook, Dr Blass has provided an important new tool for the education of the next generation and a valuable resource for people with a vested interest in the identification and commercialization of novel medications.

Magid Abou-Gharbia, PhD, FRSC

*Associate Dean for Research,
Laura H. Carnell Professor and Director Moulder
Center for Drug Discovery Research,
School of Pharmacy, Temple University
Philadelphia, PA, USA*

This page intentionally left blank

Copyright © 2015. Elsevier Science & Technology. All rights reserved.

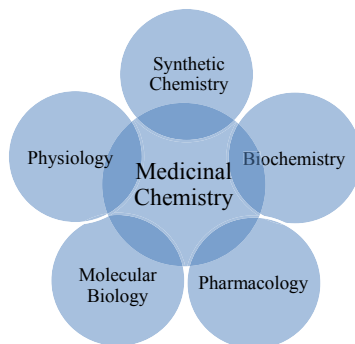
Blass, Benjamin E.. Basic Principles of Drug Discovery and Development, Elsevier Science & Technology, 2015. ProQuest Ebook Central, <http://ebookcentral.proquest.com/lib/upenn-ebooks/detail.action?docID=2036170>.
Created from upenn-ebooks on 2024-10-23 14:35:26.

Medicinal Chemistry

The ultimate goal of the vast majority of drug discovery and development programs is the identification of a single, novel compound capable of altering the outcome of a biological process. Whether the biological process is pathological, such as cancer, or a lifestyle issue, such as male pattern baldness, the end goal is essentially the same, a marketable therapeutic entity. In the previous chapters, common biochemical targets and methods of monitoring biological activity were described. While these topics are important to the process, knowledge in these areas alone is insufficient to drive a drug discovery program. If these were the only tools available, the identification of compounds with interesting biological properties would be possible, but moving beyond the initial starting set of biologically active compounds would be a difficult task. In order to move beyond initial starting points, such as literature leads, competitor therapeutics, or *in vitro* screening hits, an additional set of skills and knowledge is required. Program scientists must understand the relationship between the structures of biologically interesting compounds and their properties, both biological and physicochemical. In addition, they must have an understanding of the impact that subtle changes in chemical structure can have on these properties. In other words, knowledge of medicinal chemistry is a key component of drug discovery and development.

The American Chemical Society's division of medicinal chemistry defines medicinal chemistry as "the application of chemical research techniques to the synthesis of pharmaceuticals." This rather restrictive definition implies that the field of medicinal chemistry is the sole purview of those capable of synthesizing compounds, i.e., synthetic chemists. In reality, however, this is not the case. A more accurate description of medicinal chemistry is a field focused on understanding the chemical basis of the biological effects of compounds by integrating fundamental concepts from different fields such as synthetic chemistry, biochemistry, pharmacology, physiology, and molecular biology (Figure 5.1). Certainly, some of the team members in a drug discovery program must be well versed

FIGURE 5.1 Medicinal chemistry applies knowledge derived from synthetic chemistry, biochemistry, pharmacology, physiology, and molecular biology in an effort to understand the biological impact of candidate compounds. It is an interdisciplinary science whose primary goal is the identification of therapeutically useful compounds.



in synthetic chemistry so that new molecules can be prepared, but this is not a requirement for contribution to medicinal chemistry efforts. To be clear, an in-depth knowledge of how to make compounds for screening is not a prerequisite for being able to understand and contribute to the design aspects of medicinal chemistry, and it is the design elements of this process that enable a program to move forward toward a new therapeutic entity. This, of course, begs the question what are the design aspects of medicinal chemistry and how are they applied as a program moves forward?

STRUCTURE–ACTIVITY RELATIONSHIPS AND STRUCTURE–PROPERTY RELATIONSHIPS

A deceptively simple answer to this question begins with the premise that a compound's biological properties are a function of its chemical structure. In addition, in a given series of related compounds, structural changes across the series will lead to predictable changes in biological properties. This relationship is referred to as a structure–activity relationship (SAR). The impact of structural changes on biological properties such as binding potency, functional activity, and selectivity can be monitored in order to establish structure–activity relationships for each property. In principle, program scientists simply need to understand relationships between chemical structure and biological activity in a series of related compounds in order to identify the compound with the most promising biological properties and design compounds with improved properties.

In a similar manner, the physicochemical properties of a series of related compounds are tied to the structure of the compounds. Changes in structure will lead to changes in physicochemical properties such as solubility, membrane permeability, lipophilicity, and total polar surface area.

The relationship between a series of compounds and their physicochemical properties is referred to as structure–property relationships. In theory, one needs to only understand the relationship between chemical structure and physical properties in order to identify a compound with optimal properties.

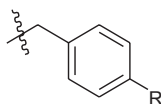
Pharmacokinetic (PK) properties such as absorption, distribution, metabolism, and excretion, all of which are useful in describing how the body deals with a compound *in vivo*, are also all a function of chemical structure. These properties will be discussed in greater detail in Chapter 6, but at this juncture, it is worth noting that changes in chemical structure will lead to changes in these properties that can be predicted using the same principles noted for structure–activity relationship and structure–property relationships. Knowledge of the impact of structural changes on these properties should provide the ability to quickly identify compounds with optimal pharmacokinetic properties.

In practice, none of these principles are as simple as they appear on paper. As discussed in earlier chapters, biologically relevant macromolecular targets are complex, three-dimensional structures whose configurations are controlled by a multitude of molecular interactions (e.g., hydrogen bonds, salt bridges, hydrophobic interactions, etc. See Chapter 3). These interactions give rise to their function and, more importantly within the context of drug discovery, afford-binding sites that are critical to their function. Binding of a ligand, natural or otherwise, requires that the ligand interact with a macromolecule at a preferred binding site, employing the same forces that control the configuration of the macromolecule as a whole. As such, in order to have a *complete* understanding of the structure–activity relationship between a target macromolecule and a series of related compounds, one must understand how these interactions change as subtle structural changes are applied across a series of potential ligands. Understanding these interactions is further complicated by the fact that macromolecules are not static entities. Binding of a ligand to a macromolecule can induce changes in the binding site that create additional favorable interactions and increase binding strength. It is also important to note that in any given chemical series, it may be possible to access hundreds, if not thousands, of possible candidate compounds. It is neither economically feasible nor practical to prepare and screen more than a small fraction of chemical space within a given series.

Fortunately, it is not necessary to have a *complete* picture of all aspects of the structure–activity relationships for a series of compounds against a given target in order to understand how one can manipulate biological properties by making subtle changes in structure across the series. A structure–activity relationship can be viewed as a jigsaw puzzle. Long before all of the pieces are put into place, it is possible to anticipate the full image that will be provided upon completion of a jigsaw puzzle. In the same sense, a small series of compounds with appropriately applied structural changes can provide a wealth of information. These data can then be used to predict

biological properties of potential candidate compounds and direct program choices toward desirable properties and away from potential dead ends. These concepts are best explained through simplified examples in which one assumes that the structure of the macromolecular target is static and that the overall structure of the series of compounds in question are identical in all respects, except in the areas indicated in each example.

Consider, for example, a class of compounds that contains a substituted aromatic ring in the same position, but are otherwise identical in structure (Figure 5.2). If the aromatic ring in question is directed into solvent space



Entry	R	IC ₅₀ (nM)	Entry	R	IC ₅₀ (nM)
1	Methoxy	10000	4	Fluoro	50
2	Methyl	1000	5	CN	20
3	Chloro	100	6	NO ₂	2

FIGURE 5.2 Altering the substituents on an aromatic ring can have a significant impact on biological activity. In this hypothetical set of compounds, potency at the biological target increases as the benzene ring becomes increasingly electron poor. These data suggest that additional electron-donating R-groups such as an NH₂ would have lower potency than R-groups that are electron withdrawing, such as a CF₃.

(that is, outside of the binding pocket), then changes in the substitution pattern of the aromatic ring would be expected to have little impact on the potency of the compounds. On the other hand, if the aromatic ring is positioned toward the binding pocket, differences in substitution patterns on the aromatic ring could have a significant impact on potency. In this hypothetical example, potency at the target is strongly influenced by the electronic nature of the R-group. Electron-donating substituents, such as methyl and methoxy, decrease potency, while electron-withdrawing groups, such as chlorine, bromine, or trifluoromethyl, increase potency. Using this structure activity information, one would conclude that other electron-donating groups in this position will not provide potent compounds, while alternative electron-withdrawing groups would be likely to provide compounds with a high degree of potency. For example, an analog wherein the R-group is a trifluoromethyl group (–CF₃), an electron-withdrawing substituent, would be predicted to be relatively potent in this assay. Importantly, if one were interested in identifying additional analogs in this series of compounds by changing other areas of the molecule, the above structure–activity relationship analysis would suggest that new compounds should contain aromatic rings with electron-withdrawing groups in the position shown.

In a similar manner, physical restrictions of a ligand-binding site can be elucidated by examining binding data. As in the previous case, consider a series of compounds that are identical, except for the presence of a variable R-group on an aromatic ring. In this case, however, rather than altering the electronic nature of the R-group, a series of increasingly sterically demanding substituents are employed (Figure 5.3). The blue line represents the outer

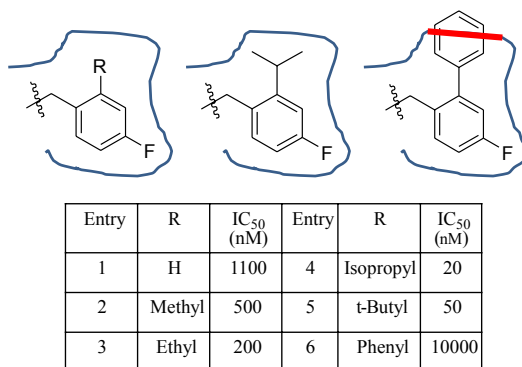


FIGURE 5.3 In any given ligand-binding site, the “walls” of the binding site are comprised of the backbone and side chain features of the macromolecular target. In this example, the outer perimeter of the binding cavity is designated by the blue line. Binding of the hypothetical ligand becomes more efficient as the R-group fills the available space and hydrophobic interaction increases (entries 2–5). Exceeding the size of the cavity (entry 6), however, will cause a dramatic loss in binding efficiency, as the candidate compound no longer fits within the allowed space (as indicated by the red line).

boundary of a hydrophobic pocket of a binding site. Increasing the size of the lipophilic substituent increases hydrophobic interactions between the ligand and the binding pocket, increasing binding strength as long as the R-group does not exceed the size of the lipophilic space. Once the ligand exceeds the size limit of the hydrophobic pocket, binding potency falls off rapidly. The compound no longer fits within the binding site, as indicated by the red line in the hypothetical binding pocket in Figure 5.3. With this information in hand, one can reasonably predict that compounds in which the R-group is an aromatic ring will be significantly less potent than those where smaller substituents are employed, irrespective of the nature of the appended aromatic ring. There is not enough room in the binding site to accommodate larger groups in this position.

The placement and orientation of functionality capable of forming hydrogen bonds can also have a significant impact on ligand binding, and an analysis of binding data within the framework of a structure-activity relationship can provide a great deal of insight when considering the design

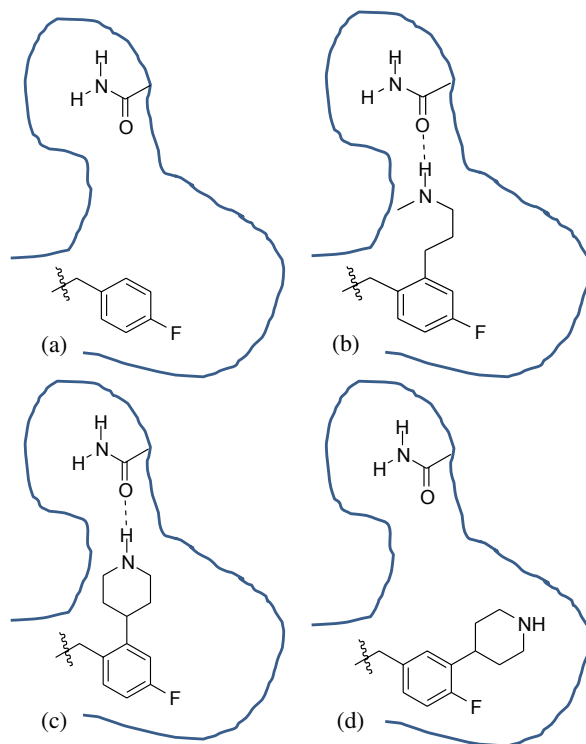


FIGURE 5.4 In this hypothetical scenario, the binding site of a protein is defined by the blue line. In (a), a 4-fluorobenzyl substituent occupies a small portion of the binding site. Addition of an amine side chain meta to the fluorine atom (b) allows the hypothetical ligand to form a hydrogen bond with an amide side chain, strengthening the binding interaction. Replacing the propyl amine side chain with a piperidine ring (c) adds rigidity to the ligand while maintaining the newly accessible hydrogen bonding interaction. Relocating the piperidine ring so that it is ortho to the fluorine atom (d), however, prevents the ligand from accessing the amide hydrogen bond partner, as the orientation of the ligand in the binding pocket no longer favors this interaction.

of next-generation compounds. Once again, a hypothetical series of compounds that are identical with the exception of the substitution pattern of an aryl ring (Figure 5.4) is considered. In this case, an amide side chain of an amino acid in the binding site is available to form a hydrogen bond with a suitably functionalized ligand. The monofluoronated species (a) is incapable of forming a hydrogen bond, and thus derives no additional binding strength from the presence of the amide side chain. If, however, the ligand also contained a properly positioned propyl amine side chain (b), the

resulting compound would be able to form the hydrogen bond indicated, strengthening the interaction between the ligand and the binding pocket. Knowledge of the increased binding energy, as indicated by IC₅₀ screening data against the target of interest, would suggest that replacing the propyl amine side chain with a piperidine (c) would provide a compound of similar potency, as both compounds would provide the framework necessary for the formation of a critical hydrogen bond. In contrast, one could predict that relocating the amine hydrogen bond donor ortho to the fluorine atom (d) would position the potential hydrogen bond donor away from the hydrogen bond acceptor (amide). No significant increase in binding strength would be expected in this scenario.

THE ROLE OF CHIRALITY

All of the aforementioned examples describe the relationship between structural changes and biological activity in a two-dimensional framework. In reality, of course, biological activity, cellular function, and life in general occur in three-dimensional space. Macromolecular targets that are the focus of drug discovery and development efforts possess complex three-dimensional structures. This forms the basis of their ability to differentiate between potential ligands, even in cases where a pair of compounds are simply mirror images of each other. As such, a discussion of structure-activity relationships must include an examination of how changes in the three-dimensional structure of a series of compounds impact biological activity.

The concept of chirality is fundamental to understanding structure-activity relationships and was originally postulated by Lord Kelvin more than 100 years ago. He stated "I call any geometrical figure, or group of points, 'chiral', and say that it has chirality if its image in a plane mirror, ideally realized, cannot be brought to coincide with itself."¹ Although this statement applies to any arrangements of materials, within the context of drug discovery the most common occurrence of chirality is a tetrahedral carbon atom with four different substituents (Figure 5.5). In this scenario, the substituents can be

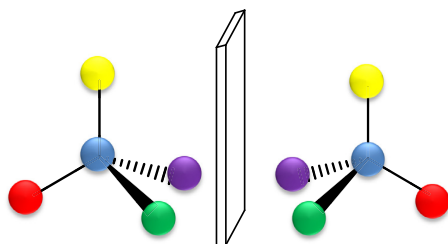


FIGURE 5.5 Chiral compounds cannot be superimposed on their mirror image. The two isomers are referred to as enantiomers (R and S). Chirality often plays a key role in biological activity, as nature is a chiral environment.

attached to the central carbon atom in one of two arrangements, which are non-superimposable mirror images of each other. The two separate isomers are referred to as enantiomers or optical isomers, and in the absence of a chiral environment, they are identical in all respects, except for the direction in which they rotate plane polarized light. Biological macromolecules, being chiral, offer an environment that predictably discriminates between enantiomers, leading to dramatic changes in the functional activity between pairs of enantiomers. Darvon² and Novrad^{2a,3} (Figure 5.6), for example, are identical

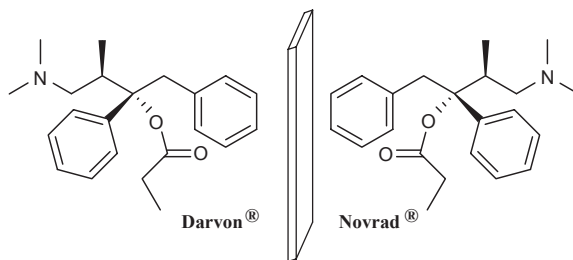


FIGURE 5.6 Darvon[®] and Novrad[®] are identical, save for the nature of the chiral center present in each. The enantiomers have very different properties in a biological setting. Darvon[®] is an analgesic that activates the μ -opioid receptor, while Novrad[®] is an antitussive agent with minimal efficacy as an analgesic. Unlike its enantiomer, Novrad[®] has little affinity for the μ -opioid receptor.

except for the difference in chirality, and this leads to substantially different biological activity. The former is a potent analgesic that acts as a μ -opioid receptor agonist,⁴ while the latter is an antitussive with minimal potency at the μ -opioid receptor.⁵ This underscores the importance of how a deceptively simple change in molecular structure, the inversion of a chiral center, can have a significant influence on therapeutic action. In fact, a single change at a chiral center is a drastic change in three-dimensional space and is certainly not viewed as a subtle change to a medicinal chemist. From the perspective of a drug discovery scientist, there is also the underlying question of how one could employ knowledge of the impact of chiral inversions in order to design better candidate compounds. In other words, what is the relationship between structural changes induced by chiral inversion and biological activity?

An understanding of the impact of chiral inversion on biological activity, and thus understanding the role of chirality in structure–activity relationships, is best accomplished through the examination of a practical example. Consider, for example, Tamiflu[®] (oseltamivir), an influenza neuraminidase inhibitor pro-drug useful for the treatment of the common flu.⁶ X-ray crystallographic studies of this compound in the binding site of influenza

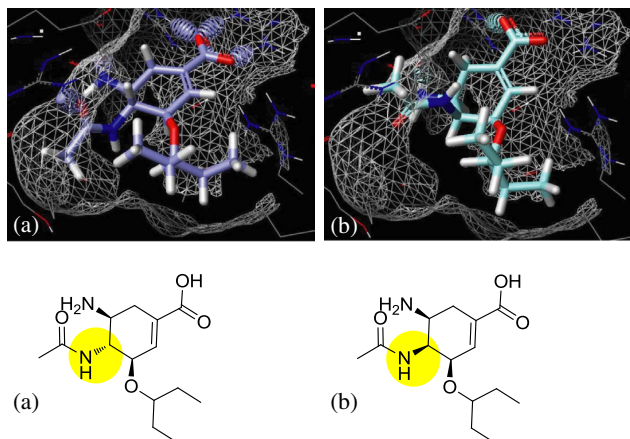


FIGURE 5.7 (a) The active agent derived from Tamiflu[®] (oseltamivir) occupies the binding site of influenza neuraminidase, the boundaries of which are delineated by the wire mesh. (b) Inversion of the amide chiral center causes the acetamide to protrude from the binding cavity in this model, demonstrating the importance of the orientation of this chiral center (RCSB 2QWK).

neuraminidase (Figure 5.7(a)) portray a tight-fitting interaction, which explains its potent impact on enzymatic activity. The translucent gray surface represents the outer edges of the binding site, marking the physical limits for potential inhibitors of this enzyme.⁷ Hypothetically, one could invert a single chiral center of this compound and produce an alternative structure that, in all other respects, is structurally identical to the active component of Tamiflu[®]. As indicated in Figure 5.7(b), however, this modification has a significant impact on the interaction of the hypothetical ligand with the binding site. Rather than fitting within the confines of the binding site, the proposed ligand protrudes through the theoretical boundary defined by the atoms of the binding site. The negative interaction produced by the “bumping” of atoms of the ligand and the target have a negative impact on binding interactions, which in turn decreases biological potency.

In the case of Tamiflu[®], knowledge of the X-ray structure and the availability of advanced protein modeling techniques can provide insight into how theoretical compounds might fit into the binding site of influenza neuraminidase, providing guidance for prioritization. If an X-ray structure is not available, however, one can still apply design principles in order to predict activity and prioritize compound preparation and screening. Comparison of *in vitro* screening data for a series of related compounds that are identical except for the orientation of a single chiral center, for example, would provide indirect knowledge of the limitation of the ligand-binding site. For the purposes of illustration, consider a series of compounds with an embedded amino acid

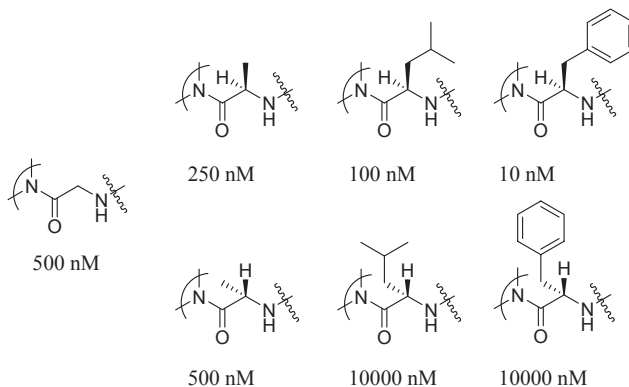


FIGURE 5.8 In a series of otherwise identical compounds, differences in chirality can have a significant impact on biological activity. Although an X-ray crystal is a valuable tool in identifying areas where changes in chirality will have the largest impact, it is not required. In this hypothetical example, potency differences between the R-series (top) and the S-series (bottom) clearly demonstrate that the R-series is heavily favored.

(Figure 5.8). When the embedded amino acid is a glycine moiety, there is no chirality at the α -carbon, as both substituents are hydrogen. Incorporation of either naturally occurring alanine or its enantiomer, however, renders the molecule chiral, as would the incorporation of the enantiomers of leucine or phenylalanine. A comparison of the screening results from this set of seven compounds indicates that the ligand-binding site strongly favors one of the two possible enantiomers. *In vitro* potency improves with increased steric bulk in one enantiomeric series (the R-series, top), but the same changes in the opposite enantiomeric series (the S-series, bottom) leads to significant loss of potency. With these screening results in hand, one could reasonably predict that compounds with the S-configuration at this chiral center would be significantly less potent than the R-series and should be deprioritized when considering potential candidate compounds.

PUSH AND PULL IN STRUCTURE–ACTIVITY RELATIONSHIPS

The examples discussed so far are all based on the assumption that changes to the candidate compound are isolated to one location within the molecular framework. In practice, however, this is not always the case. As drug discovery programs evolve through the process of identifying a hit in an assay, developing a lead series, and finally choosing a clinical candidate, multiple changes to the molecular framework of the original hit will be examined. These changes often produce similar compound classes that may have similar SAR patterns. It is important to keep in mind, however,

that “similar” is not “same.” Given two related, but ultimately different compound classes, there may be significant overlap between the SAR patterns of the two classes of compounds, but there are also likely to be differences. If, for example, a series of imidazopyridines (Figure 5.9(a)) had

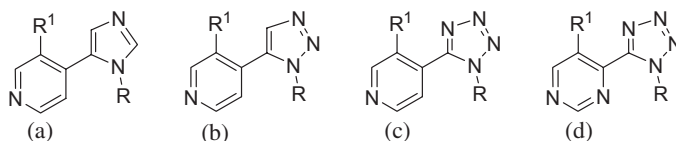


FIGURE 5.9 Compounds (a), (b), (c), and (d) have similarities that could lead to overlapping SAR, but the increasing number of nitrogen atoms present in each structure changes the character of the aromatic rings. If the aromatic rings’ sole function is structural support for the functionality required for binding, then SAR patterns may be comparable. On the other hand, if the aromatic rings are critical to binding, then changes in nitrogen counts may lead to very different SAR patterns for each series of compounds.

a known SAR pattern of activity as kinase inhibitors, it would be reasonable to expect that the related triazolopyridines series (Figure 5.9(b)) might also be active against the same kinase. Similar SAR patterns would likely emerge in the triazolopyridines class, but there may be differences as a result of the presence of an additional nitrogen atom in the five-membered ring. The corresponding tetrazolopyridines (Figure 5.9(c)) are also similar to the original lead series, as are the tetrazolopyrimidines (Figure 5.9(d)), but each step away from the original compound class changes the chemical nature of the compound class and can result in unexpected changes in SAR patterns. Making the assumption that the SAR patterns of similar compound classes will be identical is a risky proposition. It is well worth the effort to test the limits of the overlap of SAR patterns between related series.

QUANTITATIVE STRUCTURE–ACTIVITY RELATIONSHIPS

Identifying structure activity trends is an integral part of the drug discovery process, but ultimately this aspect of the process is inherently limited by the number of compounds that can be prepared in the lifetime of a discovery campaign. Although it may be possible to generate several hundred, or perhaps even a few thousand analogs in any given series, the theoretical number of compounds that could be prepared is virtually unlimited. Efforts to merge the concepts of structure–activity relationships with computational methods and growing computer power led to the development of quantitative structure–activity relationship (QSAR) models. In simple terms, QSAR models represent an attempt to mathematically quantify molecular properties and correlate changes in these properties to changes in biological activity.

This information is then used to predict the impact of potential alterations in the molecular structure on the biological activity of potential candidate compounds. In theory, if one could develop a QSAR model capable of accurately predicting biological activity in a given series of compounds, then the computational power of modern computer systems could be employed to determine which compounds to prepare, greatly simplifying the selection process.

In practice, QSAR models require that molecular properties be distilled to mathematical equations, and this is not easily accomplished. Early efforts to develop mathematical relationships between structure and activity include the Hammett equation. Originally proposed in 1937 by Louis Plack Hammett,

$$\log (K / K_0) = \rho \sigma \quad (5.1)$$

Equation (5.1): The Hammett equation: K and K_0 are equilibrium constants for the hydrolysis of benzoic acid derivatives, where K is the value for a substituted benzoic acid and K_0 is the value for the unsubstituted reference compound, ρ is a reaction constant, and σ is a substituent constant that is dependent only on the electronic effects of the substituent.

$$\log (K / K_0) = \rho^* \sigma^* + \delta E_s \quad (5.2)$$

Equation (5.2): The Taft Equation: K_0 and K are the equilibrium constants for a reference reaction and a reaction in which substituents are present, σ^* describes the electrical field and inductive effective of a substituent in the reaction, and ρ^* is a sensitivity constant that defines the impact of polar effects on the rate of the reaction. Steric effects in the reaction are defined by the term E_s , and δ is a sensitivity factor that considers the sensitivity of the reaction to steric effects.

$$\log (1 / C) = a \log P + b (\log P)^2 + \rho \sigma + \delta E_s + d \quad (5.3)$$

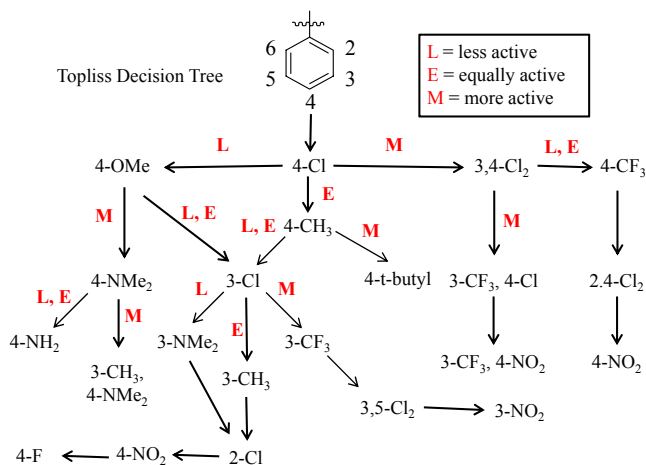
Equation (5.3): The Hansch equation: The terms ρ , σ , δ , and E_s are defined in the same manner as in [Eqns \(5.1\) and \(5.2\)](#); $\log P$ is the octanol/water partitioning coefficient; and a , b , and d are mathematical constants derived through linear regression analysis of experimental data. The concentration at which a biological response is being measured for data analysis is defined by the term “ C .”

this formula proposes that the rate of hydrolysis of a benzoic ester is a function of the electronic effects of the substituents of a benzene ring according to [Eqn \(5.1\)](#).⁸ Although limited in its scope and not directly applicable to drug discovery, the concepts proposed by Hammett were considered revolutionary at the time. Robert W. Taft modified the Hammett equation in 1952 by incorporating the concept of steric effects in an attempt to broaden the applicability of the equation ([Eqn \(5.2\)](#)),⁹ but many additional factors remained outside the

realm of this model. Corwin Hansch, often referred to as the father of QSAR, was one of the early pioneers in developing mathematical representations of the interactions of drugs and their biological targets. He was the first scientist to demonstrate that biological activity could be quantitatively related to physicochemical characteristics of compounds (Eqn (5.3)). His equations built upon the work of Hammett and Taft, but he added concepts such as lipophilicity ($\log P$) and “active concentration” (C) in order to improve the utility of the formula for predictive purposes.¹⁰

Of course, the complexity of the equations required to predict the molecular properties of candidate compounds increases with each parameter added, and there are many to consider. Additional molecular factors that could be considered as part of an effort to predict biological and physicochemical attributes of candidate compounds include molecular weight, total polar surface area, interatomic distance and angles, potential for hydrogen bonding (as a donor or acceptor), and three-dimensional shape just to name a few. In the absence of the appropriate computer technology and software, it would be nearly impossible to perform the necessary calculation on a sufficient number of compounds within the time frame of the typical drug discovery project. Up until the end of the twentieth century and the beginning of the twenty-first century, computer capacity and software technology limited the translation of computational data into drug discovery efforts and compound design.

Efforts to apply the principles of QSAR to drug design on a practical level began with the development of “decision trees” designed to guide compound design. Rather than preparing a large number of analogs, one could explore the SAR parameters of a particular compound class by following the guide points set in the decision tree. The Topliss Scheme, for example, was introduced in 1970¹¹ as a means of systematically examining the impact of changing aromatic substitution on the biological activity of candidate compounds (Figure 5.10). This system takes into consideration factors such as inductive electronic effects and lipophilicity, and begins with the unsubstituted aromatic ring. Modification of the structure by incorporation of a chlorine atom in the 4-position followed by biological screening provides data for the first decision point. At this juncture, increased activity would suggest further modification of the aromatic ring by incorporation of a second chlorine atom in the 3-position, while decreased activity would suggest that incorporation of a 4-methoxy substituent might lead to an improved outcome. If biological activity remains the same (within a reasonable margin), then the Topliss tree path indicates that the next analog of interest is the 4-methyl analog. As each analog in the tree is prepared and screened, a new decision point is reached and new compounds are suggested. Decision trees of this type are based on empirical observations across multiple data sets, but are an imperfect guide for molecular design strategies. Subtle differences between compound classes can lead to deviations from the expected results. The advantages of not requiring computational analysis in order to decrease the potential



number of analogs needed in order to establish an SAR pattern, however, cannot be understated.

Over the last several decades, molecular modeling software and computing power has increased substantially. Modern QSAR methods incorporate numerous mathematical algorithms that are designed to incorporate a wide range of physicochemical properties. Software packages such as Tripos Sybyl-X have become substantially more user friendly, and for many tasks, it is no longer necessary to have an advanced understanding of computational chemistry in order to apply the tools necessary to predict changes in biological activity, provided enough structural data are available. For example, potential analogs of a known lead compound can be created *in silico*, energy minimized through the click of a mouse (and an automated algorithm designed for this purpose), and then compared to the lead compound. Physical overlaps can be observed, electrostatic surfaces can be compared, and positions of hydrogen bond donor and acceptors can be examined to determine the differences and similarities between compounds. Automated systems capable of comparing thousands, if not millions, of compounds with a single lead compound and ranking them for similarity to the lead compound are available. *In silico*

screening of compounds can be accomplished in a relatively short time frame (hours).

The development of sophisticated computer modeling designed to provide a more comprehensive understanding of the relationship between molecular properties and biological activity also provided scientists with the ability to visualize compounds in the way that macromolecular targets “see” them. On paper, scientists use lines and juncture points to describe molecules, both large and small, but in reality, potential therapeutic entities and macromolecular targets “see” each other as something far more complicated than simply a collection of atoms joined together. They “see” overall three-dimensional shapes with contoured regions of high electron density and low electron density, areas of high and low lipophilicity, and hydrogen bond forming potential (hydrogen bond donors or acceptors). The visualization tools developed as a consequence of advances in QSAR modeling provide scientists with the ability to generate three-dimensional images of compounds and view them in the way that they are “seen” by their intended biological targets. Substantially different compounds that modulate the same target could be more readily compared to gain a better understanding of similarities between the apparently different classes of compounds that generate the same biological response. The selective serotonin reuptake inhibitors sertraline,¹² citalopram,¹³ fluoxetine,¹⁴ for example, have very different chemical structures, yet they interact with the same biological target (Figure 5.11). When viewed through the

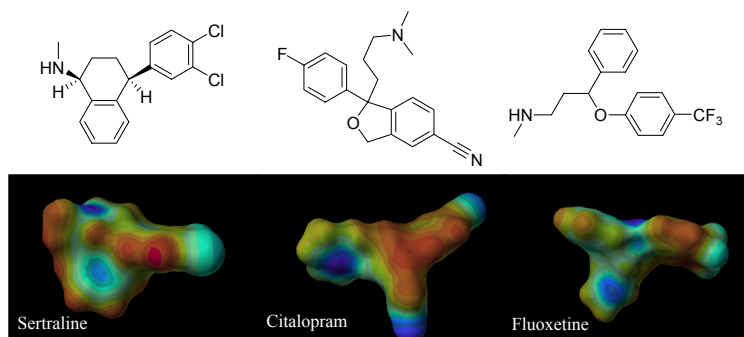


FIGURE 5.11 In order to facilitate communicating structural information, scientists typically draw compounds using line structures such as those portrayed for the three selective serotonin reuptake inhibitors (SSRI) sertraline, citalopram, and fluoxetine. It is important to remember, however, that this representation is only a shadow of the full structure of the molecule. A more accurate depiction can be created using advanced computer software and graphical display systems that can provide renderings of molecules that display features beyond the simple connectivity of the atoms. In this case, electrostatic potential surface maps are superimposed over the three-dimensional structures of the SSRIs in order to visualize the similarities and differences. Images of this type can help explain why compounds with apparently different structures can behave in a similar manner within the context of a biological system.

lens of QSAR modeling, however, similarities between the three compounds become apparent. Overlaps between regions of high and low electron density become apparent, lipophilic regions can be compared, and the orientation of hydrogen bond donors and acceptors provides a greater understanding of the relationship between these three compounds. Thus, QSAR modeling provides the scientific basis for understanding why compounds with dissimilar chemical structures interact with the same macromolecular targets by providing a more comprehensive view of candidate compounds. In simpler terms, compounds with very different chemical structures can modulate the same biological target if they “look” similar to each other through the QSAR lens.

THE PHARMACOPHORE

Another aspect of structure–activity relationships considers the relative importance of various sections of an overall molecular framework, rather than individual changes to a series of lead compounds. In many cases, a binding interaction between a lead compound and its macromolecular target requires only a subset of the atoms and functionalities available within the molecular framework of the lead compound. These sections of the molecule are referred to as the pharmacophore of the structure, and changes in these parts of the molecule can have a dramatic impact on the strength of binding interactions and biological activity. The official definition provided by the International Union of Pure and Applied Chemistry (IUPAC) states that a pharmacophore is “an ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target and to trigger (or block) its biological response.”¹⁵

The remainder of the compound, portions of which may provide structural support for the position of the functionality within the pharmacophore but play no direct role in binding events, is often referred to as the auxophore. In principle, since the auxophore is not directly involved in binding events, changes to this section of a potential lead compound should have minimal impact on binding interactions, provided that the changes do not impact the overall shape of the pharmacophore. These sections of the compound may be oriented away from the binding surface of the macromolecule, so changes in these regions, including excision from the lead compound, can often be accomplished while maintaining biological activity. It is also possible to take advantage of auxophore space for the purposes of altering physical properties of a compound. Consider, for example, a compound with poor solubility that contains a side chain that points away from the binding surface of the target of interest and into solvent space (i.e., water). Since this side chain is outside of the pharmacophore, changes to this side chain such as the addition of a basic amine should have little impact on the desired biological activity. It is, however,

likely that the addition of a basic amine will have a positive impact on solubility, thereby improving a physical property of the compound without negatively impacting biological function.

Understanding the distinction between the pharmacophore and the auxophore can provide significant design opportunities. Consider, for example, the four compounds: morphine,¹⁶ levorphanol,¹⁷ metazocine,¹⁸ and meperidine¹⁹ (Figure 5.12). All of these compounds are analgesic

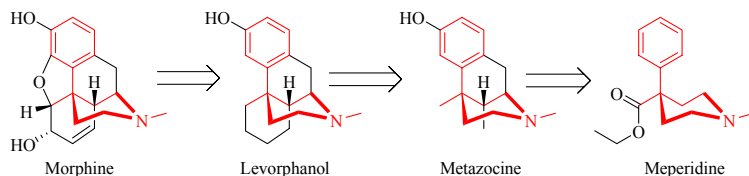


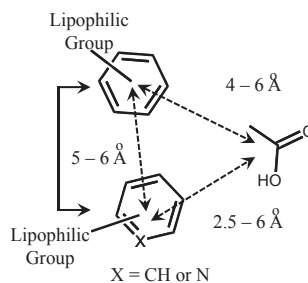
FIGURE 5.12 A pharmacophore can be described as the minimum structural requirements that must be present in order for a compound to have biological activity at a given target. The μ -opioid receptor family of analgesics, typified by morphine, levorphanol, metazocine, and meperidine, can be described by the minimum pharmacophore outlined in red in each structure. Features outlined in black can have a significant impact on biological activity, but can also be modified (levorphanol), or even removed entirely (metazocine and meperidine), while still maintaining the ability to bind to the μ -opioid receptor. Loss of the features in red, however, leads to a loss of μ -opioid receptor binding.

agents that bind to the μ -opioid receptor, but with varying degrees of potency. The common pharmacophore, outlined in red, represents the minimal structural feature required for biological activity. Morphine is, of course, a powerful analgesic, but levorphanol, a less-known structural analog in which one of the two alcohols and the oxygen-containing ring have been removed, is three to four times *more* potent than morphine. This indicates that these regions of the molecule are not part of the critical binding elements for the μ -opioid receptor pharmacophore and were actually impeding binding interactions.

Metazocine and meperidine are also part of the μ -opioid receptor family of analgesics, but they are both less potent than morphine and levorphanol. In both cases, the basic pharmacophore is intact, but molecular elements responsible for rigidity and shape have been removed. In the case of metazocine, the lower cyclohexane ring has been truncated, leading to a loss in activity relative to morphine, but the excision of both cyclohexane rings in meperidine leads to an even more dramatic loss of potency. Meperidine possesses only 10–12% of the analgesic potency of morphine, even though the overall pharmacophore remains intact. The ability to excise portions of the auxophore may be limited by their importance in maintaining the overall three-dimensional shape of candidate compounds.²⁰

Understanding the nature of a pharmacophore is greatly facilitated when a set of related compounds such as the μ -opioid receptor family of analgesics is available for analysis, but it is often the case that apparently unrelated compounds elicit a biological response by binding to the same macromolecular target. Identifying a unifying pharmacophore within a structurally diverse set of compounds may require a more abstract view-point. Rather than defining a distinct molecular framework as in the previous example, it may be necessary to define required chemical functionalities and their relative positions in space. In the case of the voltage-gated chloride channels ClC-Ka, for example, Liantonio et al.²¹ described a pharmacophore containing two aromatic rings and one carboxylic acid positioned within a specific distance from each other (Figure 5.13). The

FIGURE 5.13 A pharmacophore is not always described as a substructure of a larger class of compounds as seen in the μ -opioid receptor family of analgesics. In many cases, a pharmacophore for a macromolecular target can be defined by key features and their relative positions in space (distances, torsional angles, etc.). The pharmacophore model of the voltage-gated chloride channels ClC-Ka provides guidance for compound design and optimization even though it is not a complete structure.



nature of the linker units between each region of the pharmacophore is defined only by the distance between the regions and their ability to hold the aromatic rings and acid moiety in the proper positions relative to each other. The ability to generate constructs that are capable of defining the physical and chemical requirements for biological activity at a given target has been greatly facilitated by the development of software capable of analyzing molecular interactions on a broad scale.

DEVELOPING AN SAR DATA SET

The drug discovery process depends on the ability of scientists to make informed decisions based on the structure–activity relationships observed through the course of their program. In the beginning of a program, however, it is often the case that very little biological data is available, as only a few compounds, usually benchmark compounds and literature leads, have been examined in biological assays. This, of course, begs the question of how does one begin the process of developing an SAR database that can be used as a basis for design strategy and project decisions?

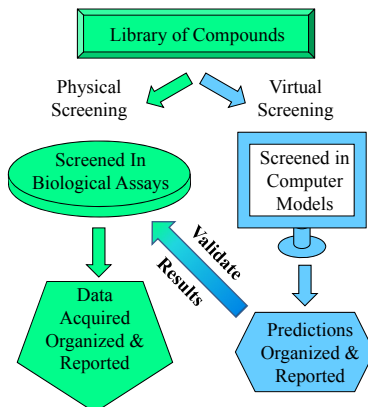
It is often the case that these efforts begin with the high-throughput screening of large compound libraries in an effort to identify groups of related compounds that possess the desired biological activity. Ideally, the compound libraries should be composed of a diverse set of pharmacophores that cover a large portion of “druglike” chemical space and contain enough members of representative chemical classes such that an initial set of SAR data will be provided in the first round of screening. The number of compounds necessary for this kind of exercise to succeed is a subject that has been debated in the literature and no clear answer has emerged. Some institutions maintain compound collections of over 1 million discrete compounds.

In practice, libraries available for screening are usually biased based on their origin. Private compound collections available to pharmaceutical companies and research institutions tend to be populated with historical compounds from internal efforts. This increases the likelihood of identifying novel leads and groups of related compounds, but creates an institutional bias in the collection’s coverage of chemical space. Large commercial compound collections are also readily available from a variety of vendors (Maybridge, Enamine, Life Chemicals, etc.) and can be effective tools as initial screening sets. However, they are composed of known compounds, making intellectual property protection potentially more difficult to obtain (see Chapter 12) and are often expensive, especially from the perspective of small biotech companies and academic institutions. Searchable databases of commercially available compounds are available, the largest of which is the Zinc database with over 21 million entries.²² Custom compound collections can be designed and acquired in order to decrease the cost burden, but intellectual property issues will remain.

Natural product and biological extract collections have also been used as starting points for screening programs designed to generate initial SAR data, but it is important to realize that these libraries may contain complex mixtures. Identifying compounds with positive results may be challenging, and it is also possible that the mixtures contain previously identified active components with previously established intellectual property. Compound supply issues should also be considered, as natural products can be exceptionally complex to prepare in scales required for follow-up activities. That being said, a large portion of marketed pharmaceutical agents are natural products, or derivatives thereof, so this avenue of approach is worth considering as compound collections are built.

Irrespective of the source of the library, once the compound collection is chosen, the next step in building an SAR data set is screening the compounds for biological activity in an assay. This can be accomplished either via physical screening of the sample collection or by virtual screening

FIGURE 5.14 An initial data set for a drug discovery program can be derived from a physical screening of a compound library or a virtual screening of the same library if an appropriate computer model is available. Each method has advantages and disadvantages that should be considered, and neither method is exclusionary of the other. It is, however, very important to validate the *in silico* screening results with physical screening methods to ensure that the virtual screening is truly predictive of real-world results.



(Figure 5.14). As discussed in Chapter 4, there are a multitude of assay platforms and technologies that can be crafted to screen thousands of compounds for biological activity in 96-, 384-, or even higher throughput modes on robotic systems, providing access to potentially millions of data points. If, for example, a library of 100,000 compounds were examined using a 12-point dilution curve to determine IC_{50} values at a biological target, this would produce 2.4 million data points when run in duplicate, not including positive and negative controls used to ensure assay integrity. Although this may seem like a large amount of data, many large pharmaceutical companies maintain compound collections containing 500,000 to 1.5 million samples, which would produce a proportionately larger number of data points to analyze if screened in its entirety. Clearly, this is a staggering level of data output that cannot be processed in a reasonable length of time in the absence of the proper computer software. Fortunately, there are a number of commercial software systems capable of analyzing screening data, generating IC_{50} data, and associating that data (or any other data set) with a structure-searchable chemical database. Once the screening data have been matched with the chemical structures, the data can be sorted to identify active compounds and collections of related compounds in order to identify any nascent SAR trends that may be discernible in the first round of screening.

While high throughput screening is an effective tool for identifying initial leads and early SAR trends, the previous discussion neglects to consider a major factor associated with high throughput screening. The cost of this process can be quite high. Even if a compound collection is already in place, the cost of consumable materials can be substantial. Consider again a compound collection with 100,000 members. A high throughput screening assay run in 96-well plates using three concentrations run in duplicate would require 7500 microtiter plates, thousands of disposable pipette tips for reagents and solvent transfer, as well as reagents and

solvents necessary for the assay itself. A single high throughput screening cycle for a compound collection of this size can cost thousands of dollars, creating a significant cost barrier that must be considered.

As an alternative to physical screening of compounds, many organizations choose to employ a virtual screening approach. Rather than acquiring large compound collections, electronic databases such as the Zinc database mentioned earlier are examined *in silico* using computer models designed to identify potentially active compounds. In some cases, the model is based on an X-ray crystallographic structure of the target macromolecule bound to a known ligand. Binding interactions of the known ligand are defined in the model system, the ligand is then removed from the model-binding site, and database compounds are then “docked” in the binding site (Figure 5.15). The relative binding strength of each member

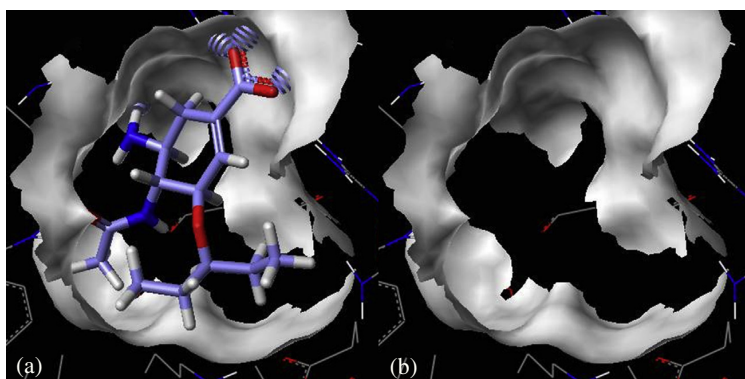


FIGURE 5.15 (a) An X-ray crystallographic structure of influenza virus neuraminidase with an inhibitor bound to the active site provides information that can be used in the design of additional inhibitors. The boundaries of the binding site (gray) and important interactions are displayed, including hydrogen bonds (arrows). (b) *In silico* excision of the inhibitor from the binding site using protein modeling software provides a template for virtual examination candidate compounds. RCSB 2QWK.

of the collection is calculated and compared with the calculated binding strength of the original ligand. The information derived from this process is then used to identify compounds with binding “scores” equal to or better than the original ligand. In theory, the compounds identified in this process would have binding affinities for the target macromolecule that are similar to or better than the original ligand.

If an X-ray crystallographic structure for the biological target is not available, it may be possible to create a homology model *in silico* using structural data from a macromolecule with a high enough degree of similarity to the target of interest. As noted earlier, it has been demonstrated that proteins with similar primary structures also possess similar secondary

and tertiary structures.²³ Given the relationship between primary structure and the overall three-dimensional structure of a protein, this should not be surprising. The process of building the model itself begins with an alignment of the linear peptide sequences (the primary structure) of the two macromolecules. Once the sequences are aligned, the known structural data (an X-ray crystal structure²⁴ or a structure built with protein NMR data²⁵) are used as a template to establish a three-dimensional model of the macromolecule of interest. Of course, there will be some areas within the model structure that will require adjustment in order to account for differences in the primary sequences of the template and the model. Structural changes to the model are generally accomplished through energy minimization, the addition of water molecules that may have been present in the original X-ray structure, and the elimination of overlaps within the structure. The model is then tested by docking ligands in the theoretical binding site, scoring their relative binding affinities, and comparing the *in silico* results with real-world data. This process has been successfully employed to create model systems for a number of GPCRs based on X-ray crystal structures of rhodopsins, a family of light-sensing transmembrane proteins.²⁶ Structural data derived from the X-ray crystal structures of bacterial rhodopsin, for example, were successfully adapted to establish a homology model for the dopamine D2 receptor (Figure 5.16). Although

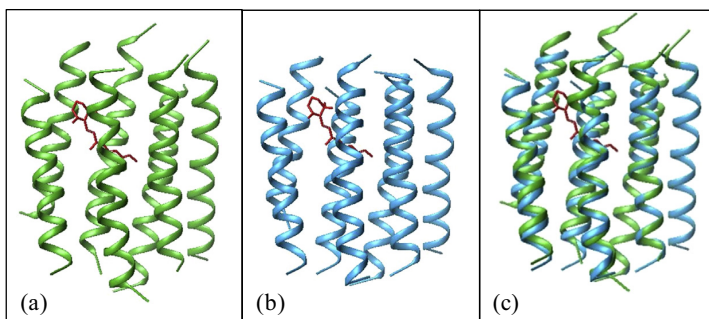


FIGURE 5.16 In the absence of an X-ray crystallographic structure, it is possible to build a model *in silico* to support virtual screening efforts. Structural data for bacterial rhodopsin ((a), RCSB 1BRD) were successfully mapped onto the dopamine D2 receptor to design a homology model ((b), RCSB 1I15) suitable for docking and evaluating candidate compounds. An overlay of the original structure with the model structure (c) demonstrates significant similarities between the two structures, but also a number of differences.

the overlap between the two structures is not perfect, the homology structure provided significant insight into the structural details of the dopamine D2 receptor that would not have been available otherwise.²⁷

In the absence of structural data for the biological target of interest, it is still possible to perform a virtual screen of compound databases in

order to focus screening efforts. Rather than docking compounds into a binding site, known ligands can be used as templates to identify potential lead compounds from within a structural database. This process is often referred to as ligand-based design. Modern molecular modeling software is applied to a starting compound, often a natural ligand or a known drug molecule, to determine its molecular properties, mapping out lipophilic potential surfaces, electrostatic potential surfaces, areas suitable for hydrogen bonding, lipophilic regions, and a wide array of molecular properties that define the compound in question (Figure 5.17). If multiple ligands

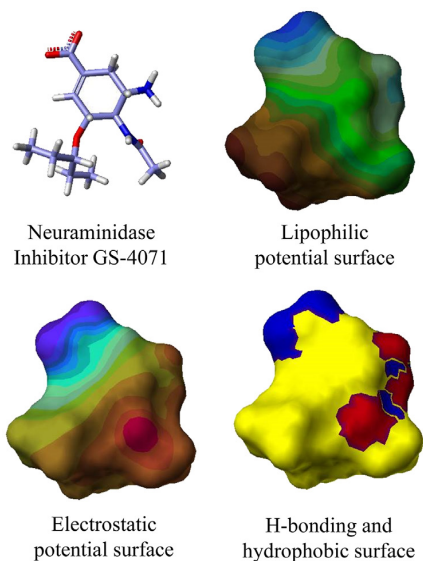


FIGURE 5.17 The neuraminidase inhibitor GS-4071 is structurally described by the classical stick figure (top left), but computer modeling can provide improved visualization of molecular properties such as lipophilic potential surfaces (top right), electrostatic potential surface (bottom left), and hydrogen bonding/hydrophobic surfaces (bottom right). These “maps” can be combined to provide a model system useful for the identification of compounds that are similar to GS-4071.

are available, it may be possible to identify common features from each ligand/drug to define a pharmacophore for the binding site of interest. The aforementioned property sets are determined for all of the members of the compound collection, the individual members of the compound collections are compared with the lead compound, and each is assigned a score based on its overall similarity to the lead compound. Compounds with higher scores (those that more closely resemble the lead compound) can be identified in this manner and similar compounds can be grouped together to select compounds for physical screening, once again significantly decreasing the overall cost of physical screening.

Irrespective of whether virtual screening results are based on an X-ray crystal structure, a homology model, or comparison of compound libraries with a known ligand, it is important to realize that virtual screening is a *model* of reality. As such, it must be validated with physical screening

to ensure that there is a correlation between the results provided by the virtual screening and the physical screening. If, for example, compounds with high binding scores in a model of an enzyme were demonstrated to be potent inhibitors of enzymatic activity, and that as scoring values decreased, inhibitory action of the compounds decreased, then the *in silico* screening is validated for identifying compounds that are inhibitors of the enzyme in question (Figure 5.18(a)). On the other hand, if there

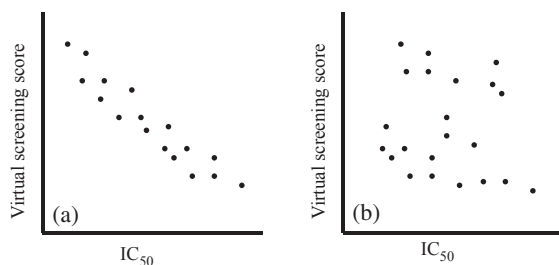


FIGURE 5.18 In order for a virtual screen to be useful, it must correlate with real-world results (a). In the absence of this correlation (b), data provided by the virtual screen cannot be used to predict biological activity of candidate compounds.

is no relationship between the scores generated through virtual screening and the inhibitory action of compounds physically screened against the enzyme, then the *in silico* model cannot be used as a predictive tool for virtual screening (Figure 5.18(b)). Of course, no model is a perfect representation of reality, so a perfect correlation should not be expected. Additional data from physical screening or new structural data for the target can be used to refine models to improve their ability to predict biological activity.

The strategic application of virtual screening can significantly reduce the number of compounds examined in a physical screen, decreasing the required monetary commitment. There are, however, costs associated with computer capacity and state-of-the-art molecular modeling software (including annual license fees) that should be considered. These costs can rival that of the purchase of a sizable physical library. Also, when virtual screens do not correlate to the physical screening system that they are designed to represent, there is a loss in productivity that is more difficult to quantify but important to consider. It is worth pointing out at this juncture that physical and virtual screening methods are not mutually exclusive strategies for drug discovery. They are often used in tandem, as they can be complimentary pathways to success.

Although screening of large compound collections can be an effective method of identifying biologically relevant compounds, the process is not without its issues. As previously mentioned, the costs associated with

screening even a medium size collection of 100,000 compounds can be significant. In addition, the process itself is essentially a random selection process. Potentially millions of compounds are screened, whether physically or *in silico*, relying on the compound designs of previously defined compound collections. Compounds that are part of a historical collection may be unique to a given institution, but they were most likely prepared for a specific purpose which may not be useful for a current program (It could, in fact, represent a problem if off-target activity becomes an issue). It is also quite possible that these compounds have already been disclosed to the public as part of the patent application process for their project of origin, potentially creating intellectual property issues. A similar issue exists for commercially acquired samples, as they are part of the public domain, again creating a potential issue for patent protection, a necessary facet of drug discovery.

The sheer volume of data resulting from a high throughput screen can also be quite daunting. High-throughput screening as a means of identifying biologically active compounds has been compared to looking for a needle in a haystack, and with good reason. A high throughput screen of a collection of 500,000 compounds producing a positive result of only 0.2% will provide a set of 1000 compounds for examination and potential follow-up. While there are software packages capable of simplifying data handling, at the end of the day, scientists must review the data and determine how to move forward. There is also the issue of false positive and false negative results. The aforementioned high throughput screening run of 500,000 compounds would produce 100 erroneous results if the false positive/negative rate 0.02%. If they are equally split between false positives and false negatives, then there will be 50 compounds incorrectly identified as biologically active. The remaining 50 compounds that are false negatives are more difficult to track down, as they will blend in with the rest of the compounds collection that did not meet the desired potency range.

While the randomized approach discussed thus far has been successfully applied to numerous drug discovery programs across a wide array of disease states and biological targets, it is by no means the only method of identifying a starting point for structure-activity relationship studies. An alternative approach employed by many drug discovery scientists takes advantage of the significant advances in our understanding of the relationship between structure and function in a more direct fashion. In this approach, often referred to as rational drug design or structure-based drug design, lead series and eventually lead compounds are developed through a detailed analysis of structural information on the targeted macromolecule, known ligands (both natural and synthetic), or both.

When an X-ray crystal structure of a macromolecule bound to a small molecule is available, for example, modern molecular modeling software can be employed to determine points of positive interaction (hydrogen bonding, salt bridges, lipophilic interactions, etc.), clash points between the

ligand and the biomolecule (negative interactions), and areas of empty space within the binding pocket that might be accessible if changes were made to the ligand. Compounds that improve upon positive interactions, perhaps by optimizing the distance between a hydrogen bond donor/acceptor pair, removing a side chain that hinders binding, or adding a new side chain to create an additional binding interaction, are then designed based on the available structural data. The new compounds can be “prepared” *in silico*, docked with the X-ray structure, and scored for binding energy relative to the original ligand to assess whether or not preparation is warranted. Biological screening of the actual compounds should be used to validate the *in silico* methods to determine the reliability of the model (i.e., docking of the compounds in the X-ray crystal structure), but if it is validated by real-world results, it can be used to prioritize medicinal chemistry efforts.

These techniques have been employed to identify inhibitors against a wide range of enzymes such as human protein tyrosine phosphatase- β (HPTP- β), a receptor-type phosphatase that is expressed in endothelial cells that form the protective lining of blood vessels.²⁸ X-ray crystallographic studies of the catalytic domain of HPTP- β bound to a vanadium tetraoxide ion indicated that this compound formed multiple hydrogen bonding interactions in the binding site (Figure 5.19(a)). It was then

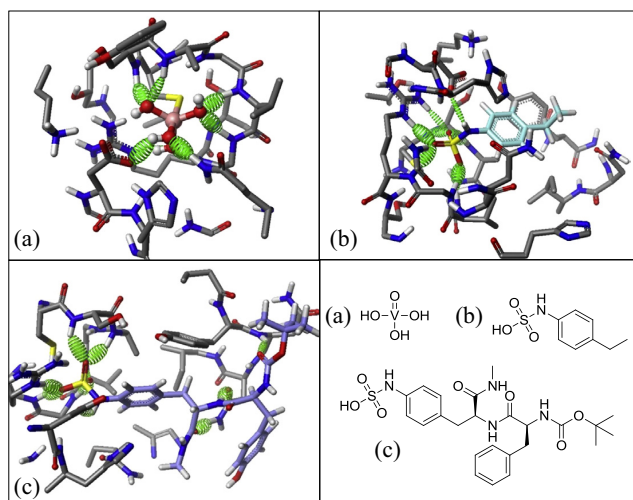


FIGURE 5.19 (a) Vanadium tetraoxide ion forms a network of hydrogen bonds (green barrels) and binds tightly to the active site of human protein tyrosine phosphatase- β (HPTP- β) (RCSB 2I4E) (b) 4-Ethylbenzene sulfamic acid forms a similar series of hydrogen bonds (RCSB 2I5X) and was used as a starting point for a drug discovery program that eventually produced a highly selective HPTP- β inhibitor ((c), RCSB 2I4H).

surmised that small molecules capable of replicating this intricate network of hydrogen bonds should be able to occupy the same space in the HPTP- β catalytic site, inhibiting enzyme action. Synthesis, biological evaluation, and additional X-ray crystallography studies verified this hypothesis, as 4-ethylbenzene sulfamic acid formed a stable complex with catalytic domain of HPTP- β , inhibiting enzyme action (Figure 5.19(b)). Molecular modeling studies using this information lead to the identification of more advanced inhibitors capable of forming additional hydrogen bonding interactions with nearby amino acid residues (Figure 5.19(c)), which lead to a significant increase in potency against the target.²⁹

Fragment-based screening³⁰ has also been employed as a means of developing structure–activity relationship, but as the name implies, the starting points for these programs are typically smaller than the average drug molecule. In this approach, low-molecular weight compounds, typically 100–250 AMU, are screened to determine their binding affinity with a macromolecule. Compounds of this size typically have low binding affinities ($>100\ \mu\text{M}$), so sensitive techniques such as X-ray crystallography,³¹ protein NMR,³² or surface plasmon resonance (SPR)³³ are required, and throughput is low compared to high-throughput screening methods. Compounds that are identified in these processes, however, generally create efficient binding interactions with the macromolecules when their low molecular weight is taken into consideration.³⁴ As such, they can be employed as starting points for further optimization. This is typically accomplished by either “growing” the fragment compound or “linking” multiple fragments molecules together. The “linking” strategy depends upon synergies that may be created by forming linkage points between fragments that bind to the target in neighboring regions. Just as the individually weak interaction in protein structure can collectively form very stable three-dimensional structures, joining together multiple fragment structures can have a synergistic effect, leading to tighter binding than is observed for the individual fragments, possibly greater than the sum of the individual components (Figure 5.20). The linker itself may also play a role in binding if additional

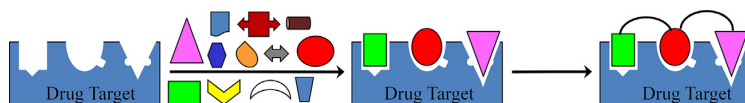


FIGURE 5.20 In one approach to fragment-based screening, low-molecular weight compounds are screened to determine binding affinities using sensitive screening techniques such as NMR, SPR, and X-ray crystallography. Low-molecular weight compounds that demonstrate binding are then stitched together with linker units designed to allow the individual components to access their individual binding sites. The tethered individual units act collectively to create a single compound with improved biological activity based on synergistic binding.

binding interactions are available. If, for example, the linker contained an amide bond, then it could form a hydrogen bond with an appropriate functional group in the targeted macromolecule, further enhancing binding strength.

As an alternative, one can “grow” a fragment molecule into unoccupied space in the binding site of target molecule. This process is greatly facilitated by the availability of either an X-ray structure of the fragment bound to the target or a homology model that can be used to visualize potential areas of interest in the binding site. Using this technique, indazole was identified as a fragment “hit” with very low binding potency (185 μM) for cyclin-dependent kinase 2 (CDK-2) and eventually transformed into AT7519, a significantly more potent CDK-2 inhibitor (47 nM, Figure 5.21).³⁵

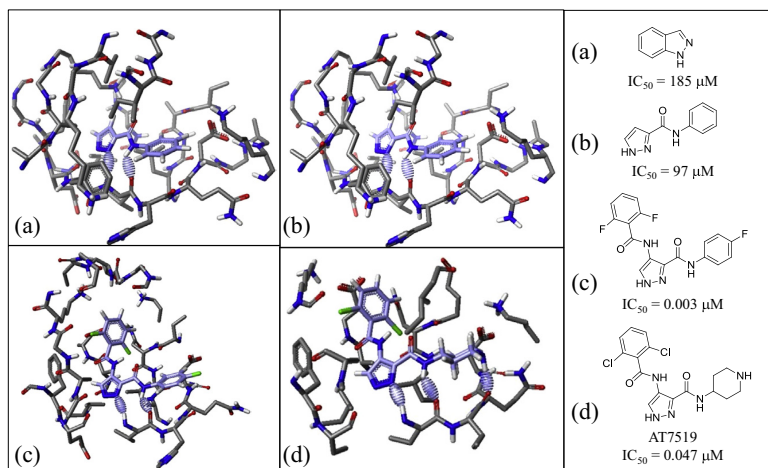


FIGURE 5.21 (a) Screening of a fragment library via X-ray crystallography identified indazole as a weak CD-2 binder (RCSB 2VTA) anchored into the binding site with a hydrogen bond (cyan barrel). (b) Removal of the benzene ring and growth of the fragment via addition of an N-phenyl-acetamide provided access to a neighboring binding site and an additional hydrogen bond interaction (RCSB 2VTL). (c) Growing the structure in the opposite direction by adding a benzamide created additional opportunities for hydrophobic interactions in a proximal section of the binding site, leading to a dramatic improvement in potency (RCSB 2VTP). (d) Incorporation of a piperidine to increase solubility and cellular activity led to the identification of the clinical candidate AT7519 (RCSB 2VU3).

Although each of the techniques described thus far has been discussed in isolation, it is important to realize that they are not mutually exclusive. In fact, it is often the case that multiple approaches are employed within the context of a given drug discovery program. Each approach has

strengths and weaknesses, but none of them have proved to be superior to the others in identifying a clinical candidate, let alone a drug. Each of these methods is most appropriately viewed as a tool to be employed when appropriate.

THE STRUCTURE–ACTIVITY RELATIONSHIP CYCLE

Ideally, once the first round of screening is complete and the data have been analyzed, it will be possible to identify a set of compounds that have varying degrees of biological activity at the target of interest. Compounds with similar core structures can then be grouped together to determine how differences in their structures impact observed biological activity, providing an early indication of structure–activity relationships within the series. An iterative cycle of design, synthesis, screening, and analysis begins at this point. New molecules with structural changes are designed based on the available compound data set, synthesized, and screened for biological activity. The resulting data are analyzed to determine the impact of structural changes on biological activity. If the structural changes incorporated into the new set of compounds have a positive impact on screening results, then the changes are retained. On the other hand, if biological activity decreases, then the changes are discarded. In either case, the new information provided by the preparation and screening of the second set of compounds is incorporated into the SAR information available, and a new round of design, synthesis, screening, and analysis is initiated. This iterative process (Figure 5.22) is repeated over and over until a molecule with the desired biological properties is identified.

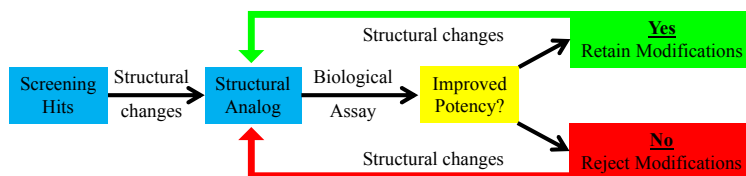


FIGURE 5.22 Identifying an initial set of hits in a drug discovery program, irrespective of their source and methods employed, is generally the beginning of an iterative process. Analogs containing specific structural changes are prepared and assessed for biological activity. Favorable changes are maintained, unfavorable changes are discarded, and a new round of compounds is prepared. Each successive round of synthesis and biological assessment builds on the previous round with the ultimate goal of optimizing biological activity.

BIOISOSTERISM

As discussed above, simple structural changes, such as increasing the length of a side chain (methyl, ethyl, propyl, etc. [Figure 5.3](#)), adding functionality to a benzene ring as described in the Topliss Tree ([Figure 5.10](#)), or the additions of hydrogen bond acceptors/donors ([Figure 5.4](#)) can lead to improvements in biological properties. There are, however, additional avenues that can be explored. The concept of isosterism, originally developed by Langmuir³⁶ and then broadened by Erlenmeyer,³⁷ can also be applied to drug discovery. Langmuir (1919) and Erlenmeyer (1932) hypothesized that compounds, functional groups, or atoms with the same arrangement and number of electrons in their outermost electron shells would have similar physical and chemical properties based on their isoelectronic features. The scope of Erlenmeyer’s original premise was eventually expanded by Harris L. Friedman in 1951 when he defined compounds as bioisosteric “if they fit the broadest definition for isosteres and have the same type of biological activity.”³⁸ In a similar manner, functional groups and atoms within a more complex molecule that can be interchanged while maintaining the same or similar biological properties are referred to as bioisosteres of each other. Modification of candidate compounds by applying the premise of bioisosterism is commonplace in drug discovery programs, as it is a useful tool for expanding the structural diversity of a lead series to identify novel chemical space and to alter physicochemical properties. These subjects will be addressed in greater detail in Chapters 12 and 6, respectively.

Bioisosteres can be divided into two broad classes: classic and non-classic bioisosteres, as first proposed by Alfred Burger.³⁹ According to this system of classification, classic bioisosteres are broadly defined as atoms, molecular fragments, or functional groups that have the same valence and ring equivalents ([Table 5.1](#)). Typical examples include exchanging halogen atoms and replacing methylene units with an oxygen or sulfur atom. A practical application of this principle is seen in efforts to develop alternatives to the antihypertensive drug rilmenidine (Albarel). In this instance, replacing the oxygen atom of rilmenidine’s 4,5-dihydro-oxazole ring system with a methylene unit provided the 4,5-dihydro-pyrrole analog

TABLE 5.1 Classic Bioisosteres

Monovalent	Divalent	Trivalent	Tetravalent
-OH, -NH ₂ , -CH ₃ , -OR	-CH ₂ -	=CH-	=C=
-F, -Cl, -Br, -I, -SH, -PH ₂	-O-	=N-	=Si=
-SiR ₃ , -Sr	-S-	=P-	=N+=

Copyright © 2015, Elsevier Science & Technology. All rights reserved.

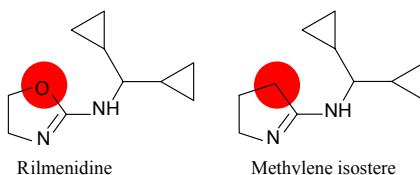


FIGURE 5.23 In this example of classic bioisosterism, exchanging a carbon atom for an oxygen atom (highlighted in red) eliminates an undesired off-target activity. While both compounds are potent modulators of the I_1 imidazoline receptor, rilmenidine's potent α_2 -adrenoceptor binding is eliminated by this simple change.

(Figure 5.23). Both compounds share similar potency at the I_1 imidazoline receptor, a key player in the regulation of blood pressure. On the other hand, the new series of compounds demonstrated improved target selectivity, as although rilmenidine potently binds to the α_2 -adrenoceptors, the methylene analogs showed no affinity for this receptor. This successful application of classic bioisosterism clearly demonstrates that simple changes in chemical structure can be employed to alter some properties (abolishing α_2 -adrenoceptor binding), while successfully preserving others (potency at the I_1 imidazoline receptor).⁴⁰

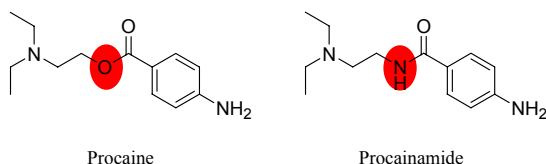


FIGURE 5.24 While both procaine and procainamide block the same sodium channel, only procainamide can be used as a treatment for arrhythmia. In this instance, bioisosteric replacement of an oxygen atom with a nitrogen atom provides a compound with similar biological activity and substantially improved metabolic stability.

Procaine and procainamide (Figure 5.24) are also excellent examples of the impact that can be achieved with a classic bioisosteric replacement. Both of these compounds exert their pharmacological impact via sodium channel blockade. Procaine, more commonly referred to as Novocain, is only useful as a topical anesthetic, while procainamide is an orally administered antiarrhythmic agent. In this instance, a classic bioisostere exchange is employed to improve metabolic properties. The ester linkage of procaine is hydrolyzed quickly *in vivo*, limiting its utility to topical application, while the amide linkage in procainamide is significantly more stable *in vivo*. The enhanced metabolic stability conferred

to procainamide versus procaine is critical to its utility as an oral antiarrhythmic agent.⁴¹

Non-classic bioisosteres represent a much broader range of structural diversity, as this category is defined as anything that does not fit the definition of a classic bioisostere. It should come as no surprise that the vast majority of examples of bioisosterism fall into this category and that a full characterization of this area of drug discovery is beyond the scope of this chapter. The concepts involved, however, can be described in a few examples. Replacing carboxylic acids in candidate compounds, for example, has been widely studied, as this functional group can create undesirable physicochemical properties. Compounds containing carboxylic acids tend to have low permeability through biological membranes, are poorly absorbed when dosed orally, and are rapidly cleared by the body. Their ability to cross the blood-brain barrier is also often limited at best. In other words, if a compound with a high degree of potency at a biological target of interest contains this functional group, moving the compound down the drug discovery and development pathway may prove challenging. Replacing this functionality with a suitable bioisostere may alleviate these issues, and some examples are shown

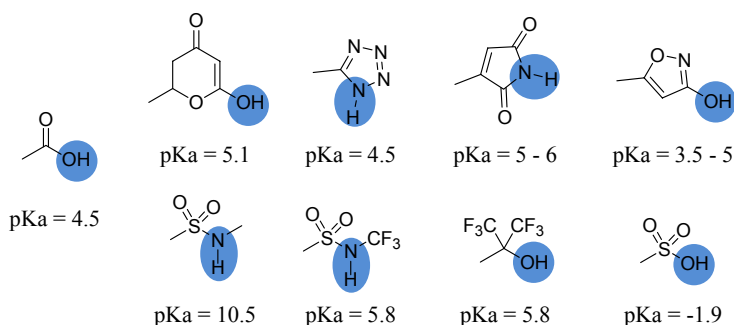


FIGURE 5.25 There are many potential non-classic bioisosteric replacements for a carboxylic acid, but they are by no means equivalent. Structural diversity is readily apparent, but an examination of the pKa values of each (highlighted in blue) provides an alternative view of the differences between the various possible carboxylic acid mimetics.

in Figure 5.25. In each instance, the dissociable proton of the carboxylic acid is maintained, but the remainder of the functionality is replaced with a collection of atoms capable of mimicking the original structure in a biological setting. The successful application of any one of these carboxylic acid bioisosteres is dependent upon the nature of the interaction between the candidate compound and the macromolecular target. In one successful example, one of the two carboxylic acids of Carbenicillin was

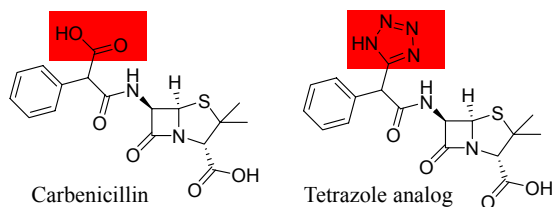


FIGURE 5.26 Although carbenicillin is a useful antibiotic, its use is hampered by spontaneous decarboxylation in acidic media. The tetrazole analog is similarly effective as an antibiotic, but does not suffer from instability in an acidic media as a result of the non-classic bioisosteric replacement (highlighted in red).

replaced with a tetrazole (Figure 5.26). In this instance, antibacterial activity was maintained, and an increase in chemical stability was imparted. Carbenicillin can spontaneously decarboxylate, while the tetrazole derivative is incapable of decomposing in this manner.⁴²

In some cases, non-classic bioisosterism can be employed to replace one scaffold with another. Biological activity is maintained, but physicochemical properties and target selectivity may be very different. Further, synthetic accessibility may be simplified, and greater intellectual property space may also be available around the new scaffold. It has been demonstrated, for example, that the steroid framework of 17 β -estradiol, a major female sex hormone, can be effectively replaced with a simple *trans*-stilbene (Figure 5.27).

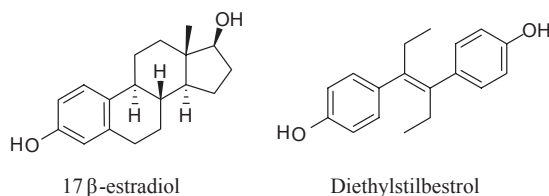


FIGURE 5.27 In this instance, three of the four rings in the steroidal framework are replaced by a stilbene framework. At the time of this discovery, this non-classical bioisosteric replacement provided novel intellectual property space and dramatically simplified the chemistry required to prepare novel estrogen receptor modulators.

Even though the two central rings of 17 β -estradiol have been removed and a third ring has been replaced with a benzene ring, *trans*-diethylstilbestrol remains a potent estrogen receptor agonist.⁴³

Bioisosteric replacement, whether classical or non-classical, is an exceptionally important concept in the field of drug discovery and development. As indicated by the examples provided, even small bioisosteric changes can have a dramatic impact on the physicochemical and biochemical properties of a compound. There are hundreds of examples of potentially useful

bioisosteres that one could consider in a drug discovery program, and a full discussion of this concept is well beyond the scope of this chapter. Fortunately, numerous review articles are available on the subject for those interested in an in-depth analysis of the subject.⁴⁴

STRUCTURE–ACTIVITY RELATIONSHIP, SELECTIVITY AND PHYSICOCHEMICAL PROPERTIES

Although the focus of the previous sections has been on the identification of increasingly more potent compounds through the analysis of structure–activity relationships, potency at a biological target is not the only consideration in a drug discovery program. As noted in Chapter 1, target selectivity is often critical to the identification of clinically relevant compounds, and physicochemical properties such as solubility and metabolic stability are major factors that must be considered. Fortunately, the same principles that have been discussed with respect to optimization of binding interactions at the intended molecular target can be employed to optimize any molecular property, provided there is a means of measuring the impact of structural changes on the property of interest. If, for example, one were targeting Kv1.5, a voltage-gated potassium channel that has clinical relevance in atrial arrhythmia,⁴⁵ optimizing this activity would be necessary. It would also be necessary, however, to establish that candidate compounds did not negatively impact the hERG channel, a voltage-gated potassium channel associated with Torsades de pointes and sudden cardiac death.⁴⁶ In this instance, a structure–activity relationship could be established for activity at the hERG channel, focusing on minimizing the potency at this channel. Thus, structural changes that *decreased* hERG activity would be maintained, while those that increased hERG activity would be dropped. Successful compounds would be identified by analyzing the SAR of *both* Kv1.5 and hERG in an effort to maximize one while minimizing the other.

In a similar manner, physicochemical properties can be optimized through structure–property relationships. Just as changes in molecular structure will have an impact of the ability of a compound to bind to a biomolecule, changes in molecular structure will also have an impact on physicochemical properties. Quantifying the impact of structural changes on a property of interest provides a means of optimizing that property through manipulation of the molecular structure. Some of these properties, for example, solubility and metabolic stability, require physical assessment. Fortunately, high-throughput screens have been developed to facilitate this process. Other properties, such as lipophilicity and polar surface area, can be calculated using the appropriate software tools. In either case, once knowledge of structural changes is paired with the appropriate data set, patterns can be established and properties can be

optimized appropriately. The importance of physicochemical assessment of compounds will be discussed in greater detail in Chapter 6, but the recognition of the high level of importance of physicochemical properties dramatically changed the way new drugs are pursued.

“DRUGLIKE” GUIDELINES

Although modern science has provided researchers with the ability to rapidly synthesize and screen compounds for biological activity across a wide range of targets, the small molecule universe is simply too large to explore exhaustively. It is estimated that there are over 10^{60} synthetically accessible small molecules, and only an infinitesimally small portion of this chemical space is likely to ever be explored.⁴⁷ Even if one were to consider only compounds within a specific class, the number of possible compounds is still extremely large. There are, for example, over 60,000 1,4-benzodiazepines (Figure 5.28)⁴⁸ in the literature as of 2013, and tens of

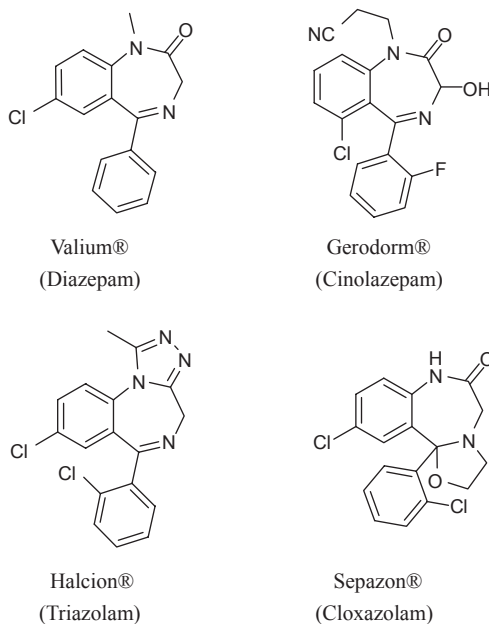


FIGURE 5.28 Valium® (Diazepam), Gerodorm® (Cinolazepam), Halcion® (Triazolam), and Sepazon® (Cloxazolam) are all members of the benzodiazepine family of compounds, a “privileged scaffold.” Chemical frameworks that have been the source of large numbers of biologically important compounds are referred to as “privileged scaffolds.”

thousands more could be prepared with commercially available building blocks. Given the enormity of chemical space within even a single series of compounds, the need to establish a general method of prioritization becomes evident. Structure–activity relationships are useful tools, but they are inherently limited by the need to synthesize and screen compounds in order to move a program forward.

In an effort to simplify what appears to be a near intractable problem, Christopher Lipinski and his colleagues examined the physicochemical properties of over 2000 drugs and drug candidates in clinical trials at the time (mid-1990s).⁴⁹ While biological activity is often the major focus of research programs, Lipinski et al. took heed of a 1991 Tufts University study on experimental drugs and why they fail. The Tufts study indicated that 30% of clinical failure could be attributed to a lack of efficacy, but that nearly 40% the programs that failed were unsuccessful due to poor pharmacokinetic profiles.⁵⁰ The candidate compounds may have been potent and selective, but they were unable to reach their intended target. Lipinski and his colleagues hypothesized that successful drugs and clinical candidates occupied specific portions of the larger chemical universe and that compounds that failed as a result of poor pharmacokinetic profiles existed outside of this “druglike” chemical space. They further proposed that if one could understand the limitation of this region of the chemical universe, then one could predict whether or not a new compound would be inside or outside of “druglike” chemical space.

The results of their studies had a profound and lasting influence on the pharmaceutical industry. They determined that the overwhelming majority of the compounds in their test set shared certain key characteristics referred to as the Lipinski Rule of Five (Figure 5.29). Successful compounds had a

Lipinski's Rule of 5 and the Veber Extension

- 1) Molecular weight <500 AMU
- 2) LogP below 5
- 3) Fewer than 5 hydrogen bond donors
- 4) Fewer than 10 hydrogen bond acceptors
- 5) Fewer than 10 rotatable bonds
- 6) Polar surface area less than 140 Å²

FIGURE 5.29 Although there are exceptions, these guidelines have been developed to help identify compounds that likely to have physicochemical properties consistent with successful drugs.

molecular weight below 500 AMUs, a logP less than 5 (the logarithm of the partition coefficient between water and 1-octanol; a measure of lipophilicity), fewer than 5 hydrogen bond donors, and fewer than 10 hydrogen bond acceptors. These guidelines to “druglike” space were based on the

90th percentile values of the properties distribution, and assume that compounds are absorbed into cells via passive diffusion (Compounds that are actively transported across a cell membrane are exceptions to these limitations.). Additional observations by Daniel Veber et al. lead to an expansion of the Lipinski guidelines to include the number of rotatable bonds (less than 10) and the concept of polar surface area (less than $140 \text{ \AA}^2$).⁵¹ There are, of course, exceptions to these guidelines, as many natural products fall outside of the scope of these rules, but the Lipinski Rule of Five has been broadly accepted by the pharmaceutical industry and are often examined via *in silico* prediction prior to preparation of proposed new compounds.

In an ideal world, a novel series of compounds could be optimized through medicinal chemistry and screening science to provide a perfect clinical candidate. Target selectivity would be absolute, eliminating the risk of off-target effects, and physicochemical properties would support a simple dosing regimen. In practice, however, it is rarely possible to identify a perfect compound. Clinical candidates, as well as the overwhelming majority of marketed drugs, have “warts” that must be considered as they are studied or used in a clinical setting. Medicinal chemistry in combination with screening science provides the tools to maximize the positive traits and minimize the negative characteristics of candidate compounds. It is up to program scientists to decide when medicinal chemistry efforts have provided candidate compounds that are suitable for further study (e.g., pharmacokinetic studies, *in vivo* animal models of efficacy, and safety studies) or that despite their best efforts, a series of compounds should be abandoned, as it is not possible to remove enough “warts” to justify further effort.

QUESTIONS

1. Define medicinal chemistry and briefly discuss what it entails.
2. Define the term structure activity relationship (SAR).
3. How does chirality impact binding of candidate compounds to their intended biological target?
4. Define the term quantitative structure activity relationship (QSAR)?
5. What is a pharmacophore?
6. What does the term auxophore refer to?
7. Describe the iterative process that is used to acquire knowledge of the SAR of a series of molecules.
8. Briefly describe fragment-based drug design.
9. What are two general approaches used to establish an?
10. What are two general approaches used in virtual screening?
11. What is a bioisostere and why are they employed?

References

1. Kelvin, W. T. *Baltimore Lectures on Molecular Dynamics and the Wave Theory of Light*; C.J. Clay and Sons, Cambridge University Press Warehouse: London, 1904.
2. a. Pohland, A. Esters of Substituted Aminobutanes US 2728779, 1955.
b. Niesenbaum, L.; Deutsch, J.; Moss, N. H. Analgesic Efficacy of Dextro-propoxyphene Hydrochloride in the Postoperative Patient. *J. Albert Einstein Med. Cent. Phila.* **1962**, *10*, 188–192.
3. Galli, A. D. Oral Compositions Containing Antitussives and Benzydamine WO 9523602 A1, 1995.
4. a. Neil, A.; Terenius, L. D-Propoxyphene Acts Differently from Morphine on Opioid Receptor-Effector Mechanisms. *Eur. J. Pharmacol.* **1981**, *69* (1), 33–39.
b. Neil, A. Affinities of Some Common Opioid Analgesics towards Four Binding Sites in Mouse Brain. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **1984**, *328* (1), 24–29.
5. a. Carter, C. H. A Clinical Evaluation of the Effectiveness of Novrad and Acetylsalicylic Acid in Children with Cough. *Am. J. Med. Sci.* **1963**, *245*, 713–717.
b. Strapkova, A.; Nosalova, G.; Korpas, J. Effects of Antitussive Drugs under Normal and Pathological Conditions. *Acta Physiol. Hung.* **1987**, *70* (2–3), 207–213.
6. Li, W.; Escarpe, P. A.; Eisenberg, E. J.; Cundy, K. C.; Sweet, C.; Jakeman, K. J.; Merson, J.; Lew, W.; Williams, M.; Zhang, L.; et al. Identification of GS 4104 as an Orally Bio-available Prodrug of the Influenza Virus Neuraminidase Inhibitor GS 4071. *Antimicrob. Agents Chemother.* **1998**, *42* (3), 647–653.
7. Varghese, J. N.; Smith, P. W.; Sollis, S. L.; Blick, T. J.; Sahasrabudhe, A.; McKimm-Breschkin, J. L.; Colman, P. M. Drug Design against a Shifting Target: A Structural Basis for Resistance to Inhibitors in a Variant of Influenza Virus Neuraminidase. *Structure* **1998**, *6*, 735–746.
8. Hammett, Louis P. The Effect of Structure upon the Reactions of Organic Compounds. Benzene Derivatives. *J. Am. Chem. Soc.* **1937**, *59*, 96–103.
9. a. Taft, R. W. Linear Free Energy Relationships from Rates of Esterification and Hydrolysis of Aliphatic and Ortho-Substituted Benzoate Esters. *J. Am. Chem. Soc.* **1952**, *74*, 2729–2732.
b. Taft, R. W. Polar and Steric Substituent Constants for Aliphatic and *o*-Benzoate Groups from Rates of Esterification and Hydrolysis of Esters. *J. Am. Chem. Soc.* **1952**, *74*, 3120.
c. Taft, R. W. Linear Steric Energy Relationships. *J. Am. Chem. Soc.* **1953**, *75*, 4538–4539.
10. a. Debnath, A. K. Quantitative Structure-Activity Relationship (QSAR) Paradigm—Hansch Era to New Millennium. *Mini-Rev. Med. Chem.* **2001**, *1* (2), 187–195.
b. Hansch, C. A Quantitative Approach to Biochemical Structure-Activity Relationships. *Acc. Chem. Res.* **1969**, *2*, 232–239.
c. Hansch, C.; Leo, A.; Taft, R. W. A Survey of Hammett Substituent Constants and Resonance and Field Parameters. *Chem. Rev.* **1991**, *91*, 165–195.
11. Topliss, J. G. Utilization of Operational Schemes for Analog Synthesis in Drug Design. *J. Med. Chem.* **1972**, *15* (10), 1006–1011.
12. Hirschfeld, R. M. A. Sertraline in the Treatment of Anxiety Disorders. *Depression Anxiety* **2000**, *11* (4), 139–157.
13. Keller, M. B. Citalopram Therapy for Depression: A Review of 10 Years of European Experience and Data from U.S. Clinical Trials. *J. Clin. Psychiatry* **2000**, *61* (12), 896–908.
14. Wong, D. T.; Perry, K. W.; Bymaster, F. P. The Discovery of Fluoxetine Hydrochloride (Prozac). *Nat. Rev. Drug Discov.* **2005**, *4*, 764–774.
15. Wermuth, C. G.; Ganellin, C. R.; Lindberg, P.; Mitscher, L. A. Glossary of Terms Used in Medicinal Chemistry (IUPAC Recommendations 1998). *Pure Appl. Chem.* **1998**, *70* (5), 1129–1143.
16. Novak, B. H.; Hudlicky, T.; Reed, J. W.; Mulzer, J.; Trauner, D. Morphine Synthesis and Biosynthesis – An Update. *Curr. Org. Chem.* **2000**, *4*, 343–362.

17. Prommer, E. Levorphanol: The Forgotten Opioid. *Supportive Care Cancer* **2007**, *15* (3), 259–264.
18. a. Berzetei-Gurske, I.; Loew, G. H. The Novel Antagonist Profile of (-)-metazocine. *Prog. Clin. Biol. Res.* **1990**, *328*, 33–36.
 b. Hori, M.; Ban, M.; Imai, E.; Iwata, N.; Suzuki, Y.; Baba, Y.; Morita, T.; Fujimura, H.; Nozaki, M.; Niwa, M. Novel Nonnarcotic Analgesics with an Improved Therapeutic Ratio. Structure-Activity Relationships of 8-(methylthio)- and 8-(acylthio)-1, 2,3,4,5,6-hexahydro-2,6-methano-3-benzazocines. *J. Med. Chem.* **1985**, *28* (11), 1656–1661.
19. Kaiko, R. F.; Foley, K. M.; Grabinski, P. Y.; Heidrich, G.; Rogers, A. G.; Inturrisi, C. E.; Reidenberg, M. M. Central Nervous System Excitatory Effects of Meperidine in Cancer Patients. *Ann. Neurol.* **1983**, *13* (2), 180–185.
20. Silverman, R. B. *The Organic Chemistry of Drug Design and Drug Action*, 2nd ed.; Elsevier Academic Press: Oxford, UK, 2004. 17–20.
21. Liantonio, L.; Picollo, A.; Carbonara, G.; Fracchiolla, G.; Tortorella, P.; Loiodice, F.; Laghezza, A.; Babini, E.; Zifarelli, G.; Pusch, M.; et al. Molecular Switch for CLC-K Cl-Channel Block/Activation: Optimal Pharmacophoric Requirements towards High-Affinity Ligands. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 41369–41373.
22. Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G. ZINC: A Free Tool to Discover Chemistry for Biology. *J. Chem. Inf. Model.* **2012**, *52* (7), 1757–1768.
23. a. Chothia, C.; Lesk, A. M. The Relation between the Divergence of Sequence and Structure in Proteins. *EMBO J.* **1986**, *5* (4), 823–826.
 b. Kaczanowski, S.; Zielenkiewicz, P. Why Similar Protein Sequences Encode Similar Three-Dimensional Structures? *Theor. Chem. Acc.* **2010**, *125*, 543–550.
24. a. Sharma, H.; Cheng, X.; Buolamwini, J. K. Homology Model-Guided 3D-QSAR Studies of HIV-1 Integrase Inhibitors. *J. Chem. Inf. Model.* **2012**, *52* (2), 515–544.
 b. Joshi, U. J.; Shah, F. H.; Tikhele, S. H. Homology Model of the Human 5-HT1A Receptor Using the Crystal Structure of Bovine Rhodopsin. *Internet Electron. J. Mol. Des.* **2006**, *5* (7), 403–415.
25. a. Menon, V.; Vallat, B. K.; Dybas, J. M.; Fiser, A. Modeling Proteins Using a Super-Secondary Structure Library and NMR Chemical Shift Information. *Structure* **2013**, *21* (6), 891–899.
 b. Kitchen, D.; Hoffman, R. C.; Moy, F. J.; Powers, R. Homology Model for Oncostatin M Based on NMR Structural Data. *Biochemistry* **1998**, *37* (30), 10581–10588.
26. a. Henderson, R.; Schertler, G. F. X. The Structure of Bacteriorhodopsin and Its Relevance to the Visual Opsins and Other Seven Helix G-Protein Coupled Receptors. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* **1990**, *326*, 379–389.
 b. Ovchinnikov, Y. A. Rhodopsin and Bacteriorhodopsin: Structure Function Relationships. *FEBS Lett.* **1982**, *148*, 179–191.
27. Teeter, M. T.; Froimowitz, M.; Stec, B.; DuRand, C. J. Homology Modeling of the Dopamine D2 Receptor and Its Testing by Docking of Agonists and Tricyclic Antagonists. *J. Med. Chem.* **1994**, *37*, 2874–2888.
28. a. Fachinger, G.; Deutsch, U.; Risau, W. *Oncogene* **1999**, *18*, 5948–5953.
 b. Krueger, N. X.; Streuli, M.; Saito, H. *EMBO J.* **1990**, *9*, 3241–3252.
 c. Wright, M. B.; Seifert, R. A.; Bowen-Pope, D. F. *Arterioscler. Thromb. Vasc. Biol.* **2000**, *20*, 1189–1198.
29. Evdokimov, A. G.; Pokross, M.; Walter, R.; Mekel, M.; Cox, B.; Li, C.; Bechard, R.; Genbaffe, F.; Andrews, R.; Diven, C.; et al. Engineering the Catalytic Domain of Human Protein Tyrosine Phosphatase Beta for Structure-Based Drug Discovery. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2006**, *D62*, 1435–1445.
30. Murray, C. W.; Rees, D. C. The Rise of Fragment-Based Drug Discovery. *Nat. Chem.* **2009**, *1*, 187–192.

31. Hartshorn, M. J.; Murray, C. W.; Cleasby, A.; Frederickson, M.; Tickle, I. J.; Jhoti, H. Fragment-Based Lead Discovery Using X-ray Crystallography. *J. Med. Chem.* **2005**, *48*, 403–413.
32. Lepre, C. A.; Moore, J. M.; Peng, J. W. Theory and Applications of NMR-Based Screening in Pharmaceutical Research. *Chem. Rev.* **2004**, *104*, 3641–3676.
33. Neumann, T.; Junker, H. D.; Schmidt, K.; Sekul, R. SPR-Based Fragment Screening: Advantages and Applications. *Curr. Top. Med. Chem.* **2007**, *7*, 1630–1642.
34. Hann, M. M.; Leach, A. R.; Harper, G. Molecular Complexity and Its Impact on the Probability of Finding Leads for Drug Discovery. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 856–864.
35. Wyatt, P. G.; Woodhead, A. J.; Berdini, V.; Boulstridge, J. A.; Carr, M. G.; Cross, D. M.; Davis, D. J.; Devine, L. A.; Early, T. R.; Feltell, R. E.; et al. Identification of N-(4-piperidinyl)-4-(2,6-dichlorobenzoylamino)-1H-pyrazole-3-carboxamide (AT7519), a Novel Cyclin Dependent Kinase Inhibitor Using Fragment-Based X ray Crystallography and Structure Based Drug Design. *J. Med. Chem.* **2008**, *51*, 4986–4999.
36. Langmuir, I. Isomorphism, Isosterism and Covalence. *J. Am. Chem. Soc.* **1919**, *41*, 1543–1559.
37. Erlenmeyer, H.; Leo, M. Über Pseudoatome. *Helv. Chim. Acta* **1932**, *15*, 1171–1186.
38. Friedman, H. L. *Influence of Isosteric Replacements upon Biological Activity*, Vol. 206; National Academy of Sciences National Research Council Publication, 1951. 295.
39. Burger, A. *Medicinal Chemistry*, 3rd ed.; John Wiley & Sons: New York, 1970. p. 127.
40. Schann, S.; Bruban, V.; Pompermayer, K.; Feldman, J.; Pfeiffer, B.; Renard, P.; Scalbert, E.; Bousquet, P.; Ehrhardt, J. D. Synthesis and Biological Evaluation of Pyrrolidinic Isosteres of Rilmenidine. Discovery of Cis-/trans-Dicyclopropylmethyl-(4,5-dimethyl-4,5-dihydro-3H-pyrrol-2-yl)-amine (LNP 509), an I₁ Imidazoline Receptor Selective Ligand with Hypotensive Activity. *J. Med. Chem.* **2001**, *44* (10), 1588–1593.
41. Wildsmith, J. A. W.; Gissen, A. J.; Takman, B.; Covino, B. G. Differential Nerve Blockade: Esters versus Amides and the Influence of PKa. *Br. J. Anaesth.* **1987**, *59* (3), 379–384.
42. Essery, J. M. Preparation and Antibacterial Activity of A-(5-Tetrazolyl)benzylpenicillin. *J. Med. Chem.* **1969**, *12*, 703–705.
43. Glass, R.; Loring, J.; Spencer, J.; Villee, C. The Estrogenic Properties *in Vitro* of Diethylstilbestrol and Substances Related to Estradiol. *Endocrinology* **1961**, *68*, 327–333.
44. a. Lima, L. M.; Barreiro, E. J. Bioisosterism: A Useful Strategy for Molecular Modification and Drug Design. *Curr. Med. Chem.* **2005**, *12*, 23–49.
 b. Sethy, S. P.; Meher, C. P.; Biswal, S.; Sahoo, U.; Patro, S. K. The Role of Bioisosterism in Molecular Modification and Drug Design: A Review. *Asian J. Pharm. Sci. Res.* **2013**, *3* (1), 61–87.
 c. Patani, G. A.; LaVoie, E. J. Bioisosterism: A Rational Approach in Drug Design. *Chem. Rev.* **1996**, *96*, 3147–3176.
45. a. Wang, Z.; Fermini, B.; Nattel, S. Evidence for a Novel Delayed Rectifier K⁺ Current Similar to Kv1.5 Cloned Channel Currents. *Circ. Res.* **1993**, *73*, 1061–1076.
 b. Fedida, D.; Wible, B.; Wang, Z.; Fermini, B.; Faust, F.; Nattel, S.; Brown, A. M. Identity of a Novel Delayed Rectifier Current from Human Heart with a Cloned Potassium Channel Current. *Circ. Res.* **1993**, *73*, 210–216.
 c. Brendel, J.; Peukert, S. Blockers of the Kv1.5 Channel for the Treatment of Atrial Arrhythmias. *Expert Opin. Ther. Pat.* **2002**, *12* (11), 1589–1598.
46. a. Taglialatela, M.; Castaldo, P.; Pannaccione, A. Human Ether-a-gogo Related Gene (HERG) K Channels as Pharmacological Targets: Present and Future Implications. *Biochem. Pharmacol.* **1998**, *55* (11), 1741–1746.
 b. Vaz, R. J.; Li, Y.; Rampe, D. Human Ether-a-go-go Related Gene (HERG): A Chemist's Perspective. *Prog. Med. Chem.* **2005**, *43*, 1–18.
 c. Kang, J.; Wang, L.; Chen, X. L.; Triggle, D. J.; Rampe, D. Interactions of a Series of Fluoroquinolone Antibacterial Drugs with the Human Cardiac K⁺ Channel HERG. *Mol. Pharmacol.* **2001**, *59*, 122–126.

47. Virshup, A. M.; Contreras-García, J.; Wipf, P.; Yang, W.; David, N.; Beratan, D. N. Stochastic Voyages into Uncharted Chemical Space Produce a Representative Library of All Possible Drug-Like Compounds. *J. Am. Chem. Soc.* **2013**, *135*, 7296–7303.
48. a. Reeder, E.; Sternbach, L. H. Process for Preparing 5-phenyl-1,2-dihydro-3H-1,4-benzodiazepines US3109843, 1963.
b. Schlager, L. H. Novel 3-hydroxy-1,4-benzodiazepine-2-ones and Process for the Preparation Thereof US4388313, 1983.
c. Hester, J. B. 6-Phenyl-4H-s-triazolo[4,3-a][1,4]benzodiazepines US3987052, 1976.
d. Tachikawa, R.; Takagi, H.; Miyadera, T.; Kamioka, T.; Fukunaga, M.; Kawano, Y. *Antidepressant Benzodiazepine*, DE 1812252; 1968.
49. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. *Adv. Drug Deliv. Rev.* **1997**, *23*, 3–25.
50. DiMasi, J. A.; Hansen, R. W.; Grabowski, H. G.; Lasagna, L. Cost of Innovation in the Pharmaceutical Industry. *J. Health Econ.* **1991**, *10*, 107–142.
51. Veber, D. F.; Johnson, S. R.; Cheng, H. Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D. Molecular Properties that Influence the Oral Bioavailability of Drug Candidates. *J. Med. Chem.* **2002**, *45*, 2615–2623.

This page intentionally left blank

Copyright © 2015. Elsevier Science & Technology. All rights reserved.