Drug Discovery Research
New Frontiers in the Post-Genomic Era

Ziwei Huang
DRUG DISCOVERY
RESEARCH
Each generation has its unique needs and aspirations. When Charles Wiley first opened his small printing shop in lower Manhattan in 1807, it was a generation of boundless potential searching for an identity. And we were there, helping to define a new American literary tradition. Over half a century later, in the midst of the Second Industrial Revolution, it was a generation focused on building the future. Once again, we were there, supplying the critical scientific, technical, and engineering knowledge that helped frame the world. Throughout the 20th Century, and into the new millennium, nations began to reach out beyond their own borders and a new international community was born. Wiley was there, expanding its operations around the world to enable a global exchange of ideas, opinions, and know-how.

For 200 years, Wiley has been an integral part of each generation’s journey, enabling the flow of information and understanding necessary to meet their needs and fulfill their aspirations. Today, bold new technologies are changing the way we live and learn. Wiley will be there, providing you the must-have knowledge you need to imagine new worlds, new possibilities, and new opportunities.

Generations come and go, but you can always count on Wiley to provide you the knowledge you need, when and where you need it!
DRUG DISCOVERY RESEARCH
New Frontiers in the Post-Genomic Era

Edited by

ZIWEI HUANG
San Diego, California

WILEY-INTERSCIENCE
A JOHN WILEY & SONS, INC., PUBLICATION
## CONTENTS

**FOREWORD** ix  
**INTRODUCTION** xi  
**CONTRIBUTORS** xv  

### PART I. COMPUTATIONAL AND STRUCTURAL APPROACHES IN DRUG DISCOVERY

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MOLECULAR DOCKING AND STRUCTURE-BASED DESIGN</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>P. Therese Lang¹, Tiba Aynechi¹, Demetri Moustakas², Brian Shoichet¹,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Irwin D. Kuntz³*, Natasja Brooijmans¹, and Connie M. Oshiro³</td>
<td></td>
</tr>
<tr>
<td></td>
<td>¹University of California, San Francisco, CA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>²University of California, Berkeley, CA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>³Roche, Palo Alto, CA</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>RECENT ADVANCES IN DRUG DISCOVERY RESEARCH USING STRUCTURE-BASED VIRTUAL SCREENING TECHNIQUES: EXAMPLES OF SUCCESS FOR DIVERSE PROTEIN TARGETS</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>Sutapa Ghosh¹, Aihua Nie¹, Jing An¹, and Ziwei Huang¹,²*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>¹Burnham Institute for Medical Research La Jolla, CA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>²University of California, San Diego, CA</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>VIRTUAL SCREENING IN DRUG DISCOVERY</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>Malcolm J. McGregor, Zhaowen Luo, and Xuliang Jiang</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Serono Research Institute, Rockland, MA</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>COMPUTER-AIDED DRUG DESIGN</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>Grace Shiahuy Chen¹ and Ji-Wang Chern²*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>¹Providence University, Shalu, Taiwan, Republic of China</td>
<td></td>
</tr>
<tr>
<td></td>
<td>²National Taiwan University, Taipei, Taiwan, Republic of China</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>FOCUSED LIBRARY DESIGN BASED ON HIT AND TARGET STRUCTURES: METHOD AND APPLICATION IN DRUG DISCOVERY</td>
<td>108</td>
</tr>
<tr>
<td></td>
<td>Weiliang Zhu¹, Jian Li¹, Zhen Gong¹, Hong Liu¹, and Hualiang Jiang¹,²*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>¹Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, P.R. China</td>
<td></td>
</tr>
<tr>
<td></td>
<td>²East China University of Science and Technology, Shanghai, P.R. China</td>
<td></td>
</tr>
</tbody>
</table>
PART II. CHEMICAL AND SYNTHETIC APPROACHES IN DRUG DISCOVERY

CHAPTER 6. BEYOND NATURAL PRODUCTS: SYNTHETIC ANALOGUES OF BRYOSTATIN 1
Paul A. Wender*, Jeremy L. Baryza, Michael K. Hilinski, Joshua C. Horan, Cindy Kan, and Vishal A. Verma
Stanford University, Stanford, CA

CHAPTER 7. THE PROMISES AND PITFALLS OF SMALL-MOLECULE INHIBITION OF POLY(ADP-RIbose) GLYCOHYDROLASE (PARG)
Amanda C. Nottbohm and Paul J. Hergenrother*
University of Illinois at Urbana-Champaign, Urbana, IL

CHAPTER 8. CYCLIC ADP-RIbose ANALOGUES WITH MINIMAL STRUCTURE: SYNTHESIS AND CALCIUM-RELEASE ACTIVITY
Lihe Zhang¹ and Andreas H. Guse²
¹Peking University, Beijing, P.R. China
²University Hospital Hamburg—Eppendorf, Institute of Biochemistry and Molecular Biology, Hamburg, Germany

CHAPTER 9. EVALUATION OF NEUROIMMUNOPHILIN LIGANDS
Junhai Xiao, Aihua Nie, and Song Li*
Beijing Institute of Pharmacology and Toxicology, Beijing, P.R. China

CHAPTER 10. CHIRAL DRUGS AND THE ASSOCIATED ASYMMETRIC SYNTHESIS
Jiange Zhang¹ and Guoqiang Lin²*
¹Zhengzhou University, Zhengzhou, P.R. China
²Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Shanghai, P.R. China

CHAPTER 11. ALPHA-HELIX MIMETICS IN DRUG DISCOVERY
Hang Yin, Gui-In Lee, and Andrew D. Hamilton*
Yale University, New Haven, CT

CHAPTER 12. BASIC AND TRANSLATIONAL RESEARCH OF CHEMOKINE LIGANDS AND RECEPTORS AND DEVELOPMENT OF NOVEL THERAPEUTICS
Won-Tak Choi¹, Yohichi Kumaki², I. M. Krishna Kumar³, Jing An³, Douglas D. Richman⁴, Joseph G. Sodroski⁵, and Ziwei Huang¹,³,⁴*
¹University of Illinois at Urbana-Champaign, Urbana, IL
²Raylight Corporation, Chemokine Pharmaceutical Inc., La Jolla, CA
³Burnham Institute for Medical Research, La Jolla, CA
⁴Department of Molecular Pathology, University of California, San Diego, CA
⁵Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA
### PART III. BIOLOGICAL APPROACHES AND TRANSLATIONAL RESEARCH IN DRUG DISCOVERY

<table>
<thead>
<tr>
<th>CHAPTER 15.</th>
<th>RETINOIC ACID AND ARSENIC TRIOXIDE TREATMENT IN ACUTE promyeloctic leukemia: A MODEL OF ONCOProtein Targeted Therapy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Jian-Hua Tong, Sai-Juan Chen, and Zhu Chen*</td>
</tr>
<tr>
<td></td>
<td>Shanghai Institute of Hematology, Rui Jin Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, P.R. China</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER 16.</th>
<th>2-CHLORODEOXYADENOSINE (CLADRIBINE): RATIONAL DEVELOPMENT OF A NOVEL CHEMOTHERAPEUTIC AGENT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Howard B. Cottam and Dennis A. Carson*</td>
</tr>
<tr>
<td></td>
<td>Moores Cancer Center, University of California, San Diego, CA</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER 17.</th>
<th>APOPTOSIS-BASED DRUG DISCOVERY FOR CANCER</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>John C. Reed*, Dayong Zhai, Marc Hyer, and Kate Welsh</td>
</tr>
<tr>
<td></td>
<td>Burnham Institute for Medical Research, La Jolla, CA</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER 18.</th>
<th>MECHANISM-BASED DEVELOPMENT OF MEMANTINE AS A THERAPEUTIC AGENT IN TREATING ALZHEIMER’S DISEASE AND OTHER NEUROLOGIC DISORDERS: LOW-AFFINITY, UNCOMPETITIVE ANTAGONISM WITH FAST OFF-RATE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Huei-Sheng Vincent Chen*, and Stuart A. Lipton*</td>
</tr>
<tr>
<td></td>
<td>Burnham Institute for Medical Research, La Jolla, CA, University of California, San Diego, CA</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER 19.</th>
<th>BRIDGING BENCH TO CLINIC: ROLES OF ANIMAL MODELS FOR POST-GENOMICS DRUG DISCOVERY ON METABOLIC DISEASES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Chi-Wai Wong and Ling Chen*</td>
</tr>
<tr>
<td></td>
<td>Guangzhou Institute of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou, P.R. China</td>
</tr>
</tbody>
</table>
CHAPTER 20. STEM CELL RESEARCH AND APPLICATIONS FOR HUMAN THERAPIES

Prithi Rajan, Kook In Park, Vaclav Ourednik, Jean Pyo Lee, Jamie Imitola, Franz-Joseph Mueller, Yang D. Teng, and Evan Snyder*

Burnham Institute for Medical Research, La Jolla, CA

INDEX

*denotes the author(s) whom should be contacted for all correspondences regarding the specific chapter.
FOREWORD

NIGHT THOUGHTS OF A CHEMIST BEMUSED BY BIOTECHNOLOGY

This book covers a wide variety of topics of interest to the pharmaceutical industry, and does so very well. The writing is usually clear, the illustrations apt, and the references generous. There is some sense, however, of desperation, which probably reflects the awareness of the impossible task now facing those in that industry. No matter what they do, or how well, it will never be enough, for either the patients or the company profits. There are good reasons for such uneasiness. The task, as now posed, may be impossible.

Biology, and therefore human biomedicine, is a product of evolution, and is thus a palimpsest of all that has gone before. Each variation, genetic or epigenetic, has consequences, throughout the organism, as it modifies catalytic and structural proteins and the various feedbacks loops in which they are imbedded. If, on balance, the various changes leave the organism more fit to survive and reproduce in the environment at that time, the change may become widespread, even if some of its effects are deleterious. The effects produce cascades of small differences in different populations, leading to a balanced diversity in their genomes and in metabolisms. This means that each drug acts against a different background. Those that act on the genome or on its expression or, nominally, one simple receptor, similarly change the physiology and its potential response to further intervention. As a result of this variability and plasticity of the matrix, a drug, in extreme cases, can kill one person and cure another, so that clinical trials are always veiled by some level of statistical mystification.

What can be done to ameliorate this problem, which is rapidly making a “red queen’s race” of the industry? Bacterial infections were, and are, relatively straightforward problems, because bacteria are relatively far phylogenetically from eukaryotes and hence some key features of their metabolism are strongly uncoupled those of polycellular organisms. But metabolic and genetic aberrations in patients are very similar to those in normal tissues, and hence their selective targeting is far more subtle. Can any response but a deadly game of cut and try be used in such circumstances? Perhaps not easily.

Because biological evolution in general is exceedingly conservative and modular, the histories of biomedical cycles and genetic systems are represented, often in only slightly distorted forms, in contemporary organisms. That information, and the connections within the systems themselves, can lead to a deeper understanding of
its responses to perturbations. Obtaining and using this information will not be a quick or easy task. Some aspects may be very generally applicable, however. The lifespan of a mouse is very different from that of an elephant, but would a mouse altered to live for many decades still be a mouse? Only a very deep analysis can tell us how to do new things and what the side-effects might be. The kind of work described in this book is laying the foundation for such understanding and for the skills required to make use of it.

Paul C. Lauterbur

_Urbana, Illinois_

March 2007

*Note by the Editor: Shortly after writing this Foreword, Professor Paul C. Lauterbur, 2003 Nobel laureate in medicine, passed away on March 27, 2007. This book is dedicated to the memory of this great man who has had a life-long passion in applying scientific innovations to alleviate profound human sufferings.
INTRODUCTION

This book is a product of the collective effort and contribution of 70 scholars and researchers from 25 universities, institutes and companies around the world who are experts in a wide range of different disciplines related to biomedical and drug discovery research but share a common interest in translating research discoveries into new therapies for human diseases. In this book, they have shared their knowledge, recent research discoveries, and future perspectives of the various aspects of drug discovery research.

The recent completion of the human genome project has ushered in a new era of post-genomic scientific research and biomedical applications. As shown in the central illustration on the front cover of the book, the rapid advances in many scientific disciplines such as genomics, structure, chemistry, computer-aided drug design (CADD), biology, and pharmacology and their integration are changing the face of drug discovery research and facilitating the translation of the vast information from the human genome to the understanding of the biology of the human body and development of new medicine. As such, the book focuses on the interplay of the research in these fields for the discovery of new drugs and describes some of the most fascinating and cross-disciplinary frontiers in post-genomic biomedical and biopharmaceutical science, as written by experts working in such research frontiers. The intent of the book is to provide both graduate students and researchers who study, work, or have an interest in biotechnology and pharmaceutical science with a general introduction to and practical examples of some of the many methods and techniques that can be used in biopharmaceutical research in the post-genomic era.

The book has a total of 20 chapters, divided into three different sections: Part I, Computational and Structural Approaches in Drug Discovery (Chapters 1–5); Part II, Chemical and Synthetic Approaches in Drug Discovery (Chapters 6–14); and Part III, Biological Approaches and Translational Research in Drug Discovery (Chapters 15–20). Their arrangement and order reflect the overall scheme of modern drug discovery research shown on the cover illustration, which involves structure-based and computer-aided drug design, chemical synthesis and optimization of drug molecules, and biological and preclinical evaluation. The five chapters in Part I describe the recent advances in the development of computational methods and their application in virtual screening to identify new ligands or inhibitors of a protein target. In addition to receptor binding or so-called molecular docking simulations of ligands by virtual screening (Chapters 1–4), this section describes the development and use of computational techniques to predict other properties such as metabolism and toxicity of drug candidates (Chapter 3) or to design novel combinatorial
compound libraries to be synthesized (Chapter 5). Part II has a collection of nine chapters devoted to the advances in chemical synthesis and chemical strategies for the development or delivery of both small and large molecule drugs or probes of protein functions. Chapters 6–11 cover a wide range of chemistries of small molecules of various natures and their application for the discovery of drugs and research probes. As small molecules have been a major source of drugs for the pharmaceutical industry, these chapters provide an important account of the state-of-the-art chemical methods in the research and discovery of these small molecule drugs or probes. Besides small molecules, synthetic molecules of larger sizes – such as chemically modified proteins or peptides – are becoming another important class of therapeutic and research agents, as described in Chapter 12 which uses synthetic chemokine analogs as an example. Chapter 13 reviews the chemical strategies for studying post-translational modification proteomics which has a significant implication on the development of protein-based therapeutics. The merging of nanotechnology and biomedicine is becoming an important trend. The role of chemistry in nano-medicine is highlighted in Chapter 14, which presents the development of nano-materials for the delivery and controlled release of drugs, using antibiotics as the example. Part III focuses on the biology of drug discovery, with six chapters reviewing biological and translational research with the goal of bringing novel laboratory research agents to the clinic. Chapters 15–18 provide an account of some of the most remarkable achievements and successful case stories in modern research and development of anticancer drugs and drugs for treating neurodegenerative disease. Chapter 19 reviews the role of animal models in drug discovery. Last but not least, research and applications of stem cells are becoming a highly promising new avenue for biomedical and drug discovery research, as described in Chapter 20.

It should be pointed out that the classification of the chapters into the three different sections and the overtly simplified summary as described above are limited and partial. As readers will find out when reading each of the 20 chapters in detail, many chapters are actually multidisciplinary and cross the boundaries of these sections. Indeed, one salient feature that this book wishes to present is the cross-disciplinary and collaborative nature of modern drug discovery research as demonstrated by these 20 chapters written by biologists, chemists, and clinicians whose works are highly interrelated to each other. Another important feature of the book is its focus on post-genomic drug discovery research in the United States and China. As China is becoming a global powerhouse in research, this book hopes to showcase some of the latest and most exciting research in drug discovery in China, together with the work of colleagues in the United States.

Readers of this book should note that a corresponding author is denoted for each chapter and his/her e-mail address provided in the Contributors pages. Readers who have questions or inquiries about a particular chapter are encouraged to communicate with the corresponding author directly through the e-mail address provided. A complete list of authors, corresponding authors, and affiliations can be found in the Table of Contents. Please also note that many chapters have color figures. All color figures are placed together in the center of the book. While black-and-white copies (so stated specifically in the figure captions) of these color figures are provided within
each chapter, readers should refer to the color versions for better comprehension of these figures.

I would like to thank all those who have helped with and contributed to this book. Most importantly, I want to acknowledge all my co-authors of this book. I won’t name them one by one here as there are 69 of them! Not only are they the pioneers, leaders, or experts in their fields of research, many also are or have become good friends and colleagues through our work on this book. It has been a great experience for me to work with these outstanding individuals and bring their fascinating research to the readers. The knowledge, insights, and perspectives shared by these wonderful co-authors are truly valuable to me and, I hope, to the readers. The final piece of the book was completed by my former colleague at the University of Illinois at Urbana-Champaign (UIUC), Professor Paul C. Lauterbur, who kindly wrote the Foreword. The idea of putting together this book was first conceived in 2003 when I was at UIUC where Paul and I were both faculty members of the Department of Chemistry and the Center for Biophysics and Computational Biology. The book was started there and continued in San Diego with the moving of my laboratory to San Diego in 2004. In this sense, it is personally satisfying to see that, in a somewhat serendipitous way, this book began and ended at Illinois. Recognized by the 2003 Nobel Prize in Physiology or Medicine, Paul Lauterbur’s pioneering research has led to the development of modern magnetic resonance imaging (MRI), which represents one of the most significant medical diagnostic discoveries of the 20th century and has benefited patients throughout the world. This classic work demonstrates the power of cross-disciplinary research to discover innovative medical applications, which is the theme and goal of this book.

In addition to the science, many people deserve recognition for their efforts and help in other aspects of the book. Cindy Cook, Jonathan Rose and Danielle Lacourciere have tirelessly worked on the editing and production of the book. Sameer Kawatkar, Serena Chen, Pak-Nei Hon, Mi Youn Lim and Harsukh Gevariya have helped with the cover illustrations and design.

As the editor and on behalf of my 69 co-authors, I hope that this book can be a helpful tool and reference for your study and research of post-genomic drug discovery, an endeavor filled with difficult challenges for scientific intellect and great promises for human health.

Ziwei Huang

San Diego, California

March 2007
CONTRIBUTORS

Jing An, University of Illinois at Urbana-Champaign, Urbana, IL 61801; The Burnham Institute for Medical Research, La Jolla, CA 92037; and Raylight Corporation, Chemokine Pharmaceutical Inc., La Jolla, CA 92037, USA

Tiba Aynechi, Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, CA 94143, USA

Jeremy L. Baryza, Department of Chemistry, Stanford University, Stanford, CA 94305, USA

Natasja Brooijmans, Graduate Program in Chemistry and Chemical Biology, University of California, San Francisco, San Francisco, CA 94143, USA

Dennis A. Carson, Moores Cancer Center, University of California, San Diego, La Jolla, CA 92093, USA; dcarson@ucsd.edu

Grace Shiahuy Chen, Department of Applied Chemistry, Providence University, Shalu, Taiwan, Republic of China

Huei-Sheng Vincent Chen, The Burnham Institute for Medical Research, University of California, San Diego, La Jolla, CA 92037, USA; hsv_chen@burnham.org

Ling Chen, Guangzhou Institute of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou, People’s Republic of China; chen_ling@gibh.ac.cn

Sai-Juan Chen, Shanghai Institute of Hematology (SIH) and State Key Lab of Medical Genomics, Rui Jin Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, 200025, People’s Republic of China

Zhu Chen, Shanghai Institute of Hematology (SIH) and State Key Lab of Medical Genomics, Rui Jin Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, 200025, People’s Republic of China; zchen@stn.sh.cn

Ji-Wang Chern, School of Pharmacy, National Taiwan University, Taipei 100, Taiwan, Republic of China; chern@jwc.mc.ntu.edu.tw

Won-Tak Choi, Department of Biochemistry, University of Illinois at Urbana—Champaign, Urbana, IL 61801, USA

Howard B. Cottam, Moores Cancer Center, University of California, San Diego, La Jolla, CA 92093, USA

Sutapa Ghosh, The Burnham Institute for Medical Research, La Jolla, CA 92037
Zhen Gong, Drug Discovery & Design Centre, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, People’s Republic of China

Yanting Guo, Key Laboratory of Bioorganic Phosphorus Chemistry and Chemical Biology, Department of Chemistry, Tsinghua University, Beijing 100084, People’s Republic of China

Andreas H. Guse, University Hospital Hamburg—Eppendorf, Center of Experimental Medicine, Institute of Biochemistry and Molecular Biology 1: Cellular Signal Transduction, 20246 Hamburg, Germany

Andrew D. Hamilton, Department of Chemistry, Yale University, New Haven, CT 06520, USA; andrew.hamilton@yale.edu

Paul J. Hergenrother, Department of Chemistry, Roger Adams Laboratory, University of Illinois, Urbana, IL 61801, USA; hergenro@uiuc.edu

Michael K. Hilinski, Department of Chemistry, Stanford University, Stanford, CA 94305, USA

Joshua C. Horan, Department of Chemistry, Stanford University, Stanford, CA 94305, USA

Jia Hu, Key Laboratory of Bioorganic Phosphorus Chemistry and Chemical Biology, Department of Chemistry, Tsinghua University, Beijing 100084, People’s Republic of China

Ziwei Huang, University of Illinois at Urban-Champaign, Urbana, IL 61801; The Burnham Institute for Medical Research and University of California, San Diego, La Jolla, CA 92037, USA; ziweihuang@burnham.org

Marc Hyer, The Burnham Institute for Medical Research, La Jolla, CA 92037, USA

Jamie Imitola, Program in Stem Cells & Regeneration, Center for Neuroscience and Aging, Burnham Institute for Medical Research, La Jolla, CA 92037, USA

Hualiang Jiang, Drug Discovery & Design Centre, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203; and School of Pharmacy, East China University of Science and Technology, Shanghai 200237, People’s Republic of China; hljiang@mail.shcnc.ac.cn

Xuliang Jiang, Serono Research Institute, Rockland, MA 02370, USA; xuliang.jiang@serono.com

Cindy Kan, Department of Chemistry, Stanford University, Stanford, CA 94305, USA

Yohichi Kumaki, Raylight Corporation, Chemokine Pharmaceutical Inc., La Jolla, CA 92037, USA

I. M. Krishna Kumar, The Burnham Institute for Medical Research, La Jolla, CA 92037, USA
Irwin D. Kuntz, Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, CA 94143, USA; kuntz@cgl.ucsf.edu

P. Therese Lang, Graduate Program in Chemistry and Chemical Biology, University of California, San Francisco, San Francisco, CA 94143, USA

Gui-In Lee, Department of Chemistry, Yale University, New Haven, CT 06520, USA

Jean Pyo Lee, Program in Stem Cells & Regeneration, Center for Neuroscience and Aging, The Burnham Institute for Medical Research, La Jolla, CA 92037, USA

Jian Li, Drug Discovery & Design Centre, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, People’s Republic of China

Song Li, Beijing Institute of Pharmacology and Toxicology, Haidian District, Beijing 100850, People’s Republic of China; lis@nic.bmi.ac.cn

Yanmei Li, Key Laboratory of Bioorganic Phosphorus Chemistry and Chemical Biology, Department of Chemistry, Tsinghua University, Beijing 100084, People’s Republic of China; liym@mail.tsinghua.edu.cn

Guoqiang Lin, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Shanghai 201203, People’s Republic of China; lingq@mail.sioc.ac.cn

Stuart A. Lipton, The Burnham Institute for Medical Research, University of California, San Diego, La Jolla, CA 92037, USA; slipton@burnham.org

Hong Liu, Drug Discovery & Design Centre, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, People’s Republic of China

Zhaowen Luo, Serono Research Institute, Rockland, MA 02370, USA

Malcolm J. McGregor, Serono Research Institute, Rockland, MA 02370, USA

Demetri Moustakas, Joint Graduate Program in Bioengineering, University of California, Berkeley, Berkeley, CA 94720, USA

Franz-Joseph Mueller, Program in Stem Cells & Regeneration, Center for Neuroscience and Aging, The Burnham Institute for Medical Research, La Jolla, CA 92037, USA

Aihua Nie, Beijing Institute of Pharmacology and Toxicology, Haidian District, Beijing, People’s Republic of China, 100850; and The Burnham Institute for Medical Research, La Jolla, CA 92037, USA

Amanda C. Nottbohm, Department of Chemistry, Roger Adams Laboratory, University of Illinois, Urbana, IL 61801, USA

Connie M. Oshiro, Roche, Palo Alto, CA 94304, USA

Vaclav Ourednik, Program in Stem Cells & Regeneration, Center for Neuroscience and Aging, The Burnham Institute for Medical Research, La Jolla, CA 92037, USA
Kook In Park, Program in Stem Cells and Regeneration, Center for Neuroscience and Aging, The Burnham Institute for Medical Research, La Jolla, CA 92037, USA

Prithi Rajan, Program in Stem Cells & Regeneration, Center for Neuroscience and Aging, The Burnham Institute for Medical Research, La Jolla, CA 92037, USA

John C. Reed, The Burnham Institute for Medical Research, La Jolla, CA 92037, USA; jreed@burnham.org

Douglas D. Richman, Department of Molecular Pathology, University of California, San Diego, La Jolla, CA 92037, USA

Brian Shoichet, Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, CA 94143, USA

Evan Snyder, Program in Stem Cells & Regeneration, Center for Neuroscience and Aging, The Burnham Institute for Medical Research, La Jolla, CA 92037, USA; esnyder@burnham.org

Joseph G. Sodroski, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02115, USA

Yang D. Teng, Program in Stem Cells & Regeneration, Center for Neuroscience and Aging, The Burnham Institute for Medical Research, La Jolla, CA 92037, USA

Jian-Hua Tong, Shanghai Institute of Hematology (SIH) and State Key Lab of Medical Genomics, Rui Jin Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, 200025, People’s Republic of China

Vishal A. Verma, Department of Chemistry, Stanford University, Stanford, CA 94305, USA

Yen Wei, Department of Chemistry, Drexel University, Philadelphia, PA 19104, USA; weiyen@drexel.edu

Kate Welsh, The Burnham Institute for Medical Research, La Jolla, CA 92037, USA

Paul A. Wender, Department of Chemistry, Stanford University, Stanford, CA 94305, USA; wenderp@stanford.edu

Chi-Wai Wong, Guangzhou Institute of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou, People’s Republic of China

Junhai Xiao, Beijing Institute of Pharmacology and Toxicology, Haidian District, Beijing, 100850, People’s Republic of China

Hang Yin, Department of Chemistry, Yale University, New Haven, CT 06520, USA

Houping Yin, Department of Chemistry, Drexel University, Philadelphia, PA 19104, USA

Dayong Zhai, The Burnham Institute for Medical Research, La Jolla, CA 92037, USA
Jiangle Zhang, Department of Medicinal Chemistry, School of Pharmacy, Zhengzhou University, Zhengzhou 450052, People’s Republic of China

Lihe Zhang, State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University, Beijing 100083, People’s Republic of China; zdszlh@bjmu.edu.cn

Laura Zheng, Department of Chemistry, Drexel University, Philadelphia, PA 19104, USA

Weiliang Zhu, Drug Discovery & Design Centre, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, People’s Republic of China
PART I

COMPUTATIONAL AND STRUCTURAL APPROACHES IN DRUG DISCOVERY
1.1 INTRODUCTION

The discovery of new drugs is a complex process. It generally starts with the identification of compounds that bind to a target or show efficacy in a simple screen. Molecules that show good affinity are called “hits.” The next step is to find compounds that have attractive pharmaceutical properties—for example, low toxicity and sufficient aqueous solubility to be orally active. Such compounds are often called “leads.” Traditionally, “hits” have been found by screening, while “leads” are developed from “hits” through chemical synthesis. Screening normally involves large numbers of compounds from natural products, corporate databases, or organic chemistry companies that can be examined for biological activity in high-throughput assays. Commercial systems can process millions of tests per day for enzyme targets. The best compounds are moved forward in a process aimed at modifying their chemical structure to improve potency, specificity, and in vivo activity while lowering toxicity and side effects. Synthetic methods include combinatorial chemistry and library synthesis (Fig. 1.1).

*Corresponding author: kuntz@cgl.ucsf.edu

*Drug Discovery Research: New Frontiers in the Post-Genomic Era, Edited by Ziwei Huang
Copyright © 2007 John Wiley & Sons, Inc.*
Computational methods have proved useful in many aspects of the discovery process (Alvarez, 2004). A variety of strategies are available. If an active “lead” is known, it is straightforward to query a database for molecules with similar properties using pharmacophore searches or quantitative structure–activity relationships (QSAR). These methods typically use information available for the ligand, inhibitor, or substrate rather than the receptor. However, there is growing interest in mapping out receptor properties—either from known family relationships with other members of the receptor family or through pharmacophore strategies applied directly to the receptor structures (Arnold et al., 2004; Hajduk et al., 2005). This chapter focuses on a set of strategies in which direct knowledge of the receptor structure is used to identify or design ligands that possess good steric and chemical complementarity to specific sites on the target macromolecule. This process is referred to as “structure-based drug design” (Brooijmans and Kuntz, 2003).

The structure-based drug design paradigm is analogous to experimental screening. Structures for the receptor or target are obtained either from the literature or from in-house operations. These structures come from crystallography or nuclear magnetic
resonance (NMR) experiments, but there is increasing interest in high-quality homology-modeled structures (Chance et al., 2004). Computer analogs of ligands are generated. These families are often called “virtual libraries” and may consist of compounds from corporate, academic, or commercial holdings (Laird and Blake, 2004; Webb, 2005). A virtual library might also include molecules that are not physically available but might be obtained through chemical synthesis, perhaps using combinatorial chemistry (Jorgensen, 2004; Kick et al., 1997). Screening of the virtual library against the target structure involves some form of positioning the putative ligand in three-dimensional space and evaluating the intermolecular interactions for that particular geometry. Typically, the process is an iterative one: The ligand is moved, and the new geometry is evaluated. This cycle is repeated until some “best scoring” geometry is identified for the particular ligand under test. Then, the next ligand in the list is chosen, and the whole procedure begins again. The goal of virtual screening is to identify the best binding candidates from the library for experimental testing. Because the virtual libraries can be huge—upwards of a billion compounds—this triage procedure is a critical step.

Once a binding candidate has been found, structure-based design can be used to optimize binding affinity. For this operation, one starts with a “hit” or “lead” with a known activity. Often, a structure for the ligand–target complex is available. There are many computational methods available for evaluating chemical variants of the “lead” that offer suggestions about the direction for the next round of synthesis. Ideally, a number of such variants are prepared, and their properties and structures are obtained so that a selection of molecules are available to take into further biological testing. Optimization methods are typically much more computationally intensive than other virtual screening approaches.

There has recently been a merging of these two ideas. Virtual libraries containing 1000–10,000 molecular fragments (sometimes called “anchors” or “scaffolds”) are used in the initial screening. The most promising are then expanded using computer synthesis in a combinatorial fashion (Jorgensen, 2004; Kick et al., 1997; Miranker and Karplus, 1991).

In this chapter, we will focus on a particular subset of molecular design strategies called “docking,” in which candidate molecules are matched to receptor structures and evaluated for chemical and geometric complementarity. We will not discuss the broad field of quantitative structure–activity relationships (QSAR) that focus on the chemical structures of ligands alone (Bender and Glen, 2005).

1.2 MOLECULAR DOCKING

1.2.1 Overview

The basis of molecular docking is the calculation or estimation of the free energy of binding of a ligand to a specific receptor site in a fixed environment. The free energy of binding yields, directly, an equilibrium binding constant and, indirectly, the preferred binding mode of the ligand–receptor complex. There are important scientific and mathematical issues involved. For example, it is currently much easier to calculate the energy/enthalpy of interaction than to obtain the free energy because we lack efficient
ways to obtain the entropic contributions. Second, the interactions of the ligand and receptor with the solvent (water, salts, and other components) are not easy to quantitate. Third, searching through the large number of conformations of the receptor and the ligand and their relative positions are difficult computer science problems. The need to repeat these calculations for large numbers of putative ligands and many possible targets requires serious attention to the algorithms. Docking protocols have adopted a variety of heuristics to make useful calculations with the knowledge that a complete high-level calculation is not feasible for the systems of biological or therapeutic interest.

What can we expect from current approaches? The best-case calculations are accurate to within approximately 0.5 kcal/mol of experimental results, but these are generally free energy differences obtained using perturbation techniques on a related family of ligands (Jorgensen, 2004). Routine results are rarely within 1 kcal/mol of experimental results, and library searches of diverse chemical types have larger inaccuracies. Work continues on improving the force fields that model the enthalpic terms (Bernacki et al., 2005). Estimates of entropic contributions are empirical, and adequate sampling of configurations and conformations search is a complex combinatorial problem. Consequently, searching large databases for new leads requires protocols that deal with four specific tasks:

1. Receptor site identification
2. Receptor site characterization
3. Orientation of the ligand within the site
4. Evaluation of the ligand

These steps are described in turn in the succeeding sections, and examples are given.

1.2.2 Receptor Site Identification

With the sequencing of the human genome and recent advances in structural techniques, the number of publicly available biomolecular structures has exploded over the last few years with over 40,000 in the Protein Data Bank (Berman et al., 2000). The two main experimental sources for three-dimensional (3D) structures of biomolecules are X-ray crystallography and high-resolution NMR spectroscopy. X-ray crystallography provides structures of biomolecules in the crystalline states, and with the crystallization of the ribosome, the upper limit of the experiment has been extended to the 1000-kDa range. NMR spectroscopy is limited to approximately 50 kDa, but the method has the advantage of providing additional information about the dynamics of the structure. As an alternative, if the structure has not been solved experimentally, computational techniques such as homology modeling can be used to predict 3D structures. We next discuss some pros and cons of each of these sources for structure-based drug design.

For X-ray crystal structures to be sufficiently accurate for drug designing purposes, a resolution of approximately 2 Å, an $R$-factor below 20%, and an $R_{\text{free}}$ factor below 30% are preferred. It is important to note that the majority of 3D crystal
structures of biomolecules do not have hydrogens or highly flexible residues included in the file. The missing atoms must be considered before structure-based drug design can begin. In addition, crystal packing forces may locally influence protein conformation, particularly for nucleic acids and surface active sites.

The result of structure determination with NMR is an ensemble of structures that agree equally well with experimental data. Although an averaged structure can be derived, it has been shown that the entire ensemble provides a more complete description of the system from an experimental perspective (Staunton et al., 2003). For structure-based drug design purposes, though, there are several methods for choosing the appropriate structure, including selecting the member of the ensemble closest to the average as measured by some distance metric or cross-docking to all members of the ensemble. Unfortunately, there is no generally accepted standard of accuracy for NMR-generated structures. As a rough rule of thumb, a high-resolution NMR structure should preferably have approximately 20 (distance or dihedral) restraints per residue (Berman and Westbrook, 2004).

If no experimental structural information is available for the target biomolecule, homology modeling can provide structures to guide the search for novel lead compounds. It should be noted that, depending on the method, homology modeling yields average errors of 3-Å root-mean-square deviation of proteins with greater than 50% sequence similarity, with larger errors for increasing sequence dissimilarity (Nissen et al., 2000). Nevertheless, homology modeling has proven to be successful in several cases, including discoveries of a highly potent DNA methylation inhibitor and a compound that discriminates between two voltage-gated K⁺ channels with 20-fold accuracy (Staunton et al., 2003).

A number of more general issues associated with the selection and preparation of a receptor structure should be noted. In many structures, ions are required for structural or functional purposes. However, modeling this type of chemistry is often difficult because the formal charge and associated desolvation energy of ions are extremely complicated to compute accurately. The protonation states of residues such as histidine, lysine, glutamic acid, or aspartic acid are highly dependent on the environment in the active site and may even change in response to the ligand (Hensen et al., 2004). In particular for crystal structures, critical water molecules may be present in the active site, and it is often difficult to predict whether the water can be replaced or should be included in the model of the receptor. All these issues can affect the quality of the model and should be carefully considered.

1.2.3 Receptor Site Characterization

Once an accurate structure has been determined for the target, ligands are typically restricted to lie within one geometric region of the macromolecule, generally known as the “binding site.” This region is generally selected because, upon ligand binding, normal function is altered. A given receptor can have one or more binding sites, such as the active sites of enzymes, allosteric sites, the binding or recognition sites of receptors, or even a dimer interface. The exact location of the binding site may be well known through experiment. However, if the binding site is not known, automated methods exist to identify potentially interesting regions.
Experimental data that indicate the binding site is the best source, if available. Experimentally derived structures of the biomolecule complexed with the natural substrate or a known inhibitor directly indicate the binding site. The Protein Data Bank and the Nucleic Acid Data Bank contain a large number of these types of structures (Berman et al., 1992, 2000). The Cambridge Crystallographic Data Centre/Astex has compiled a subset of biologically relevant protein–ligand complexes identified as being reliable for structure-based drug design purposes (Nissink et al., 2002). If direct observations are not available, binding regions can be identified through mutational experiments (such as alanine screening) or other biochemical assays.

There are cases where binding site information is not available. In these instances, computational tools can be used to indicate probable binding areas. We will describe two methods: The first is based upon geometric features, and the second uses chemical functionality of the receptor surface. For illustration, we consider SPHGEN from the DOCK suite of programs and the Multiple Copy Simultaneous Search (MCSS) approach (Ewing et al., 2001; Kuntz et al., 1982; Meng et al., 1992; Miranker and Karplus, 1991). In addition, an interesting statistical characterization has been recently described (Hajduk et al., 2005).

SPHGEN automatically identifies a target site by computing a set of site points or sphere centers, which serves to create a negative image of the surface. The algorithm begins by mapping the geometric features of the receptor surface, as defined by Lee and Richards, using the dms program (Richards, 1977). Then, spheres of varying radii are analytically generated to touch the molecular surface at two points, with the sphere center lying along the surface normal and with no portion of the sphere intersecting with a receptor atom. These overlapped spheres indicate various surface features, including invaginations and clefts. A clustering protocol, using radial overlap as a metric, is then used to indicate potential areas for ligand binding. The largest cluster is generally used as the binding site, and, once generated, the cluster is used as a template for possible ligand atom positions.

MCSS identifies binding sites by mapping the chemical properties of the biomolecular surface. Thousands of copies of specific molecular fragments are distributed in the target region of the protein. Then, energy minimization is performed on the ensemble, creating distinct local minima for each fragment. This process is repeated on a variety of chemical functionalities until the surface is adequately described. Although this method does not capture every geometric detail of the binding site, it does provide a basic pharmacophore that can be used in later studies (Arnold et al., 2004; Miranker and Karplus, 1991).

1.2.4 Orientation of the Ligand in the Target Site

There are two basic strategies for exploring the orientational degrees of freedom for putative ligands. The first uses a search grid for both the translational space and the euler angle space. This brute-force method is feasible if one is studying a few ligands in a restricted site. It can be extended to larger libraries if parallel processing is available (Jorgensen, 2004). Alternatively, protocols have been developed to prescreen orientation space. For example, the DOCK program uses a geometric pairing...
algorithm, matching the sphere centers (described above) with ligand centers (usually ligand atoms). The matching criterion is based on a comparison of intersphere and interatom distances. Exhaustive or selective searches can be done over the match matrix. Careful placement of the spheres is an important step in getting good-quality results. The second type of method selectively samples all of orientation and conformation space using search engines. For example, the Metropolis algorithm and simulated annealing are used in QXP, and the genetic algorithm has been implemented an Autodock (McMartin and Bohacek, 1997; Morris et al., 1998). These algorithms have been studied extensively for many applications. Their strengths and limitations are well understood.

1.2.5 Evaluation of Ligand Orientations

The many configurations (orientations and conformations) of the ligand need to be evaluated with a scoring function to identify the energetically most favorable ligand binding pose. Ideally, the scoring function would calculate the ligand free energy of binding in aqueous solution (Beveridge and Dicapua, 1989; Kollman, 1993). However, the large computational expense of these calculations leads to the introduction of scoring functions that calculate a range of simplifications of the ligand binding free energy.

Scoring functions can be broadly classified into two categories: those based on first-principles-derived molecular mechanics force fields, and those based on functions fit to empirically derived binding data. For the purposes of this review, scoring functions that employ quantum mechanics are not considered as the extreme computational cost of these calculations make them prohibitive for use during small-molecule docking.

Of the first-principles-derived scoring functions, the most computationally efficient are those that approximate the binding free energy ($\Delta G_{\text{bind}}$) as the molecular mechanics protein–ligand interaction energy. Molecular mechanics treats the molecule as a collection of atoms governed by a set of classical mechanical potential functions (Weiner and Kollman, 1981). Parameters for these potentials are derived from small-molecule experiments and refined to yield correct structural and thermodynamic quantities such as bond stretching frequencies or heats of formation. The primary DOCK energy scoring function approximates the ligand–receptor binding energy using the AMBER molecular mechanics intermolecular interaction energy, a sum of the Lennard-Jones 6–12 van der Waals (vdW) potential, and the Coulombic potential, given in equation (1.1):

$$E = \sum_{i}^{\text{Lig}} \sum_{j}^{\text{Rec}} \left( \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} + K \frac{q_i q_j}{D r_{ij}} \right)$$  \hspace{1cm} (1.1)

where $i$ indexes the ligand atoms and $j$ indexes the receptor atoms; $A$ and $B$ are the vdW attractive and dispersive parameters, respectively; $q$ is the partial charge on the atom; $K$ is the scaling constant that converts electrostatic energy into kcal/mol; $D$ is the dielectric constant of the medium; and $r_{ij}$ is the distance between ligand atom $i$ and receptor atom $j$ (Pearlman et al., 1995). This scoring function is limited by its use of a distance-dependent dielectric screening function to mediate all charge–charge interactions.
interactions. This dielectric treatment assumes that the dielectric value of the solvent is uniform between all charge pairs.

A class of scoring functions has been developed that utilizes implicit models of solvation to calculate the electrostatic component of the molecular mechanics intermolecular interaction energy in a more sophisticated fashion than the simple Coulombic approach previously described. Both the Generalized Born (GB) and Poisson–Boltzmann (PB) terms have been combined with an empirically derived surface area (SA) term to include the energy of desolvating nonpolar atoms, and the resulting GB/SA and PB/SA methods have been implemented into molecular dynamics and docking methods (Feig and Brooks, 2004; Feig et al., 2004; Honig and Nicholls, 1995). While these functions are more computationally intensive than the Coulombic electrostatic energy functions, their proper treatment of solvation effects yields more accurate energy scores, and they are therefore frequently used in a hierarchical fashion to rescore docked ligand poses. DOCK 5 implements a GB/SA scoring function that is recommended for use in a rescoring capacity (Zou et al., 1999).

The most computationally intensive class of first-principles-derived scoring functions combine molecular dynamics (MD) simulations with implicit or explicit solvation to average the interaction energies from a Boltzmann-weighted ensemble of complex structures, yielding accurate estimates of the binding free energy that takes into account protein flexibility. The implicit solvation methods, MM-PB/SA and MM-GB/SA, perform a short, explicit solvent MD simulation from which a set of snapshots of the protein–ligand complex structure are saved (Gohlke and Case, 2004; Wang et al., 2001; Zhou and Madura, 2004). These snapshots, representing a Boltzmann-weighted ensemble of complex structures, are rescored with either the PB/SA or GB/SA scoring functions, and the average interaction score of the snapshots is taken as the free energy of binding for the ligand. Several research groups in academia and industry use these methods in a hierarchical fashion to rescore ligands identified as potential binders (Kollman et al., 2000).

The second major class of scoring functions models the binding free energy as a weighted sum of several different types of interaction energies, with and without explicit vDW and electrostatic terms. Many of these functions are based upon a comparison of receptor–ligand complexes and experimental binding data \( K_i \). Programs such as AutoDock, FlexX, GOLD, and Glide implement a variety of empirically derived energy score functions, including the well-known ChemScore and PLP functions (Eldridge et al., 1997; Friesner et al., 2004; Gehlhaar et al., 1995; Halgren et al., 2004; Hoffmann et al., 1999; Morris et al., 1998; Verdonk et al., 2003).

### 1.3 LIGAND STRUCTURE GENERATION

To identify virtual screening “hits,” databases consisting of three-dimensional structures of putative ligands are searched using the methodology described above. How are these structures obtained? In some cases, structures can be taken directly from a database of experimentally determined structures; for example, the Cambridge Crystallographic Database contains over 360,000 small-molecule...
crystallographically determined structures. In cases where only the two-dimensional information is known, one can create computer representations of the covalent structure using conformation generation programs such as Concord (Tripos), ChemX (Accelrys), Rubicon (Daylight), and Omega (OpenEye). These programs provide one or more conformers consistent with the chemical connectivity and general rules of physical organic chemistry. Often, these conformers will be ranked by some energy formula. Finally, some vendors make libraries of three-dimensional structures directly available, including the Advanced Chemical Development (www.acdlabs.com), MDL Drug Data Report (www.mdl.com), National Cancer Institute Open Database Compounds (cactus.nci.nih.gov), Tripos Discovery Research Screening Libraries (www.tripos.com), InfoChem GmbH database (www.infochem.de), Thomson Index Chemicus database (scientific.thomson.com), and ZINC (blaster.docking.org/zinc).

Two strategies have emerged to study flexible ligands. The first, generally referred to as incremental construction, breaks down the ligand into smaller pieces and then rebuilds it during the docking calculation. One example of this technique starts with a fragment of the compound (an “anchor”) and then adds atoms in layers during a docking or an optimization cycle. This approach has been called “anchor and grow.” In the alternate method, conformers for each compound can be pregenerated, stored in a database, and then rigidly docked. Molecular dynamics and Monte Carlo techniques offer a combination strategy where the starting point is a single conformation of the ligand that then explores alternatives during the dynamics phase. There have been some tests of the two strategies, but there is no strong consensus of which is better (Lorber and Shoichet, 2005; Moustakas et al., 2006).

Other important issues that influence ligand structure and docking are choices of partial charges, tautomer preferences, and pK_a values. A new database (ZINC) that deals with many of these concerns is now available through the Shoichet group at UCSF (Irwin and Shoichet, 2005).

1.4 DESCRIPTION OF DOCKING PROGRAMS

While we do not have the space to provide descriptions of the many different approaches to molecular docking, Table 1.1 summarizes features of several of the most frequently used programs. We also present a brief synopsis of DOCK, developed at UCSF.

The current version of the DOCK program, is written in C++ and provides an object-oriented implementation in which each major component of the DOCK algorithm is a class with a documented interface, allowing these DOCK functions to be modified or replaced easily. As a result, it has been possible to independently validate and optimize the rigid body sampling, the flexible sampling, the energy scoring functions, and our minimizers. DOCK features an energy scoring function based on a molecular mechanics force field, solvation corrections using implicit solvent models, integration with the complete AMBER force field score, rigid body docking, ligand conformational searching, binding pose cluster analysis, and local minimization methods and also includes support for parallel computing using the MPI standard.
1.5 TESTS OF DOCKING AND STRUCTURE-BASED DESIGN

Despite well-known methodological weaknesses, structure-based screening using molecular docking has had important successes. Pragmatically, docking has suggested new, nonobvious ligands for multiple targets; these have been subsequently tested and shown to bind experimentally. Hugo Kubinyi, in a recent review, describes over 50 macromolecular targets for which ligands have been discovered using docking-based approaches (see Table 1.2 for a partial list) (Kubinyi, 2006). Most of these projects used experimental X-ray structures to represent the protein. In several cases, homology-modeled structures were employed (Evers and Klebe, 2004; Schapira et al., 2003).
<table>
<thead>
<tr>
<th>Target</th>
<th>Representative Hit</th>
<th>Lead Inhibitor IC&lt;sub&gt;50&lt;/sub&gt;</th>
<th>Follow-Up Inhibitor IC&lt;sub&gt;50&lt;/sub&gt;</th>
<th>Complex Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>p56 Lck SH2 domain (Huang et al., 2004)</td>
<td><img src="image" alt="p56 Lck SH2" /></td>
<td>10 μM</td>
<td>NR</td>
<td>No</td>
</tr>
<tr>
<td>Neurokinin-1 receptor (Evers and Klebe, 2004)</td>
<td><img src="image" alt="Neurokinin-1" /></td>
<td>0.25 μM</td>
<td>NR</td>
<td>No</td>
</tr>
<tr>
<td>AICAR transformylase (Li et al., 2004)</td>
<td><img src="image" alt="AICAR transformylase" /></td>
<td>0.15 μM</td>
<td>NR</td>
<td>No</td>
</tr>
<tr>
<td>IMPDH (Pickett et al., 2003)</td>
<td><img src="image" alt="IMPDH" /></td>
<td>31 μM</td>
<td>NR</td>
<td>No</td>
</tr>
<tr>
<td>Checkpoint kinase 1 (Lyne et al., 2004)</td>
<td><img src="image" alt="Checkpoint kinase 1" /></td>
<td>0.11 μM</td>
<td>NR</td>
<td>No</td>
</tr>
<tr>
<td>DNA gyrase (Boehm et al., 2001)</td>
<td><img src="image" alt="DNA gyrase" /></td>
<td>10,000 μM</td>
<td>0 μM</td>
<td>Yes</td>
</tr>
<tr>
<td>Aldose reductase (Iwata et al., 2001)</td>
<td><img src="image" alt="Aldose reductase" /></td>
<td>4.3 μM</td>
<td>0 μM</td>
<td>No</td>
</tr>
<tr>
<td>CDK4 (Honma et al., 2001)</td>
<td><img src="image" alt="CDK4" /></td>
<td>44 μM</td>
<td>0 μM</td>
<td>Yes</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Target</th>
<th>Representative Hit</th>
<th>Lead Inhibitor IC$_{50}$</th>
<th>Follow-Up Inhibitor IC$_{50}$</th>
<th>Complex Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matriptase (Enyedy et al., 2001)</td>
<td><img src="image" alt="Matriptase structure" /></td>
<td>0.92 μM</td>
<td>0 μM</td>
<td>No</td>
</tr>
<tr>
<td>Bcl-2 (Enyedy et al., 2001)</td>
<td><img src="image" alt="Bcl-2 structure" /></td>
<td>10.4 μM</td>
<td>NR</td>
<td>No</td>
</tr>
<tr>
<td>Adenovirus protease (Abarbanel 1984; Pang et al., 2001b)</td>
<td><img src="image" alt="Adenovirus protease structure" /></td>
<td>3.1 μM</td>
<td>NR</td>
<td>No</td>
</tr>
<tr>
<td>AmpC β-lactamase (Powers et al., 2002)</td>
<td><img src="image" alt="AmpC β-lactamase structure" /></td>
<td>26 μM</td>
<td>1 μM</td>
<td>Yes</td>
</tr>
<tr>
<td>Retinoic acid receptor (Schapira et al., 2001)</td>
<td><img src="image" alt="Retinoic acid receptor structure" /></td>
<td>2 μM</td>
<td>NR</td>
<td>No</td>
</tr>
<tr>
<td>TGT (Gradler et al., 2001)</td>
<td><img src="image" alt="TGT structure" /></td>
<td>8.3 μM</td>
<td>0 μM</td>
<td>Yes</td>
</tr>
<tr>
<td>Carbonic anhydrase II (Gruneberg et al., 2001)</td>
<td><img src="image" alt="Carbonic anhydrase II structure" /></td>
<td>0.0008 μM</td>
<td>NR</td>
<td>Yes</td>
</tr>
<tr>
<td>HPRTase (Freymann et al., 2000)</td>
<td><img src="image" alt="HPRTase structure" /></td>
<td>2.2 μM</td>
<td>NR</td>
<td>No</td>
</tr>
<tr>
<td>Dihydro-dipicolinate (Paiva et al., 2001)</td>
<td><img src="image" alt="Dihydro-dipicolinate structure" /></td>
<td>7.2 μM</td>
<td>NR</td>
<td>No</td>
</tr>
</tbody>
</table>
In recent work, the structures of known ligands in complex with their receptors have been predicted by docking, beginning with the structures of the independent molecules (Fig. 1.2) (Rizzo et al., 2000; Rosenfeld et al., 2003). In these studies, where the binding affinity is known but the structure of the complex is not, the docking predictions have been relatively accurate. The caveat to this is that there are many cases where docking mispredicts geometries in retrospective tests. Still, in published cases where the goal was genuine prediction, the docked geometry has often turned out to correspond closely to the subsequent experimental result.

A more difficult test is comparing the predicted geometries of novel ligands that emerge from the docking screens themselves. There are many examples of such predictions of ligand and geometry from docking screens against simple model cavity sites. These sites are small, completely enclosed by the protein, and dominated by one particular type of interaction, such as hydrophobicity, a single hydrogen bond acceptor, or a single electrostatic interaction. These features have allowed for multiple predictions of new ligands that are tested experimentally, often including structure

**Figure 1.2** Predicted complexes versus X-ray crystallographic structures that were subsequently determined. (A) Predicted (carbons in gray) and experimental (green) structures for sustiva in HIV reverse transcriptase (Rizzo et al., 2000; Rosenfeld et al., 2003). (B) Predicted (magenta) and experimental (white carbon atoms) structures of 2,3,4-trimethylthizole in the W191G cavity of cytochrome c peroxidase (Rizzo et al., 2000; Rosenfeld et al., 2003). (C) Predicted (green) and experimental structure (carbons in gray) of an HIV protease inhibitor (ligands with thick bonds, enzyme residues with thin bonds) (Brik et al., 2003, 2005; courtesy of Art Olson, TSRI). See color plates.
determination (Fig. 1.3) (Brenk et al., 2006; Graves et al., 2005; Wei et al., 2002, 2004). Thus, X-ray crystal structures have been determined for about 25 ligands bound to three different cavities; in every case, the docking prediction corresponds closely to the X-ray crystallographic result. These results suggest that current docking algorithms are adequate to capture first-order determinants of binding fidelity (Fig. 1.4) (Gradler et al., 2001; Gruneberg et al., 2002; Powers et al., 2002; Wei et al., 2002).

How does performance in simple sites translate into larger, more drug-like sites? The consensus of many retrospective and prospective docking screens is that the ability to predict ligands and their geometries diminishes considerably in biology-relevant targets. In most cases, this failure reflects the increased complexity of the binding sites and the greater opportunities to find decoy ligand geometries. Nevertheless, there are examples of successful ligand prediction followed by structural

Figure 1.3  Docking predicted ligands from virtual screening against simple cavity sites. (A) The docked prediction (carbons in green) superposed on the crystallographic result (carbons in cyan). The surface of the L99A/M102Q cavity of T4 lysozyme (yellow) is cut away to reveal the complex. (i) Phenol, (ii) Chlorophenol, (iii) fluoroaniline, (iv) methylpyrrole, (v) difluorophenol. (B) The docked prediction (carbons in green) superposed on the crystallographic result (carbons in yellow) in the W191G cavity of cytochrome c peroxidase. (i) Thiophene-amidine, (ii) diaminopyridine, (iii) 2-amino-5-methylpyridine, (iv) 2-amino-4-methylpyridine, (v) diaminopyrimidine, (vi) hydroxymethylimidazole, (vii) 3-methyl-4-n-methylpyridine, (viii) 4-hydroxymethyl-pyridine, (ix) aminomethyl-cyclopentane, (x) aminomethyl-benzene, (xi) aminomethyl-furan. See color plates.
determination, and in these cases the docking prediction is often close to the experimental result (Fig. 1.4) (Gradler et al., 2001; Gruneberg et al., 2002; Powers et al., 2002; Wei et al., 2002). These studies suggest that when the method does correctly predict a new ligand, even for a complicated, drug-like binding site, it does so for the right reasons.

An important question is whether structure-based screening is worth the effort, assuming that groups have access to high-throughput screening for ligand discovery. The two types of screens have been compared only a few times publicly, though rather more in unpublished industrial work. In the few published studies, the virtual screens had “hit rates” 10- to 1700-fold higher than the empirical screens (Table 1.3) (Doman et al., 2002; Kick et al., 1997; Oshiro et al., 2004; Paiva et al., 2001; Wyss et al., 2003). In the case with the best hit-rate enhancement, that of the diabetes-associated enzyme PTP-1B, the comparison was an imperfect one. Here, different libraries were targeted by the virtual and high-throughput screen, and a slightly different assay was used. In very recent work, Eric Brown and colleagues at McMaster University challenged the virtual screening community to predict the affinities of 50,000 molecules, none of which had been tested before but which were

**TABLE 1.3 Hit Rates and Drug-Like Properties for Inhibitors Discovered with High-Throughput and Virtual Screening Against the Enzyme PTP-1B (Doman et al., 2002)**

<table>
<thead>
<tr>
<th>Technique</th>
<th>Number Compounds Tested</th>
<th>Hits with IC$_{50}$ &lt; 100 µM</th>
<th>Hits with IC$_{50}$ &lt; 10 µM</th>
<th>Rule of Five Compliant Hits$^a$</th>
<th>Hit Rate$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>HTS</td>
<td>400,000</td>
<td>85</td>
<td>6</td>
<td>23</td>
<td>0.021%</td>
</tr>
<tr>
<td>Docking</td>
<td>365</td>
<td>127</td>
<td>18</td>
<td>73</td>
<td>34.8%</td>
</tr>
</tbody>
</table>

$^a$Number of 100 µM or better inhibitors that passed all four “rule of five” criteria (Paiva et al., 2001).

$^b$The number of compounds experimentally tested divided by the number with IC$_{50}$ values of 100 µM or less.
about to be tested in a high-throughput screen against dihydrofolate reductase (DHFR). In this experiment, the docking and HTS libraries were precisely the same, as were the experimental conditions. One of the startling results of this experiment was the very small number of hits to emerge from the screen. Indeed, whereas several groups were able to enrich putative inhibitors among their high-scoring molecules, the experimental group eventually concluded that they had no reliable hits at all (Elowe et al., 2005; Lang, et al., 2005). Intriguingly, several of the computational groups were able to indicate the lack of binders as part of their predictions (Brenk et al., 2005). Whereas the lack of experimental hits prevents definitive conclusions from this study, what does seem clear is that there is room for more of these comparative studies and "competitions."

1.6 CONCLUSIONS AND FUTURE DIRECTIONS

We have described the general steps involved in docking as well as typical algorithms for addressing each stage of the methodology. We have also listed examples in which prepackaged programs have been successfully used to discover novel inhibitors of a wide range of medically relevant applications. In the future, these methods will become progressively more integrated in the drug design process.

However, there are still several components of docking that need improvement. The two most-debated open questions in the field involve improving the scoring functions and developing algorithms for receptor flexibility. For scoring functions, research focuses on improving the treatment of solvent and the effects of entropy loss upon binding. Most docking approaches currently include drastic approximations of both of these properties, which have shown improvements over older methods. However, it is necessary to develop new schemes that treat these issues more accurately while preserving the speed of the calculation. Configurational entropy contributions are also difficult to calculate. Techniques that use molecular dynamics simulations to generate ensembles of ligand positions that are then rescored with high-accuracy scoring functions generate very accurate free energies of binding; however, they are computationally expensive (Kollman et al., 2000). It will be necessary to develop sampling techniques able to generate Boltzmann-weighted ensembles of ligand poses without requiring the use of expensive molecular dynamics calculations. Similarly, for receptor flexibility, many prepackaged programs have recently begun to develop new algorithms that allow for some measure of induced fit in the binding site. Several of these methods show great promise and will be further perfected with time, allowing for more and more structural rearrangement.

ACKNOWLEDGMENTS

We thank the sponsors of this work: the National Institute of General Medical Sciences of the National Institutes of Health through grants GM 56531 (P. O. Montellano, PI to IDK) and GM 59957 (BKS) as well as the AFPE fellowship to PTL.
REFERENCES


2.1 INTRODUCTION

The identification of a proper lead compound for a given molecular target is a critical step in the process of drug discovery. Traditionally, high-throughput screening (HTS) of large chemical libraries has been a primary source of identification of novel lead compounds. In recent years, the rapid progress in the human genome project has provided an ever-increasing number of potential drug targets to be screened (Westbrook et al., 2003). Moreover, advances in high-throughput (HTP) strategies for structure determination (Blundell and Abell, 2002) have yielded a large number of three-dimensional (3D) structures of therapeutically relevant targets. To this end, computational tools are becoming increasingly important to efficiently integrate
this rapidly increasing wealth of structural data with the more traditional lead optimization techniques. Amongst them, virtual screening is bringing a more focused approach to drug discovery by using high-performance computing to analyze chemical databases and prioritize compounds for synthesis and analysis.

2.2 VIRTUAL SCREENING OF CHEMICAL LIBRARIES

Virtual screening is a computational filter to reduce the size of a chemical library to be screened experimentally and offers an opportunity to drastically reduce the time and effort associated with lead identification. The benefits are a focused subset with enhanced hit rates and a prioritized library for screening and synthesis. There are two fundamental approaches for virtual screening: a ligand-based approach (Bajorath, 2002) and a receptor-based approach (Gohlke et al., 2000). The ligand-based approach aims to identify molecules with physical and chemical similarities (pharmacophore-based, descriptor-based) to known ligands that are likely to interact with the target. This type of approach limits the diversity of the hits as they are biased by the properties of known ligands. Receptor-based virtual screening (protein–ligand docking, active site-directed pharmacophores) uses knowledge of the target protein’s 3D structure to impose a structure-based filter on a chemical database to select candidate compounds that are likely to interact favorably with the protein’s active site residues. This is a more open-ended approach that allows the identification of structurally novel ligands that may have similar interactions like known ligands or may have different interactions with other parts of the binding site. Receptor-based virtual screening (RBVS) and some of its recent applications will be discussed here.

2.2.1 Receptor-Based Virtual Screening

Molecular recognition (Fische, 1894) is the fundamental basis for drug action in which drug molecules exhibit pharmacological activity by binding to a target protein and forming a stable protein–ligand complex. The major factors contributing to the stability of such complexes include shape complementarity of the ligand to the ligand-binding site, formation of key hydrogen bonds, and good hydrophobic and electrostatic interactions with the active site residues as well as stable conformation of the bound ligand when the complex is formed. Receptor-based virtual screening (RBVS) aims to exploit the molecular recognition between a ligand and a target protein to select chemical entities that bind strongly to the active sites of biologically relevant targets for which the three-dimensional structures are known or inferred. This approach uses docking and scoring (Muegge and Enyedy, 2004) to sort the candidates in a virtual library. The docking algorithms (Halperin et al., 2002; Brooijmans and Kuntz, 2003) deal with the prediction of ligand conformation and orientation (or pose) within the targeted active site of the receptor. The scoring methods (Wang et al., 2003) evaluate the binding interactions between the target and the small molecule and aim to predict the biological activity of the compound based on the computed binding interactions.
In RBVS, one starts with a 3D structure of a target protein and a 3D database of ligands and uses virtual filtering to dock and score compounds as a means to identify potential lead candidates for further analysis and improvement. Some of the most widely used software programs in high-throughput RBVS are DOCK (Ewing et al., 2001), FlexX (Rarey et al., 1996), GOLD (Jones et al., 1997), Glide (Halgren et al., 2004), ICM (Totrov and Abagyan, 1997), FRED (McGann et al., 2003), and AutoDock (Goodsell et al., 1993). All such software programs have docking and scoring algorithms to generate subsets of a compound collection with higher affinity against a target by predicting their binding mode (by docking) and affinity (by scoring) and retrieving those with the highest scores. These programs can dock single or multiple conformers of small-molecule libraries. Although docking algorithms have matured significantly over the years, some challenges still remain, such as the treatment of protein-flexibility issues arising from induced fit or other conformational changes of the protein that occur upon ligand binding and the participation of water molecules in protein–ligand interactions. In some cases, limited resolution of available crystal structures of target proteins also affect the outcome of the virtual screening effort.

Receptors are usually represented in three forms in the different docking algorithms. They are treated as either atoms, surfaces, or potential energy grids (Halperin et al., 2002). Ligand flexibility is also treated in three different ways (Brooijmans and Kuntz, 2003) in these algorithms. The three ligand flexibility treatment methods are: (a) systemic search (Leach, 1996), using conformational search or incremental construction; (b) random or stochastic search (Kitchen et al., 2004), using Monte Carlo (Hart and Read, 1992) genetic algorithm (Oshiro et al., 1995), or tabu search (Baxter et al., 1998); and (c) simulation methods using molecular dynamics (Di Nola et al., 1994). The treatment of protein flexibility (Carlson and McCammon, 2000) in docking algorithms is less advanced than that of ligand flexibility and approaches such as rotamer libraries (Leach, 1994), molecular dynamics, and Monte Carlo simulations, have, in some cases, been applied to a limited region of the protein binding site. Docking programs such as DOCK, FlexX, and Glide use the systematic approach for treating small-molecule flexibility; AutoDock, GOLD, and MOE use the random/stochastic method for treating ligands; and DOCK, Glide, MOE (2003), and AutoDock incorporate the simulation option for ligand flexibility treatment.

The evaluation and ranking of the predicted ligand poses are critical steps in the success of the receptor-based virtual screening technique. The scoring methods (Ewing et al., 2001) implemented in various docking software programs are empirically or semiempirically derived, and these functions attempt to estimate the tightness of target–ligand interactions in bound complexes. Most scoring functions implemented in docking programs make various assumptions in the evaluation of the docked complex and oversimplify certain physical characteristics (such as hydrophobic, entropic, and solvation effects) that determine molecular recognition. Three major categories of scoring functions are currently used in the major docking programs; empirical, force-field based, or knowledge-based scoring functions.
The various force-field scoring functions are based on the different molecular mechanics of the force-field parameters used in the different programs. These scoring functions typically measure only the potential energy of the system by quantifying the sum of the energies between the receptor–ligand interaction energy and the internal ligand energy induced by binding. Most force-field scoring functions consider only one protein conformation, which greatly simplifies the actual scenario. The major drawback of this scheme is that it often relies on nonbonded interaction energy terms of standard force fields formulated to model gas-phase contributions to structure and energetics and do not include solvation and entropic terms (Kitchen et al., 2004). Some examples of force-field-based scoring functions include: G-score (Kramer et al., 1999a,b), DOCK energy score (Meng et al., 1992), and GOLD score (Verdonk et al., 2003).

The empirical scoring functions are fit to reproduce experimental data such as binding energy as the sum of individual uncorrelated terms. These coefficients were determined by either regression analysis or partial least-squares fitting on a large training set of experimentally available receptor–ligand complexes (Bohm, 1994, 1998). The success of these scoring functions are measured by comparing the predicted to experimental binding affinity. The strength of these types of functions is that they are trained on a large number of receptor–ligand complexes available in the PDB. The weakness is that because they are constructed empirically, they are sensitive to the complexes represented in their training sets and therefore are unsuitable for VS projects, where these functions are likely to encounter novel ligands and novel ligand orientations. Some examples of empirical scores include: LUDI (Bohm, 1992), Chemscore (Eldridge, 1997), SCORE (Tao and Lai, 2001), and F-Score (Rarey et al., 1996).

Knowledge-based scoring functions are designed to reproduce experimental structures rather than binding energies. In this case, the structural information from a large number of complexes in the PDB was used to construct atomic interaction potentials, where a number of atom-type interactions are defined depending on their molecular environment. Similar to the empirical scoring functions, the knowledge-based functions are essentially biased toward the limited information that is implicitly encoded in their derivation. However, their computational simplicity allows them to be effectively used in virtual screening and in scoring large chemical databases. Some examples of knowledge-based scores include PMF (Muegge and Martin, 1999; Muegge, 2000, 2001) and DrugScore (Gohlke et al., 2000).

Regardless of the choice of docking software used and even when the binding conformations are predicted, the major drawback in the RBVS procedure is the unavailability of appropriate scoring functions to differentiate between correct and incorrect poses of bound ligands and to identify false-negative and false-positive hits. Given the limitations of the current scoring functions, a recent trend in this field has been the use of consensus scoring schemes (Charifson et al., 1999) and visual inspection to select likely candidates. Consensus scoring combines the information from different scores to balance errors in individual scores, reduces the number of false positives identified by individual scoring functions, and improves the odds of identifying the true ligands.
The output of a docking-based screen is a set of 3D models of the predicted binding mode of each compound against the receptor, together with a ranking that is a measure of the quality of fit. Thus, it represents the most detailed and relevant model for identifying a receptor-focused subset of ligands. A few important factors to consider when selecting docking software are the capability of the software for iterative refinement of docking parameter/protocol based on new results, adaptability to additional scoring functions, pre- and/or post-docking filters, design, components, and results of validation studies, user learning curve, customer support, cost, speed, user interface, input/output structural file formats, code availability, and upgrading possibility. Like any HTP technology, RBVS workflows are designed to balance accuracy and time commitment. A typical RBVS scheme is shown in Fig. 2.1. Over the last few years, a number of laboratories have published virtual screening studies using a variety of workflow and docking programs, and the number is steadily rising, demonstrating the level of interest in this approach and its broad applicability.

For the small-molecule compound database, it is desirable to have maximum structural diversity in the virtual library so as to maximize the chances of finding a hit for the target macromolecules. Some of the commonly used small-molecule databases in virtual screening are large public databases such as ZINC (3.3 million commercially available compounds; free), Available Chemicals Directory (ACD, 4 million entries; not free), National Cancer Institute compound database (NCI, 400,000 entries; free), and MDDR (MDL Drug Data Report, >147,000 entries).
Other possibilities for small-molecule collection include CMC (Comprehensive Medicinal Chemistry, >8600 entries), CSD (Cambridge Structural Database), Beilstein, and SciFinder. Large pharmaceutical companies have corporate databases of a few million compounds.

2.2.2 Recent Successes from RBVS

2.2.2.1 Bcl-2 Family Proteins

Proteins in the Bcl-2 family are central regulators of programmed cell death, and members that inhibit apoptosis, such as Bcl-xL and Bcl-2, are overexpressed in many cancers, making them important targets for anticancer drug development (Shangary and Johnson, 2003). The development of inhibitors of these proteins as potential anticancer therapeutics has been first explored in our laboratory by virtual screening of a chemical database using RBVS techniques. At the time the study was carried out, no experimental structure of Bcl-2 was available. Because Bcl-2 and Bcl-xL have 47.2% sequence identity, a homology model of Bcl-2 was constructed based on the NMR structure of Bcl-xL complexed with the Bak peptide (Sattler et al., 1997). The Bak BH3 peptide binds to a largely hydrophobic binding pocket formed by the BH1, BH2, and BH3 domains of in Bcl-2 as well as in Bcl-xL. This binding pocket in Bcl-2 appears to be essential for its anti-apoptotic function since mutations at this site abolish its function (Yin Oltvai et al., 1994; Cosulich et al., 1997). Using the Bak binding surface pocket in the homology model of Bcl-2 as the targeted active site, virtual screening of the Available Chemicals Directory (ACD) database of 193,833 compounds was carried out (Wang et al., 2000) using the program DOCK3.5. Fifty-three candidate molecules that displayed relatively lower binding energy, favorable shape complementarity, and/or potential of forming hydrogen bonds with the Bcl-2 protein residues were picked for the next phase of biological testing. Twenty-eight of these compounds with diverse scaffold and possible druglike properties were actually obtained for biological testing. Subsequent fluorescence polarization assays verified that the most potent compound, HA14-1, competed with Bak BH3 peptide for binding to Bcl-2 (IC50 \( \approx 9 \mu M \)). The structure of this compound is shown in Table 2.1, and a structural model for the complex of HA14-1 with the Bcl-2 surface pocket as predicted by computer docking calculations is shown in Fig. 2.2. Apoptosis induced by HA14-1 in HL-60 cells is characterized by activation of caspase protease and loss of \( \Delta \psi_m \). This was the first small-molecule ligand identified for this family of proteins, and this identification opened new avenues for the development of anticancer agents targeting members of this family of proteins.

Another small organic molecular inhibitor of Bcl-2, named compound 6 (structure shown in Table 2.1), was discovered by virtual screening in another laboratory (Enyedy et al., 2001), starting with a homology model of Bcl-2. The binding site was identified as all residues within 8 Å of the Bak BH3 peptide binding site of Bcl-2. The National Cancer Institute database of 206,876 compounds was virtually screened against the modeled structure of Bcl-2 using the program DOCK3.5. The top 500 compounds with the best scores were further filtered based on those with nonpeptidic character. Thirty-five compounds were
<table>
<thead>
<tr>
<th>Target Protein</th>
<th>Disease Treatment Area</th>
<th>Small Molecule Library</th>
<th>VS Software Used</th>
<th>Potency in Binding Assay</th>
<th>Structure of a Representative Hit Compound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bcl-2</td>
<td>Cancer</td>
<td>ACD</td>
<td>DOCK3.5</td>
<td>IC₅₀ = 9 μM</td>
<td><img src="image" alt="HA14-1" /></td>
</tr>
<tr>
<td>(Wang et al., 2000)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bcl-2</td>
<td>Cancer</td>
<td>NCI DB</td>
<td>DOCK</td>
<td>IC₅₀ = 4 μM</td>
<td><img src="image" alt="Compound-6" /></td>
</tr>
<tr>
<td>(Enyedy et al., 2001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bcl-2</td>
<td>Cancer</td>
<td>Chembridge Chemnavigator</td>
<td>TreeDOCK</td>
<td>IC₅₀ = 8.5 ± 0.6 μM</td>
<td><img src="image" alt="BH3I-ISCH₃" /></td>
</tr>
<tr>
<td>(Lugovskoy et al., 2002)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XIAP</td>
<td>Cancer</td>
<td>TCM-3D</td>
<td>DOCK</td>
<td>IC₅₀ = 4.1 ± 1.1 μM</td>
<td><img src="image" alt="Embelin" /></td>
</tr>
<tr>
<td>(Nikolovska-Coleska et al., 2004)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Acetyl Cholinesterase (Mizutani and Itai, 2004)
Alzheimer’s disease  ACD & Maybridge  ADAM & EVE  IC₅₀ = 590 nM

Acetyl Cholinesterase (Rollinger et al., 2004)
Alzheimer’s disease  NPD  LigandScout  IC₅₀ = 168.6 μM

ER (Zhao and Brinton, 2005)
Alzheimer’s disease  Plant-based library  GOLD 2.0  IC₅₀ = 680 nM

DPP-IV (Ward et al., 2005)
Diabetes  Commercially available compounds  GLIDE  % inhibition: 81.9% (30 μM)

Human aldose reductase (Kraemer et al., 2004)
Diabetes  ACD  FlexX  IC₅₀ = 2.4 ± 0.5 μM

BTB02 809

(Continued)
<table>
<thead>
<tr>
<th>Target Protein</th>
<th>Disease</th>
<th>Treatment Area</th>
<th>Small Molecule Library Used</th>
<th>Assay</th>
<th>Hit Compound</th>
<th>Structure of a Representative Hit Compound</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTP1B</td>
<td>Diabetes</td>
<td></td>
<td>ACD BioSpecs Maybridge</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADAM12</td>
<td>Cardiovascular disease</td>
<td>Virtual Library PharmoMap/PharmaScan</td>
<td>DOCK3.5</td>
<td></td>
<td>Compound-1</td>
<td><img src="image" alt="Compound-1" /></td>
</tr>
<tr>
<td>PDE-1 PDE-5</td>
<td>Cardiovascular disease</td>
<td>Virtual Library SPECS</td>
<td>SPSS ver. 10.0</td>
<td></td>
<td>Compound-5</td>
<td><img src="image" alt="Compound-5" /></td>
</tr>
<tr>
<td>FT</td>
<td>Cancer</td>
<td></td>
<td>ACD EUDOC</td>
<td></td>
<td>Compound-29</td>
<td><img src="image" alt="Compound-29" /></td>
</tr>
</tbody>
</table>

**TABLE 2.1 (Continued)**

IC$_{50}$ values:
- IC$_{50}$ = 4.1 µM (Doman et al., Maybridge 2002)
- IC$_{50}$ = 16.7 nM (Oh et al., disease PharmoScan 2004)
- IC$_{50}$ = 1.9 µM vs. IC$_{50}$ = 0.7 µM vs. PDE-5 (Yamazaki et al., 2005)
- IC$_{50}$ = 0.7 µM vs. PDE-5 (Yamazaki et al., 2005)
- IC$_{50}$ = 25 µM (Perola et al., 2000)

SOFTbank E-Book Center Tehran, Phone: 66403879,66493070 For Educational Use.
<table>
<thead>
<tr>
<th>Receptor</th>
<th>Disease / Response</th>
<th>Source</th>
<th>Software</th>
<th>EC50</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stat3 (Song et al., 2005)</td>
<td>Cancer</td>
<td>NCI DB Merck Index Sigma-Aldrich Ryan Scientific</td>
<td>DOCK 4.0</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>CB 2 receptor (Salo et al., 2005)</td>
<td>Alleviating pain inflammation</td>
<td>Maybridge</td>
<td>GOLD 4.0</td>
<td>$-\log EC_{50} = 5.3 \pm 0.2 \mu M$</td>
<td></td>
</tr>
<tr>
<td>NK 1 receptor (Evers and Klebe, 2004)</td>
<td>Inflammation immune response</td>
<td>Commercially available compounds</td>
<td>Selector Unity FlexX</td>
<td>IC$_{50} = 251$ nM</td>
<td></td>
</tr>
<tr>
<td>Alpha 1 Adrenergic Receptor (Evers and Klabunde, 2005)</td>
<td>Cardiovascular disease</td>
<td>Aventis compound library</td>
<td>GOLD 2.0</td>
<td>$K_i = 1.4$ nM</td>
<td></td>
</tr>
<tr>
<td>MCH 1 receptor (Clark et al., 2004)</td>
<td>Obesity</td>
<td>Commercially available compounds</td>
<td>FlexS Unity FlexX</td>
<td>IC$_{50} = 55$ nM</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Target Protein</th>
<th>Disease Treatment Area</th>
<th>Small Molecule Library</th>
<th>VS Software Used</th>
<th>Potency in Binding Assay</th>
<th>Structure of a Representative Hit Compound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hsp90 (Barril et al., 2005)</td>
<td>Cancer</td>
<td>Commercially available compounds-</td>
<td>rDOCK</td>
<td>IC$_{50}$ = 600 nM</td>
<td><img src="Compound-13" alt="Compound-13" /></td>
</tr>
<tr>
<td>AICAR Transformylase (Li et al., 2004)</td>
<td>Cancer</td>
<td>NCI DB</td>
<td>AutoDock 2.4</td>
<td>IC$_{50}$ = 600 nM</td>
<td><img src="NSC-30171" alt="NSC-30171" /></td>
</tr>
<tr>
<td>Rac GTPase (Gao et al., 2004)</td>
<td>Cancer</td>
<td>NCI DB</td>
<td>FlexX</td>
<td>IC$_{50}$ = 50 μM</td>
<td><img src="NSC-23766" alt="NSC-23766" /></td>
</tr>
<tr>
<td>Parasitic cysteine protease (Desai et al., 2004)</td>
<td>Parasitic disease</td>
<td>ChemBridge</td>
<td>GOLD</td>
<td>IC$_{50}$ = 1.0 μM</td>
<td><img src="Compound-1" alt="Compound-1" /></td>
</tr>
<tr>
<td>Compound Name</td>
<td>Disease</td>
<td>Provider 1</td>
<td>Provider 2</td>
<td>EC$<em>{50}$/IC$</em>{50}$/Ki (μM)</td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>---------------</td>
<td>---------------------</td>
<td>---------------------</td>
<td>-----------------------------</td>
<td></td>
</tr>
<tr>
<td>HRV coat protein</td>
<td>Infective</td>
<td>Maybridge</td>
<td>Cerius Software</td>
<td>EC$_{50} = 4.3$ μM</td>
<td></td>
</tr>
<tr>
<td>(Steindl et al., 2005)</td>
<td>disease</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SARS-CoV protease</td>
<td>Infective</td>
<td>NCI ACD MDDR</td>
<td>Dock 4.01</td>
<td>K$_i = 61$ μM</td>
<td></td>
</tr>
<tr>
<td>(Liu et al., 2005)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EGFR</td>
<td>Cancer</td>
<td>ChemBridge</td>
<td>ICM</td>
<td>IC$_{50} = 15.0$ μM</td>
<td></td>
</tr>
<tr>
<td>Chk1-kinase</td>
<td>Cancer</td>
<td>AstraZeneca</td>
<td>FlexX-Pharm</td>
<td>IC$_{50} = 450$ nM</td>
<td></td>
</tr>
</tbody>
</table>

**Chemical Structures:**

1. **Compound-13**
   - F$_3$CO- spacer
   - NH group
   - N$_2$N$_2$N$_2$ atom
   - CH$_3$ group

2. **C3930**
   - Chlorine atoms
   - Benzene rings

3. **MSK-039**
   - Nitrogen atoms
   - Oxygen atoms
   - H$_2$N group

4. **Pyrimidine hit**
   - Br atoms
   - N$_2$N$_2$ atom
   - OH group

*For Educational Use.*
obtained from NCI and tested using an in vitro fluorescene polarization (FP) binding assay at an initial dose of 100 μM, and seven were found to have binding affinity (IC\textsubscript{50} value) ranging from 1.6 to 14.0 μM). Compound 6 was the most potent, with an IC\textsubscript{50} value of 4 μM in inhibition of cell growth using the 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide assay. It was found that compound 6 induced apoptosis in cancer cells expressing high levels of Bcl-2. Furthermore, using NMR methods, compound 6 was shown to bind the BH3 binding site in Bcl-X\textsubscript{L}, with a binding constant of 7 μM.

In a different study, Lugovskoy et al., (2002) outlined a novel approach based on a combination of structural and computational approaches to discover potent inhibitors, called BH3Is, of Bcl-2 family proteins. In this case, the authors used NMR chemical shift perturbation as an efficient tool for rapidly mapping the interaction interfaces and used direct NMR-derived constraints to restrict the conformational space for molecular modeling routines. The authors utilized the program TreeDock as a molecular modeling module. TreeDock uses the Lennard-Jones potential as the scoring function to obtain the protein–compound complexes based primarily on shape complementarity. The models of complexes were validated through an independent set of NMR restraints. This method was employed to analyze structure–activity relationships in the BH3I/Bcl-X\textsubscript{L} complexes, and it was found that the free energies of the complexes calculated using the TreeDock routine correlated well with in vitro Bcl-X\textsubscript{L} binding affinities of the compounds. The close homology between in vitro affinities of BH3Is and computed interaction energy allowed the authors to set
up virtual screening of available online compound libraries. Ninety-three compounds were selected based on their similarity to BH3Is (more than 80%) and solubility (log $P$ less than 6.0) from Chembridge (www.hit2lead.com) and Chemnavigator (www.chemnavigator.com) libraries. To incorporate ligand flexibility, nonfixed dihedral angles in the compounds were varied, and their backbones were superimposed with ones of the structurally closed BH3Is. Finally, the TreeDock module was used to calculate the interaction energy of these compounds with Bcl-X_L, keeping the same anchoring-docking atom pair as was used in the modeling of the structurally closed BH3I compound. From this screen, a novel analogue of BH3I-1, which was termed BH3I-1SCH3 (structure shown in Table 2.1), was found, and the *in vitro* binding affinity of this compound to Bcl-X_L was tested ($IC_{50} = 8.5 \pm 0.6 \mu M$). The results indicated that the compound binds Bcl-X_L with affinity comparable to that of BH3I-1.

### 2.2.2.2 IAPs

Current knowledge identifies the inhibitor of apoptosis proteins (IAPs) as a major control point in the execution of cell death (Deveraux and Reed, 1999). IAPs comprise a family of caspase-inhibiting proteins characterized by a shared conserved sequence region, termed the baculoviral IAP repeat (BIR) domain (Salvesen and Duckett, 2002). Eight endogenous IAPs are known in the human system, all of which inhibit apoptosis. So far, their main physiological roles seem to be (1) the establishment of a threshold under which caspases are kept at bay and (2) providing a pool of active caspases, which can rapidly execute death after release. Growing evidence also suggests the participation of IAP proteins in other cellular functions apart from inhibiting caspases, including protein degradation, cell-cycle control, and signal transduction.

The X-linked inhibitor of apoptosis (XIAP) is a promising new molecular target for the design of novel anticancer drugs aiming at overcoming apoptosis resistance of cancer cells to chemotherapeutic agents and radiation therapy. Recent studies demonstrated that the BIR3 domain of XIAP where caspase-9 and Smac proteins bind is an attractive site for designing small-molecule inhibitors of XIAP (Srinivasula et al., 2001). Using the DOCK program, the authors (Nikolovska-Coleska et al., 2004) have performed computational structure-based database searches of the TCM-3D database containing 8221 small organic molecules with diverse chemical structures isolated from traditional Chinese medicinal herbs to identify potential small-molecule inhibitors that bind to the XIAP BIR3 domain where caspase-9 and Smac bind. The high-resolution structure of the XIAP BIR3 domain in complex with the Smac protein was used to define a binding site for the database searches. The sum of the electrostatic and van der Waals interactions as calculated in the DOCK program was used as the ranking score. The top 1000 candidate small molecules with the best scores were rescored. After reranking, the top 200 compounds were considered as potential small-molecule inhibitors of XIAP. A total of 36 potential small-molecule inhibitors, primarily from the Developmental Therapeutics Program of the National Cancer Institute and some from commercial sources, were tested for their binding affinities to the XIAP BIR3 protein in the optimized FP-based biochemical binding assay. The authors have discovered five natural products from the TCM-3D that bind to the XIAP.
BIR3 protein and directly compete with the SM7F peptide. Among these five inhibitors, embelin (structure shown in Table 2.1) is the most potent inhibitor, with an IC\textsubscript{50} value of 4.1 ± 1.1 μM from three independent experiments in triplicates, which is slightly less potent than the natural 9-mer Smac peptide (IC\textsubscript{50} value = 2.8 ± 0.7 μM). Embelin inhibits cell growth, induces apoptosis, and activates caspase-9 in prostate cancer cells with high levels of XIAP, but it has a minimal effect on normal prostate epithelial and fibroblast cells with low levels of XIAP. In stable XIAP-transfected Jurkat cells, embelin effectively overcomes the protective effect of XIAP to apoptosis, enhances etoposide-induced apoptosis, and has a minimal effect in Jurkat cells transfected with vector control. Taken together, the results showed that embelin is a fairly potent, nonpeptidic, cell-permeable, small-molecule inhibitor of XIAP and represents a promising lead compound for designing an entirely new class of anticancer agents that target the BIR3 domain of XIAP.

### 2.2.2.3 Acetylcholinesterase

A variety of neurological and neuromuscular disorders involve a diminution of cholinergic activity. Often the most effective treatments are ligands, which inhibit the breakdown of acetylcholine. Acetylcholinesterase inhibitors have been used clinically in the treatment of Alzheimer’s disease. Alzheimer’s disease is characterized by amyloid plaques and neurofibrillary tangles and is associated with a loss of cholinergic neurons. The loss of cholinergic neurons is progressive and results in profound memory disturbances and irreversible impairment of cognitive function. Recent efforts have been directed toward the development of novel strategies for the treatment of Alzheimer’s disease. One strategy for the treatment of Alzheimer’s patients has been the use of acetylcholinesterase inhibitors to increase the levels of acetylcholine in the synapse, thereby enhancing cholinergic activity in the affected brain regions. Novel inhibitors of acetylcholinesterase (AChE) have been discovered for the treatment of Alzheimer’s disease using RBVS. Mizutani and Itai (2004) used the X-ray crystal structure of \textit{Torpedo californica} AChE complexed with decamethonium (PDB code 1ACL). AChE has a deep cavity containing the catalytic triad (also called the “active site gorge”), and the residues in the active site gorge were defined as the ligand binding site. The software ADAM & EVE was used for the receptor-based virtual screening studies. This software uses the automated docking procedure and takes into account the flexibility of each small molecule in the database by exploring the conformational space fully and continuously. The authors screened commercially available databases (ACD and Maybridge) of 160,000 molecules. Promising ligand candidates were selected according to a number of user-defined parameters. The small-molecule database was filtered prior to docking, and compounds with molecular weight ≥300, number of atoms ≥40, number of heteroatoms ≥2, and ≤10, number of ring structures ≥2, and no carboxylate, phosphate, or sulfate functional groups were selected for docking. The criteria for the most stable docked model of each compound included intermolecular energy < 20.0 kcal/mol, intramolecular energy of each compound < 50.0 kcal/mol, and intermolecular hydrogen bonds ≥1. The intermolecular energy \( E_{\text{inter}} \) was used as a score function in the docking.
process and in selecting hit compounds, where $E_{\text{inter}}$ was defined as

$$E_{\text{inter}} = E_{\text{vwd}} + E_{\text{elc}} - 2.5N_{\text{hb}}$$

where, $E_{\text{vwd}}$ is the intermolecular van der Waals energy, $E_{\text{elc}}$ is the electrostatic energy, and $N_{\text{hb}}$ is the number of hydrogen bonds.

One hundred fourteen hit compounds were available and purchased for assay. The resulting hit compounds had different core structures, and a majority had structures that were quite different from the known AChE inhibitors. Of the 114 compounds tested, 35 showed inhibitory activities with IC$_{50}$ values less than 100 $\mu$M. Thirteen of these compounds had IC$_{50}$ between 0.5 and 10 $\mu$M. The most potent compound had an IC$_{50}$ of 0.59 $\mu$M. The structure of a representative hit is shown in Table 2.1.

In another study of AChE inhibitors, Rollinger et al., (2004) generated a structure-based pharmacophore model utilizing an in silico filtering experiment for the discovery of promising candidates out of a 3D multiconformational database consisting of more than 110,000 natural products. Scopoletin and its glucoside scopolin emerged as potential AChE inhibitors by the virtual screening procedure. They showed moderate, but significant, dose-dependent and long-lasting inhibitory activities. In the in vivo experiments, scopoletin (structure shown in Table 2.1) and glucoside scopolin increased the extracellular acetylcholine (ACh) concentration in rat brain to about 170% and 300% compared to basal release, respectively. At the same concentration, the positive control galanthamine increased the ACh concentration to about the same level as scopoletin. These are the first in vivo results indicating an effect of coumarins on brain ACh.

### 2.2.2.4 Estrogen Receptors

Two nuclear estrogen receptors (ERs), ER$\alpha$ and ER$\beta$, have been identified. There is roughly 60% homology in the ligand-binding domains between the human ER$\alpha$ and ER$\beta$ (Mosselman et al., 1996). ER$\beta$ is considered to be a pharmacological target to promote memory function and neuronal defense mechanisms against age-related neurodegeneration, such as Alzheimer’s disease. Zhao and Brinton (2005) have used RBVS in this case to identify ER$\beta$-selective compounds. It is a challenging task because there are only two residue differences in the active sites of ER$\alpha$ and ER$\beta$. Structural and mutational studies have indicated that Met336 and Ile373 in ER$\beta$ replace the variant amino acids Leu384 and Met421 in the ligand-binding pocket of ER$\alpha$, and these are the residues that need to be targeted to increase the selectivity of the ligand for one of the receptor subtypes. Using the crystal structure of ER$\beta$ LBD complexed with genistein and a plant-based natural library of 25,000 small molecules, the authors used an RBVS approach to identify ER$\beta$-selective compounds. The receptor binding site was defined based on the position of genistein in the receptor and specified as all atoms within 10 Å of the center carbon of genistein. The program GOLD was used in the docking and scoring studies. Prior to database screening, 20 validation runs were performed on test complexes to figure out the GOLD parameter settings required to capture the reported binding mode of genistein. Using these parameters, the 3D natural source chemical database was docked into the ER$\beta$ active
site and scored based on the GoldScore fitness function. Five hundred resulting top-scoring hits were filtered via visual screening. Based on visual analysis, 100 of these were redocked using Affinity. These 100 were selected based on their binding characteristics in the active site pocket. Finally, the compounds were subjected to a final round of screening in which the ones that passed the Lipinski’s Rule of Five, and blood–brain barrier penetration properties were selected. Twelve representative hits from the database screening were assessed for their binding profiles to both ERs. Of these, five compounds showed binding selectivity to ERβ over ERα, and the most active compound (structure shown in Table 2.1) had an IC₅₀ of 680 nM in ERβ binding assay, with over 100-fold binding selectivity to ERβ over ERα.

### 2.2.2.5 Dipeptidyl Peptidase IV (DPP-IV)

DPP-IV is a membrane-bound, homodimeric class II protein with a molecular weight of 110–150 kD per subunit. It is bound to the membrane by a transmembrane sequence of approximately 22 amino acids. The six cytosolic amino acids play no role in the binding functions. In addition to its protease function, DPP-IV has receptor properties and is an extracellular binding protein (Lambeir et al., 2003). DPP-IV has been regarded as one of the most promising therapeutic targets for the treatment of type 2 diabetes in recent years (Demutha et al., 2005). Inhibition of DPP-IV in normal and diabetic rats has been shown to prevent the degradation of GLP-1, leading to enhanced insulin secretion and improved glucose tolerance (Ahren et al., 2002). DPP-IV is able to selectively cleave Xaa-Pro and Xaa-Ala dipeptides from the N-terminus of peptides and proteins (where “Xaa” is any amino acid).

Several DPP-IV crystal structures have recently been published, which allow structure-based drug design to be used in the search for new and different classes of DPP-IV inhibitors. Published DPP-IV crystal structures show a variety of bound ligands, including Val-Pyr, in the 1N1M (Rasmussen et al., 2003) structure. Other published structures including 1PGQ (Oefner et al., 2003), 1NU6, and 1NU8 (Thoma et al., 2003) have very similar conformations around the active site, affirming that this may be a suitable target for virtual screening. Ward et al. (2005) used a subset of 800,000 compounds for the starting point of the virtual screen. Filtering on physical properties and chemical filters reduced this number to around 500,000 compounds. Tautomeric, protonation state, and stereogenic center enumeration using the in-house software and Corina increased the number of structures to around 750,000 in the single-conformer database. The number of structures in the multiconformer database is substantially higher, because up to 500 conformations from each compound are built. This multiconformational database was then processed against the two designed DPP-IV pharmacophores. Twenty thousand hits from each of the pharmacophores were elected based on the RMSD overlap of the compound with the pharmacophore and the overlap with the excluded volume of the active site. A single-conformer database of the 40,000 selected molecules was then docked into the DPP-IV crystal structure using GLIDE. Ten docked poses for each compound were postprocessed after which the top 8000 compounds were picked. Clustering and then visual inspection were used to select the final 4000 compounds for screening. These had activities ranging from 30% to 82% when tested at a concentration of 30 μM in an
enzyme inhibition assay, and the most potent example (structure shown in Table 2.1) had an inhibition of 81.9\% at 30 \mu M.

2.2.2.6 Human Aldose Reductase

Aldose reductase (AR, EC 1.1.1.21), a member of the aldo–keto reductase superfamily, is a monomeric enzyme that catalyzes the reduction of aldo sugars and a variety of aldehydes to their corresponding alcohols (Yabe-Nishimura, 1998). AR adopts a (\beta/\alpha)_8 TIM-barrel fold with the catalytic site located in the center of the barrel at the bottom of a deep hydrophobic cleft. The enzyme is NADPH-dependent, and the nicotinamide moiety of the cofactor is largely buried at the bottom of the barrel where it closes up the catalytic center (Borhani et al. 1992).

Inhibition of AR provides a therapeutic means for the prevention of complications associated with chronic diabetes. Kraemer et al. (2004) performed a virtual screen based on the high-resolution crystal structure of the inhibitor IDD594 in complex with human AR. AR operates on a large scale of structurally different substrates. To achieve this pronounced promiscuity, the enzyme can adapt rather flexibly to its substrates. Likewise, it has a similar adaptability for the binding of inhibitors. The authors applied a protocol of consecutive hierarchical filters to search the Available Chemicals Directory. In the first selection step, putative ligands were chosen that exhibit functional groups to anchor the anion-binding pocket of AR. Subsequently, a pharmacophore model based on the binding geometry of IDD594 and the mapping of the binding pocket in terms of putative “hot spots” of binding was applied as a second consecutive filter. In a third and final filtering step, the remaining candidate molecules were flexibly docked into the binding pocket of IDD594 with FlexX and ranked according to their estimated DrugScore values. Of the 206 compounds selected by this search and complemented by a cluster analysis and visual inspection, nine compounds were selected and subjected to biological testing. Of these, six compounds showed IC_{50} values in the micromolar range, with the best inhibitors BTB02809 (structure shown in Table 2.1) possessing an IC_{50} value of 2.4 ± 0.5 \mu M.

2.2.2.7 Protein Tyrosine Phosphatase-1B (PTP-1B)

Protein tyrosine phosphatase-1B (PTP-1B) is a tyrosine phosphatase that has been implicated as a key target for type 2 diabetes (Møller et al., 2000). This phosphatase hydrolyzes phosphotyrosines on the insulin receptor, thus deactivating it. Doman et al. (2002) defined the site targeted in the docking calculations by the positions of the two phosphotyrosine molecules observed in the complex with PTP-1B (Puius et al., 1997). The closed, ligand-bound conformation of the active site was used in the docking studies. Molecules were docked into the active site of PTP-1B in multiple conformations using the Northwestern University version of DOCK3.5. Two databases of molecules were used as follows: about 152,000 compounds from the ACD98.2 database and about 82,000 compounds from the BioSpecs 1999 and the Maybridge 1999 catalogs. Only molecules that had at least 17 and no more than 60 nonhydrogen atoms were docked. This amounted to 165,581 molecules docked in total. The atomic
properties of these molecules, including van der Waals parameters and partial atomic charges, were calculated, and conformations were precalculated using SYBYL and stored in a flexibase. Compounds selected by the docking program were screened for their ability to inhibit the PTP-1B dephosphorylation of the insulin receptor peptide. The top-scoring 500 molecules from the flexible docking screen of the ACD database and the top-scoring 500 molecules from the flexible docking screen of the combined BioSpecs and Maybridge databases were considered for further evaluation. Of the 500 high-scoring ACD molecules, 118 were tested experimentally; 38 had IC$_{50}$ values of 100 µM or better against PTP-1B. Additionally, 15 molecules from a preliminary rigid body docking screen of the ACD were also assayed; eight had IC$_{50}$ values of 100 µM or lower (a combined hit rate of experimentally tested docking molecules of 34.6%). Of the 500 high-scoring BioSpecs/Maybridge molecules, 232 were tested experimentally; 81 had IC$_{50}$ values of 100 µM or better against PTP-1B, a hit rate of 34.9%. Overall, 34.8% of the high-scoring docked molecules that were tested (127 of 365) had IC$_{50}$ values of 100 µM or lower. The structure of a typical representative is outlined in Table 2.1. This representative had an IC$_{50}$ of 4.1 µM for PTP-1B.

2.2.2.8 ADAMs

ADAMs (a disintegrin and metalloproteases) are type 1 transmembrane metalloproteases that are involved in cell–cell adhesion, surface proteolysis, and transmembrane signaling. Recently, ADAM12 was shown to play an important role in the development of cardiac hypertrophy by shedding heparin-binding epidermal growth factor (Asakura et al., 2002). Therefore, inhibition of ADAM12 could be a potent therapeutic strategy for cardiac hypertrophy and congestive heart failure. Some ADAM12 inhibitors have been reported to repress ADAM12 in both in vitro and in vivo experiments (Tokumaru et al., 2000). However, these inhibitors have broad spectra to MMPs. Oh et al. (2004) identified potent and selective ADAM12 inhibitors through structure-based virtual screening. The authors constructed the homology mode of ADAM12 based on the published structures of reference proteins that had over 30% sequence similarity to ADAM12. The reference proteins in this study were adamanxyn (Crotalus adamanteus; PDB code 4AIG), acutolysin-C (Agkistrodon acutus; PDB code 1QUA), and acutolysin-C (Agkistrodon acutus; PDB code 1BUD). Modeling of ADAM12 was initially performed by using the HOMOLOGY module of INSIGHT2000, and it was further refined by using Discover (version 2.98) of INSIGHT2000. A pharmacophore map for the catalytic site of ADAM12 was generated by using PharmaMap/PharmoScan. Since ADAM12 is a Zn$^{2+}$-containing metalloenzyme, the Zn$^{2+}$-binding moieties are prerequisite for potent inhibitors against the enzymatic activity. Therefore, a focused virtual library, composed of 67,062 compounds containing a Zn$^{2+}$-binding moiety such as carboxyl, hydroxyl sulfur, or hydroxamate, was prepared. Virtual screening of this library for the ADAM12 pharmacophore using the PharmaMap/PharmoScan system resulted in 1217 compounds as initial hits, which were further refined by a consensus scoring function composed of Ligscore, PLP, JAIN, PMF, and LUDI. Applying a criterion of the top 30% of each score, 64 compounds were finally selected for biological testing. The best compound had an IC$_{50}$ value of 16.7 nM in a cell-based assay for selective
inhibition of ADAM12. The structure of this compound is shown in Table 2.1. To understand the mechanism underlying the highly selective inhibition of ADAM12 by the most active compound (called compound 5), the authors compared the docking modes of this active compound in the active sites of ADAM12 and TACE (also known as ADAM17). Although the compositions of amino acids in the active sites are highly conserved and the total binding energy is similar (−78 kcal/mol for the 5-ADAM12 complex and −53 kcal/mol for the 5-TACE complex), the binding modes of compound 5 in the active sites of ADAM12 and TACE were significantly different. Consistent with the proper docking of compound 5 in the pharmacophore of the Zn$^{2+}$-binding site of ADAM12, compound 5 fit into the active site of ADAM12 with the conserved Zn$^{2+}$-binding motif and hydrophobic interactions. However, compound 5 could not be docked in the active site of TACE because the hydrophobic pocket in this case was much smaller than that of ADAM12. Therefore, the hydrophobic interaction as well as the Zn$^{2+}$–carboxy group interaction might serve as an essential factor for the specific inhibition of ADAM12 by compound 5. The reason for this low nanomolar hit for virtual screening could be due to the combined use of the pharmacophore of the Zn$^{2+}$-binding site of ADAM12 and receptor-based docking to identify likely hits.

2.2.2.9 Phosphodiesterase-1 and -5

Phosphodiesterase types 1 (PDE-1) and 5 (PDE-5) are the major cGMP hydrolyzing enzymes in blood vessels and regulate the level of the mediator in concert with guanylyl cyclase, which catalyzes the synthesis of cGMP from GTP. PDE-1 and -5 have also been found in platelets. Therefore, inhibitors of PDE-1 and -5 are expected as therapeutics for cardiovascular diseases, such as hypertension, angina, cardiac failure, and obstructive arteriosclerosis (Wallis et al., 1999). Yamazaki et al. (2005) used classification and regression tree (CART) analysis to configure a prediction model for virtual screening. An optimum set of structural descriptors were selected as explanatory variables for CART analysis. A library of commercially available chemical compounds supplied by SPECS Inc. was screened for PDE-5 inhibitory activity by virtual screening. The authors used only 50,520 compounds that were not listed on the catalog of SPECS Inc. in September 1998 but were listed in October 1999 so that there would be a higher possibility that these compounds were actually in stock. In the first step of the virtual screening, the compounds with unfavorable physicochemical and pharmacokinetic properties were filtered out based on a modified Lipinski’s Rule of Five; that is, molecular weight below 500, number of hydrogen bond acceptors below 10, number of hydrogen bond donors below 5, AlogP below 6, and number of rotatable bonds below 12. In the next step, the compounds with favorable biological activities were selected with the CART model for PDE-5 inhibition. SPSS ver.10.0 was used in both steps of the virtual screening. In the first step of virtual screening of a commercially available chemical library, 43,365 compounds were selected by a modified Lipinski’s Rule of Five (85.8%). Among them, 1821 compounds were finally selected as PDE-5 inhibitor candidates (3.6%). Nineteen diverse and druglike compounds selected from the virtual screening hits were assayed for the inhibitions of PDE-1 and -5. Among them, seven compounds (37%) showed inhibitory activities.
against both PDE-1 and -5 below 10 μM as IC₅₀. The structure of a typical representative is shown in Table 2.1.

### 2.2.2.10 Metalloenzymes

**Carbonic Anhydrase II, Matrix Metalloproteinase-3, Neutral Endopeptidase, Peptide Deformylase, and Xanthine Oxidase**

Metalloenzymes have therapeutic implications in many diseases such as cancer (Kruger and Figg, 2000; Bradshaw and Yi, 2002), arthritis (Beaudeux et al., 2004), glaucoma (Cvetkovic and Perry, 2003), and infectious diseases (Eswaramoorthy et al., 2002). However, these proteins are challenging targets for virtual screening because of the covalent-like interactions between the metal centers and their ligands and the large electrostatic potentials of the metals. This is because the molecular mechanics potential functions that underlie most modeling approaches have been parameterized to treat atoms as either interacting covalently or noncovalently, and metal–ligand interactions do not fit into either of these categories. Irwin et al. (2005) have explored the robustness of RBVS methods using classical scoring functions to identify potential ligands for several metalloenzymes. To do this, the authors have treated the metals in the enzymes like any other atoms in the enzyme, with van der Waals radii, potentials, and point charges. They have tested the protocol by docking a functionally annotated database of 95,000 compounds from the MDL Drug Data Report (MDDR) against the X-ray crystal structures of five metalloenzymes, such as carbonic anhydrase II, matrix metalloproteinase-3, neutral endopeptidase, peptide deformylase, and xanthine oxidase. The quality of the docking scheme was judged based on the enrichment of the annotated ligands among the top-scoring hits and reproduction of experimental geometries. In all five test cases, the annotated ligands within the MDDR were enriched at least 20 times over random. They found this approach to be reliable for zinc metalloenzymes, such as carbonic anhydrase II (where the highest-scoring docked geometry had an rms deviation of 0.4 Å from the crystal structure and 82-fold enrichment of hits in the top 0.1% of the database), matrix metalloproteinase-3 (where the highest-scoring docked geometry had an rms deviation of 0.7 Å from the crystal structure and 32-fold enrichment of hits in the top 0.1% of the database), and neutral endopeptidase (where the highest-scoring docked geometry had an rms deviation of 1.2 Å from the crystal structure and 189-fold enrichment of hits in the top 0.1% of the database). Docking against a nickel-containing enzyme, peptide deformylase (where the highest-scoring docked geometry had an rms deviation of 0.9 Å from the crystal structure and 47-fold enrichment of hits in the top 0.8% of the database), and against a molybdenum enzyme, xanthine oxidase (where the highest-scoring docked geometry had an rms deviation of 0.7 Å from the crystal structure and 36-fold enrichment of hits in the top 0.6% of the database), was reported as satisfactory. The authors applied this protocol to identify inhibitors of zinc-β-lactamase, an antibiotic resistance target from *Bacteroides fragilis*, for which there are no inhibitors in clinical use. The 33,000 fragment-like molecules used in this case were from the chemical library of the ZINC database (Irwin and Shoichet 2005). Out of the 50 top-ranked compounds from docking and scoring, 15 were purchased and evaluated. Of these, five
compounds were found to have $K_i$ values less than 120 $\mu$M, the best of which had a $K_i$ of 2 $\mu$M. In the next application they identified 25 potential substrates of a Zn-dependent phosphotriesterase from *Pseudomonas diminuta* by screening 167,000 compounds in the ACD. Eight of the 25 hits were acquired, tested, and confirmed experimentally to be substrates. This study suggested that a simple, noncovalent scoring function may be used to identify inhibitors of at least some metalloenzymes.

**Farnesyltransferase** Farnesyltransferase (FT) catalyzes the transfer of the farnesyl group from the cosubstrate farnesyl pyrophosphate (FPP) to a cysteine residue in the C-terminal fragment of pro-Ras proteins. FT consists of an $\alpha$ subunit (48 kDa) and a $\beta$ subunit (46 kDa). The two subunits form a large cationic residue-rich active site in the middle of the enzyme. The zinc divalent cation present in the active site is known to play a functional role in FT catalysis and to facilitate the binding of the substrates (Park et al., 1997). FT is a promising anticancer drug target (Gibbs and Oliff, 1997) because selective inhibitors of FT have been found to be able to inhibit tumor cell proliferation without substantially interfering with normal cell growth (Kohl et al., 1993; Prendergast et al., 1994; Sebti and Hamilton, 1997). From the standpoint of ligand docking, it is challenging to work with FT since the active site of FT is very large and contains a metal atom (Zn, in this case). Additionally, the force-field parameters of the zinc divalent cation required in docking studies has not been adequately represented in docking software programs. Perola et al. (2000) performed a virtual screening, containing 219,390 chemicals, of the ACD to identify prototypic inhibitors of zinc-bound FT. To achieve this, the authors developed a new docking program, EUDOC, and proposed an integrated, broad, conformation substitution-based (IBS) virtual-screening approach. Using the docking program and the IBS virtual-screening approach, 21 compounds were identified from the ACD database, four of which inhibited FT *in vitro* with IC$_{50}$ values in the range of 25–100 $\mu$M. The most potent inhibitor (structure shown in Table 2.1) also inhibited FT in human lung cancer cells.

### 2.2.2.11 Stat3

Signal transducers and activators of transcription 3 (Stat3) are activated in response to cytokines and growth factors. The main domains of Stat3 protein include the tetramerization and leucine zipper at the N terminus, the DNA-binding domain, and the SH2 transactivation domain at the C-terminal end. The SH2 region is responsible for the binding of Stat3 to the tyrosine-phosphorylated receptors and for the dimerization that is necessary for DNA binding and gene expression (Zhong et al., 1994). Stat3 is activated by phosphorylation at tyrosine residue 705 (Tyr-705), which leads to dimer formation, nuclear translocation, recognition of Stat3-specific DNA binding elements, and activation of target gene transcription. The constitutive activation of Stat3 is frequently detected in breast carcinoma cell lines but not in normal breast epithelial cells. Blockade of Stat3 signaling inhibits cancer cell growth, indicating that Stat3 plays a role in the survival or growth of tumor cells. It has been reported that ~60% of breast tumors contain persistent activated Stat3. Thus, Stat3 represents an attractive target for the development of new anticancer therapy in the treatment of breast cancer.
To identify potential small-molecule compounds that can disrupt Stat3 dimerization, Song et al. (2005) used the X-ray crystal structure of Stat3β (PDB code 1BG1) and a docking and scoring strategy to screen a virtual library of 429,000 compounds collectively from NCI, Merck Index, Sigma-Aldrich, and Ryan Scientific compound databases. The docking program DOCK (version 4.0) was used to perform the virtual screening. The binding cavity on the Stat3 SH2 domain was the region targeted for docking. SYBYL software (version 6.9, Tripos Associates, St. Louis, MO) was used to assign the standard AMBER (a set of molecular mechanical force fields for the simulation of biomolecules) atomic partial charges on the Stat3β protein and the Gasteiger–Hückel atomic partial charges on each ligand molecule to be docked. Each compound with a molecular weight between 200 and 700 was docked into the targeted binding site in Stat3. The top 10% scored compounds from each database, as selected by DOCK, were extracted and combined together to provide a total of 35,000 candidate compounds. Based on the binding models of these compounds predicted by DOCK, the X-SCORE program (version 1.1) was applied to obtain an estimate of the binding affinities of these compounds for Stat3. The preselected 35,000 compounds were then reranked according to their binding affinities as estimated by X-SCORE. Of the best-scored 200 compounds selected by X-SCORE from the 35,000 compounds, physical samples of 100 compounds were obtained from NCI or purchased from Sigma-Aldrich or Ryan Scientific. One compound, termed STA-21 (structure shown in Table 2.1), was identified as the best match for this selection criteria. STA-21 inhibited Stat3 DNA binding activity, Stat3 dimerization, and Stat3-dependent luciferase activity at 20 μM. Moreover, STA-21 reduced the survival of breast carcinoma cells with constitutive Stat3 signaling but had minimal effect on the cells in which constitutive Stat3 signaling was absent at 20 μM or 30 μM.

### 2.2.2.12 GPCRs

**Serotonin, Tachykinin NK1, Dopamine D2, and Chemokine CCR3 Receptors**

G-protein-coupled receptors (GPCRs) are implicated in many human diseases, such as inflammation, hypertension, pain, depression, obesity, depression, and anxiety. As a result, they are the targets for nearly 50% of all drug-discovery programs (Klabunde and Hessler, 2002). GPCRs are especially challenging targets because they are membrane-bound proteins, which make their experimental 3D structure determination difficult. Although ligand-based *in silico* approaches have been used to identify ligands for this family of proteins, nonetheless, structure-based drug discovery remains highly desirable for this family of proteins. Becker et al. (2004) have used *in silico* 3D models of several GPCRs and have used RBVS to identify novel, potent analogs targeting the serotonin receptor, tachykinin NK1 receptor, dopamine D2 receptor, and chemokine CCR3 receptor. In each case, the authors have constructed the PREDICT 3D model of the respective target protein from its amino acid sequence using the PREDICT algorithm and methodology (Shacham et al., 2001, 2004) and have used this model in RBVS studies. PREDICT is a *de novo* GPCR-modeling methodology that combines the properties of the protein sequence with those of its membrane environment without relying on any GPCR.
experimental structure. The algorithm searches through the receptor conformational space for the most stable 3D structure(s) of the transmembrane domain of the GPCR protein within the membrane environment. A binding pocket was defined in each case while docking and scoring were carried out using the program DOCK4.0 to screen more than 100,000 compounds for each target. This method was successful in identifying promising lead candidates for these GPCRs. For the serotonin 5-HT1A receptor, 78 hits were tested using \textit{in vitro} assays. Of the 16 hits with $K_i$ less than 5 \textmu M, five were found to be novel compounds (new chemical entities—that is, compounds not previously described in literature), with the best compound having a potency of 1 nM. For the serotonin 5-HT4 receptor, 93 virtual hits were tested using \textit{in vitro} binding assays, and 19 of those had $K_i$ less than 5 \textmu M. Of these, four had distinct chemical scaffolds, and the best novel hit had a $K_i$ of 21 nM. In the case of the tachykinin NK1 receptor, there were 53 virtual hits that were tested using \textit{in vitro} binding assays, and eight had $K_i$ less than 5 \textmu M. They were grouped into five distinct chemical scaffolds, and the best novel hit was a 56 nM compound. Finally, for the chemokine CCR3 receptor, of the 43 virtual hits identified by this method and tested using \textit{in vitro} binding assays, five compounds had $K_i$ less than 20 \textmu M. The best novel hit had a $K_i$ of 12 \textmu M. This work has paved the way for the application of RBVS approaches for the GPCR family of proteins.

\textbf{Human Cannabinoid CB2 Receptor} The cannabinoid CB2 receptor is a GPCR located mainly in immune tissues. It is part of the endogenous cannabinoid system (ECS) that includes the endocannabinoids (e.g., 2-arachidonoylglycerol [2-AG]), their receptors (CB1 and CB2), metabolizing enzymes (fatty-acid amide hydrolase [FAAH] and monoglyceride lipase [MGL]), and a specific cellular uptake system. Since ECS is an attractive target for drug development, various cannabinergic ligands that act on one or more ECS target proteins have been developed during recent years (Di Marzo et al., 2004). The CB2 receptor ligands possess potential in alleviation of pain and inflammation as well as in treatment of chronic cough, gliomas, lymphomas, and osteoporosis. Many cannabinoid receptor ligands are nonselective in that they bind both to the central CB1 receptor as well as to the peripheral CB2 receptor. It is claimed that selective CB2 ligands should be devoid of the unwanted central nervous system (CNS) effects typical of $\Delta^9$-tetrahydrocannabinol ($\Delta^9$-THC), the major psychoactive component of \textit{Cannabis sativa} L (Gaoni et al., 1964).

To identify novel selective CB2 lead compounds, Salo et al. (2005) constructed a comparative model of the CB2 receptor using the high-resolution bovine rhodopsin X-ray structure as a template. The HOMOLOGY module of InsightII (version 2000) modeling software was utilized to construct the CB2 receptor model, whereas structure optimization, visualization, and ligand structure construction were carried out using the SYBYL6.9 modeling package. Molecular dynamics (MD) simulations were performed with GROMACS v.3.1.4, and the stereochemical quality of the protein structures was checked with PROCHECK. GOLD2.0 was used for ligand docking, and both CScore and X-Score v. 1.1 were used for ranking the docking results. CoMFA, as implemented in SYBYL, was used for generating 3D-QSAR models. Lead-structure searches in the Maybridge molecular database were carried out with the UNITY4.4 module of SYBYL. Five different receptor conformers and
three different database queries were used to virtually screen the Maybridge Database (~55,600 compounds) for the discovery of selective CB2 ligands. Both the docked CB2 ligand structures and the modeled receptor binding site were employed in building the 3D database queries. Query 1 (Q1) was based on the pharmacophoric points of the nonselective cannabinoid receptor agonist HU-210 and the multiple volume surface of six other CB2 agonists all in their docking conformations. The second query (Q2) was based on the binding cavity of the docked CB2 ligands, whereas the third query (Q3) was based on the combination of features taken from both the receptor and bound ligands. Finally, 86 hit molecules (0.15% of the whole database) were ordered from Maybridge to be tested for their in vitro activity at the CB2 receptor. In G-protein activation assays, 1-isoquinolyl[3-(trifluoromethyl)phenyl]methanone (NRB 04079) ($-\log EC_{50} = 5.3 \pm 0.2$) was found to act as a selective agonist at the human CB2 receptor. NRB 04079 (structure shown in Table 2.1), therefore, can serve as a lead structure for designing selective CB2 agonists that are useful in the management of pain.

**Neurokinin-1 (NK1) Receptor** The neurokinin-1 (NK1) receptor belongs to the family of GPCRs, which represents one of the most relevant target families in small-molecule drug design. Evers and Klebe (2004) describe the use of a homology model of the NK1 receptor based on the high-resolution X-ray structure of rhodopsin and the successful virtual screening based on this protein model. The NK1 receptor model was generated using the new MOBILE (modeling binding sites including ligand information explicitly) approach developed in the authors’ laboratory. Starting with preliminary homology models, this software generates improved models of the protein binding pocket together with bound ligands. Ligand information is used as an integral part in the homology modeling process. For the construction of the NK1 receptor, antagonist CP-96345 was used to restrain the modeling. The quality of the obtained model was validated by probing its ability to accommodate additional known NK1 antagonists from structurally diverse classes. On the basis of the generated model and on the analysis of known NK1 antagonists, a pharmacophore model was deduced, which was subsequently used to guide the 2D and 3D database search with UNITY. As a following step, the remaining hits were docked into the modeled binding pocket of the NK1 receptor. Finally, seven compounds were selected for biochemical testing, from which one showed affinity in the submicromolar range.

**Alpha1A Adrenergic Receptor** The authors started with the constructed homology model of the alpha1A receptor based on the X-ray structure of bovine rhodopsin. The protein model was generated by applying ligand-supported homology modeling, using mutational and ligand SAR data to guide the protein modeling procedure. Evers and Klabunde (2005) performed a virtual screen of the company’s compound collection to test how well this model is suited to identify alpha1A antagonists. They applied a hierarchical virtual screening procedure guided by 2D filters and 3D pharmacophore models. The ~23,000 filtered compounds were docked into the alpha1A homology model with GOLD and scored with PMF. From the top-ranked compounds, 80 diverse compounds were tested in a radioligand displacement assay. Thirty-seven compounds revealed $K_i$ values better than 10 $\mu$M; the most active
compound bound with 1.4 nM to the alpha1A receptor. These findings suggest that rhodopsin-based homology models used in conjunction with ligand-based pharmacophore models and 3D-QSAR modeling may aid in the identification of highly active compounds for the GPCR family of proteins.

**Melanin-Concentrating Hormone-1 Receptor (MCH-1R)** Melanin-concentrating hormone (MCH) has been known to be an appetite-stimulating peptide for a number of years. However, it is only recently that MCH has been discovered to be the natural ligand for a previously “orphan” G-protein-coupled receptor, now designated MCH-1R. This receptor has been shown to mediate the effects of MCH on appetite and body weight; consequently, drug-discovery programs have begun to exploit this information in the search for MCH-1R antagonists for the treatment of obesity. Clark et al. (2004) report the rapid discovery of multiple, structurally distinct series of MCH-1R antagonists using a variety of virtual screening techniques. The most potent of these compounds (12) demonstrated an IC$_{50}$ value of 55 nM in the primary screen and exhibited antagonist properties in a functional cellular assay measuring Ca$_{2+}$ release. More potent compounds were identified by follow-up searches around the initial hit.

**2.2.2.13 Hsp90** Docking-based virtual screening was used to identify a new class of Hsp90 inhibitor (Barril et al., 2005) with submicromolar potency. The molecular chaperone Hsp90 is involved in the onset and progression of cancer (Maloney and Workman, 2002), thereby making it an attractive target for anticancer drug discovery. The N-terminal domain of this protein binds to ATP and has been reported to drive the chaperone activity of this protein. The authors used a four-step RBVS protocol (Barril et al., 2004) to identify ligands binding to this site of Hsp90. In the first step, the relevant binding pocket was selected from all the structurally available models for this enzyme. The open and helical form of the receptor was used in docking studies. The binding pocket for purine analogs with three bound water molecules was defined as the active site. In the second step, a virtual small-molecule library of 0.7 million compounds was constructed starting from 3.5 million commercially available compounds. These were selected on the basis of their druglikeness, absence of reactive groups, and vendor delivery timelines. Corina was used to generate the 3D structures of the ligand database. In the third step, rDock was used to dock the library, and the 9000 highest-scoring compounds were short-listed for further evaluation. In the fourth and final step, a postfiltering was done, and only those compounds that formed a hydrogen bond with one carboxylic oxygen of Asp93 and a second hydrogen bond with one of the interstitial water molecules, as found in all the reported cocrystal structures of Hsp90-inhibitor complexes, were selected. Of the top 1000 compounds selected for purchase, 719 were actually available and tested for their Hsp90 inhibition capability. Three of the identified hits showed greater than 50% inhibition in the primary assay, two of which showed submicromolar potency. However, all three compounds that showed good activity in fluorescent polarization assays exhibited poor cellular activity, suggesting that the compounds are not well absorbed by the
cell. The authors went on to determine the X-ray crystal structures of the three active compounds complexed with Hsp90 to identify sites in the ligand that could be potentially modified to improve the cell-penetration ability of the respective hits (structure of a representative hit compound shown in Table 2.1).

### 2.2.2.14 AICAR Transformylase

AICAR transformylase (5-aminimidazole-4-carboxamide ribonucleotide transformylase) has been proposed as a potential target for antineoplastic drug development. Li et al. (2004) carried out virtual screening of the AICAR transformylase active site against the NCI diversity set of 1990 compounds using AutoDock, and they also identified eight potential novel scaffolds with micromolar inhibition activity. Initially, 44 compounds were selected from the NCI diversity set for in vitro evaluation on the basis of computational simulation of free energy of binding to the AICAR Tfase active site. Prior to the assay, 10 of the compounds were found to be water insoluble and 18 precipitated in the assay buffer. Of the remaining 16 that were tested, eight proved to have micromolar activity in enzymatic inhibition assays. In the next step, the authors expanded their repertoire of active compounds by taking each of the five most active compounds and doing a similarity compound search on the complete NCI database of 213,628 compounds. They found 11 additional hits from the secondary round of similarity search and AutoDock screening. These 11 candidates also turned out to have micromolar potency in enzymatic inhibition assays. The best compound (structure shown in Table 2.1) in this screening turned out to have a measured IC\textsubscript{50} of 600 nM in AICAR transformylase inhibition assays.

### 2.2.2.15 Rac GTPase

Rho family GTPases are molecular switches that control signaling pathways regulating cytoskeleton organization, gene expression, cell-cycle progression, cell motility, and other cellular processes (Etienne-Manneville and Hall, 2002). They can be activated through the interaction with the Dbl family guanine nucleotide exchange factors (GEFs) that catalyze their GTP/GDP exchange and connect them to the diverse stimuli from upstream cell surface receptors such as the GPCRs, growth factor receptors, cytokine receptors, and adhesion receptors (Zheng, 2001). Accumulating evidence has implicated Rho GTPases in many aspects of cancer development (Schmitz et al., 2000), because deregulated Rho GTPases have been discovered in many human tumors, including colon, breast, lung, myeloma, and head and neck squamous-cell carcinoma (Mira et al., 2000). Rho GTPases and the signal pathways regulated by them have thus been proposed as potential anticancer therapeutic targets (Erik Sahai, 2002).

Gao et al. (2004) have applied a structure-based virtual screening approach to the search for a Rac-GEF interaction-specific small-molecule inhibitor. In the 3D structure of the Rac1-Tiam1 complex, Trp\textsuperscript{56} of Rac1 is buried in a pocket formed by residues His\textsuperscript{1178}, Ser\textsuperscript{1184}, Glu\textsuperscript{1185}, and Ile\textsuperscript{1197} of Tiam1 and Lys\textsuperscript{5}, Val\textsuperscript{7}, Thr\textsuperscript{58}, and Ser\textsuperscript{71} of Rac1. To identify Rac1-specific inhibitors based on the structural features surrounding Trp\textsuperscript{56}, a putative inhibitor-binding pocket was created with residues of
Rac1 within 6.5 Å of Trp56 in the Rac1-Tiam1 monomer, including Lys5, Val7, Trp56, and Ser71. A 3D database search was performed to identify compounds whose conformations would fit into this pocket. The database was freely available from NCI, which includes the coordinates of >140,000 small chemical compounds. Taking the flexibility of the compounds into consideration during the screening process, the program UNITY, whose Directed Tweak algorithm allows a conformationally flexible 3D search, was applied. The small-molecule hits yielded by the UNITY program were next docked into the predicted binding pocket of Rac1 containing Trp56 by using the program FlexX, which allows flexible docking to protein-binding sites. By following the docking procedures, the compounds were ranked based on their predicted ability to bind to the binding pocket with the program CScore. CScore generates a relative, consensus score based on how well the individual scoring functions of the protein–ligand complex perform. Of the top 100 compounds that displayed high consensus scores, 58 compounds whose docking did not seem to involve residue Trp56 were crossed out by visual inspection. Taking into account the solubility and availability of the remaining compounds, 15 chemicals that show promising docking affinities by the scoring method were selected for further characterization. Among the compounds tested, NSC23766 (structure shown in Table 2.1) was the only one to significantly inhibit TrioN binding to Rac1. The inhibitory effect of NSC23766 appeared to be specific for the interaction between Rac1 and its GEF because it did not interfere with the Cdc42 binding to intersectin. It had a value of 50 μM of IC50.

2.2.2.16 Parasitic Cysteine Protease

Cysteine proteases constitute an important class of enzymes that play vital roles in the life cycle of many parasites (Sajid and McKerrow, 2002). A possible strategy for combating parasitic infections is to inhibit cysteine proteases that are crucial to parasite metabolism and reproductive function. Papain-like cysteine proteases from *Leishmania donovani*, *Plasmodium falciparum* (falcipain-2 and falcipain-3), and *Trypanosoma cruzi* (cruzain) have been identified and functionally characterized (Sajid and McKerrow, 2002). Several studies have confirmed the efficacy of peptidyl cysteine protease inhibitors in arresting and killing parasites in vivo and in tissue culture models of parasite replication or cell invasion (Sajid and McKerrow, 2002). Vinyl sulfone-based peptides, such as cruzain and falcipains, are known to be irreversible inhibitors of cysteine proteases (Shenai et al., 2003) by forming covalent bonds with the thiolate of the catalytic cysteine. Although such irreversible inhibitors are quite potent with IC50 values in the nanomolar range, the poor selectivity for parasitic cysteine proteases over the human cysteine proteases remains a significant concern (Engel et al., 1998). Also, it is desirable to design reversible inhibitors to minimize the potential toxicity that can be observed with irreversible inhibitors.

Desai et al. (2004) constructed homology models of falcipain-2 and falcipain and compared them with crystal structures of cruzain. They found that the structure of the binding pocket appears to be highly conserved across the protozoal cysteine proteases. This suggests the possibility of developing common inhibitors of these enzymes to treat malaria, leishmaniasis, and trypanosomiasis. A relatively common
binding pocket in each was defined to make sure that the common residues lining the active site were all included in the binding pocket definitions. Virtual screening was performed using the ChemBridge database (Express Pick, October 2001: ChemBridge Corporation, San Diego, CA) containing a library of about 241,000 compounds. The database was filtered to collect only druglike molecules for docking using the following protocol. Considering metals, ADME, and Lipinski’s Rule of Five, the filtered database consisting of approximately 60,000 compounds was subjected to Concord to generate 3D coordinates and was subsequently used for docking. Docking was carried out in three consecutive stages employing three different settings in GOLD. Starting with about 60,000 molecules for the first stage of docking, a total of 5500 top-ranking molecules were selected for the next stage. Next, the top 1500 molecules selected based on the second stage of docking were screened against both falcipain-2 and falcipain-3. It is interesting to note that the top 100 molecules in the falcipain-3 run were among the top 200 molecules in the falcipain-2 run, with 78% of molecules in the top 100. Finally, a total of 24 druglike nonpeptide inhibitors were identified, of which 12 compounds showed dual activity against falcipain-2 and falcipain-3. Four compounds (structure of a representative hit shown in Table 2.1) were found to inhibit both the plasmodial cysteine proteases as well as the *L. donovani* cysteine protease. These leads can be optimized into broad spectrum antiprotozoal drugs.

### 2.2.2.17 Human Rhinovirus Coat Protein

The human rhinovirus (HRV) belongs to the family of picornaviruses and is the main cause for common colds and a variety of other respiratory illnesses, including otitis media and sinusitis, and for exacerbations of asthma and reactive airways disease. These illnesses still lack effective antiviral treatment. The viral capsid is a promising and intensively studied target for drug development. This protein shell encapsulates a single, positive RNA strand and consists of 60 copies of four different viral proteins. HRV coat protein inhibitors act as capsid-binding antiviral agents that block the uncoating of the viral particles and/or inhibit cell attachment (Hadfield et al., 1999). Their binding site is located within a hydrophobic pocket situated at the bottom of a depression, a so-called canyon, on the capsid surface. In the absence of an inhibitor, this pocket can be empty or occupied by a pocket factor, a lipid or fatty acid. Structural conservation in this region among the different serotypes permits the development of broad-spectrum anti-HRV agents (Kolatkar et al., 1999). Steindl et al. (2005) applied a combined approach of virtual screening techniques (pharmacophore model generation, docking, and ligand clustering in combination with a principal component analysis [PCA] performed on a set of molecular descriptors) to select a small set of possible new HRV coat protein inhibitors from a large commercially available database. The authors used the program Catalyst to generate pharmacophore models, also termed hypotheses. A pharmacophore model defines a special arrangement of chemical features that are shared by different ligands and are obviously responsible for binding and, as a consequence, for biological activity. Such a model can be used as a query tool in 3D database mining, thereby provoking an enrichment of active compounds in the obtained hit lists. Using the crystal structure of an inhibitor
complexed with the HRV coat protein and the functional group definitions contained within the Catalyst software, a representation of the binding site was produced (structure-based pharmacophore model) and applied in screening the commercially available Maybridge database. The resulting hits were submitted to further investigation using a docking study as well as a PCA of a set of molecular descriptors, thus aiming at the final selection of a few compounds for biological testing. Application of a scoring function (SF) allows the estimation of binding affinities and therefore a ranking of the substances. PCA is a widely applied multivariate technique in drug discovery. It provides a reduction of data multidimensionality (i.e., the properties of the input molecules), gives an overview of the data, and detects trends, groupings, and outliers. In this study, PCA-based clustering was applied to estimate the similarity of the Maybridge hits with already known active molecules. The docking and scoring procedures as well as the PCA-based clustering were carried out within the Cerius software. The combination of pharmacophore model generation, docking, and PCA-based clustering presents a virtual screening and selection protocol aimed at the fast and reliable identification of potential new HRV coat protein inhibitors. Thereby, each step underwent thorough theoretical validation probing their performance on sets of known active and by all probability inactive compounds.

These combined efforts finally allowed the selection of six promising candidate structures that were tested in a biological assay exemplarily against one major receptor group virus serotype for in vitro anti-HRV activities. The ability of the selected compounds to inhibit HRV serotype 39 was evaluated. The biological results were determined by multiple-cycle CPE inhibition assay and expressed as EC$_{50}$ values. Microscopic and spectrophotometric reading confirmed that all six of the Maybridge compounds inhibit viral growth on OH-I and WI-38 cells and show activities in the micromolar range. The conclusion can be drawn that at this level of activity, we could achieve a success rate of 100% in the selected hit list. Particularly, one of the six promising candidate shows an EC$_{50}$ of 4.3 µM for OH-I cells. The structure of the candidate is shown in Table 2.1.

### 2.2.2.18 SARS-CoV Protease

Severe acute respiratory syndrome (SARS) is a respiratory illness that had a widespread dramatic outbreak in Asia, North America, and Europe in early 2003 (Lee et al., 2003). Evidence indicates that a previously unrecognized coronavirus exists, called SARS coronavirus, which is the leading hypothesis for the cause of SARS (Rota et al., 2003). It was known that the cleavage process of the SARS-CoV polyproteins by a special proteinase, the so-called SARS coronavirus 3C-like proteinase (CoV Mpro), is a key step for the replication of SARS-CoV (Fan et al., 2004). The SARS coronavirus 3C-like proteinase is considered as a potential drug design target for the treatment of SARS.

Liu et al. (2005) built a “flexible” 3D model for SARS 3C-like proteinase by using a homology modeling and multicanonical molecular dynamics method, and they used the model for virtual screening of chemical databases. A 3D homology model of SARS 3C-like proteinase based on the structure of transmissible-gastroenteritis-virus coronavirus 3C-like proteinase was built by using MODELLER6.0...
program. Using the MPI Edition DOCK4.01 program, the authors screened several chemical databases including the National Cancer Institute Diversity Set (~230,000 compounds total), ACD-3D (Available Chemical Database, Release: ACD 3D 2002.2, ~280,000 compounds in total), and MDDR-3D (MDL Drug Database Report, Release: MDDR 3D 2002.2, 120,000 compounds in total). After DOCK procedures, strategies including pharmacophore modeling, consensus scoring, and “druglike” filtering were applied in order to accelerate the process and improve the success rate of virtual docking screening hit lists. A total of 126 molecules with more than 10 types of chemical structures were selected for further study. Forty compounds were purchased and tested by HPLC and colorimetric assay against SARS 3C-like proteinase. Three of them—including calmidazolium, a well-known antagonist of calmodulin—were found to inhibit the enzyme with an apparent $K_i$ from 61 to 178 μM. These active compounds and their binding modes provide useful information for understanding the binding sites and for further selective drug design against SARS and other coronaviruses.

2.2.2.19 Protein Kinase

Epidermal Growth Factor Receptor (EGFR) EGFR is a member of the protein tyrosine kinase family and is frequently overexpressed in cancer cells, correlating with poor prognosis and survival (Normanno et al., 2005). EGFR therefore represents a rational target for the development of novel anticancer therapies. Inhibition of the EGFR tyrosine kinase activity by small molecules has proved to be effective for the treatment of cancer (Shawver et al., 2002). Cavasotto et al. (2006) used the X-ray crystal structure of EGFR (Stamos et al., 2002) bound to OSI-744 (PDB code 1M17) to screen a collection of 315,102 compounds of the ChemBridge Express Library. The conserved water molecule present within the binding site was included in the description of the binding site because of its observed strong contacts with the receptor. The ICM virtual-screening program (2005) was used in this study. The ICM-based VS procedure is based on fast docking of a flexible ligand to a grid representation of the receptor (Totrov and Abagyan, 2001) followed by scoring. The ICM score takes into account the ligand–receptor interaction energy, conformational strain energy of the ligand, conformational entropy loss, and desolvation effects. Self-docking of cocrystallized ligand OSI-744 showed a score of −35. The authors thus selected a cutoff score of −28 to take into account limitations stemming from the rigid representation of the receptor and inaccuracies in the scoring function. There is also strong evidence that kinase inhibitors form two hydrogen bonds to the hinge-region backbone atoms (corresponding in EGFR to carboxylic O in Q767, and carboxylic O and HN in M769). For this reason, 1246 compounds (from a total of 315,102) scoring better than −28 and forming at least two hydrogen bonds with the hinge-region backbone were clustered according to chemical similarity and visual inspection, resulting in a final list of 212 compounds. The 50 best-scoring molecules were evaluated for their effect on EGFR tyrosine kinase activity and their antiproliferative activity in cancer cells. The authors identified several compounds with antiproliferative effects on cancer cells. Amongst them, MSK-039 was discovered as a low-
Figure 2.3  Structural model of MSK-039 docked at the EGFR kinase active site (only selected residues are shown). Dashed lines represent hydrogen bonds between ligand and active site residues. See color plates.

Figure 2.4  Structural model of a representative hit docked at the Chk-1 kinase active site (only selected residues are shown). Dashed lines represent hydrogen bonds between ligand and active site residues. See color plates.
micromolar inhibitor of EGFR tyrosine kinase activity. The predicted binding mode of MSK-039 is shown in Fig. 2.3, and its structure is shown in Table 2.1.

Checkpoint Kinase-1 (Chk-1 Kinase)  In recent years, DNA damage checkpoint kinases have emerged as attractive targets for cancer therapy (Zhou and Elledge, 2000). The checkpoint kinase-1 (Chk-1 kinase) prevents cells with damaged DNA from entering mitosis by arresting them at the G2/M checkpoint. It has been suggested that inhibition of Chk-1 kinase would lead to abrogation of the G2/M checkpoint, allowing cells to enter a lethal mitosis with damaged DNA. Thus there is a great interest in developing inhibitors of Chk-1 kinase for use as potential sensitizing agents in combination with standard therapies that induce DNA damage. Lyne et al., (2004) carried out a virtual screen of the AstraZeneca compound collection for the identification of compounds likely to bind checkpoint kinase-1 (Chk-1 kinase), using a receptor-based strategy. In the first step, the compound collection was filtered by application of generic physical properties, followed by removal of compounds with undesirable chemical functionality. Subsequently, a 3D pharmacophore screen for compounds with kinase binding motifs was applied. A filtered database of approximately 200,000 compounds was then docked into the active site of Chk-1 kinase using the FlexX-Pharm (Hindle et al., 2002) program. For each compound that docked successfully into the binding site, up to 100 poses were saved. These poses were then postfiltered using a customized consensus scoring scheme for a kinase, followed by visual inspection of a selection of the docked compounds. This resulted in 103 compounds being ordered for testing in the project assay, and 36 of these (corresponding to four chemical classes) were found to inhibit the enzyme in a dose–response fashion with IC\textsubscript{50} values ranging from 110 nM to 68 \textmu M. The binding mode of the best hit from this study is shown in Fig. 2.4, and its structure is shown in Table 2.1.

2.3 CONCLUSIONS AND PROSPECTS

Several clear successes have been demonstrated over the past few years in the field of computer-aided virtual screening for lead identification using receptor-based, ligand-based, pharmacophore-based, or similarity-based approaches. Lead compounds have been found for a wide array of biological targets such as enzymes, ligand binding domains, and even small molecules disrupting protein–protein interactions. There is no general method to carry out these studies. Every system is unique, and what has led to these successes is an in-depth understanding of the system under investigation and fine-tuning of the virtual-screening scheme to achieve the desired result. In most cases, this was done by applying all available information to generate and validate the models, which included seeding the database to be screened with known actives and then defining the workflow or the screening protocol so as to identify the known ligands as high-ranking hits. RBVS requires database filtering, target preparation, docking, and postdocking analysis. Appropriate filters can significantly reduce the number of candidates for screening by eliminating structures with undesirable properties. Some of the key challenges in VS are appropriate treatment of ionization
and tautomerization of ligand and protein residues, treatment of target/ligand flexibility, choice of force fields in docking calculations, solvation effects, dielectric constants, exploration of multiple binding modes, and, most importantly, the approximations in the scoring functions that can lead to false positives and miss true hits. In spite of the addressed limitations in this approach, it is still the best option available currently to explore a large chemical space in terms of cost-effectiveness, time commitment, and material as it allows access to a large number of possible ligands, most of which may be simply purchased and tested. Despite the challenges, this method has provided leads in several cases, and the most important feature of this approach is its ability to find novel chemical entities (NCE) for targets, which is very valuable from the medicinal chemistry perspective. With an increasing number of targets being identified by HTP genomics and proteomics efforts, virtual screening may provide an excellent complementary approach to HTP screening for valuable leads.

ACKNOWLEDGMENTS

We would like to thank members of the Huang laboratory and collaborators at Raylight Corporation and Chemokine Pharmaceutical Inc. for helpful discussion. This work was supported by grants from the National Institutes of Health, Raylight Corporation, and Chemokine Pharmaceutical Inc.

REFERENCES


ICM (2005). La Jolla, CA: Molsoft LLC.


REFERENCES


3.1 INTRODUCTION

The goal of screening small molecules for drug discovery is to deliver new hit compounds to medicinal chemists that can act as starting points for the development of drug candidates. Computational chemistry and molecular modeling provide tools that are commonly used to direct and increase the efficiency of laboratory screening by selecting or designing compounds to be tested (Bajorath, 2002; Jorgensen, 2004; Walters and Namchuk, 2003). This is driven by the fact that the number of compounds available for screening usually far exceeds the number that will actually go into the screen. Similar methods are then applied in the hit or lead optimization process; however, this review will mainly focus on the methodologies that are applied to computer-based, in silico, or “virtual” screening in the early stages of drug discovery.

The precise definition of a “hit” or “lead” molecule varies from one organization to another. At Serono, a compound from screening is regarded as a hit if it has favorable properties in the following respects: (i) potency in a biological or biochemical assay, (ii) novelty with regard to intellectual property, (iii) selectivity for the intended target, and (iv) tractability in terms of ease of synthesis. A lead has the properties of a hit plus a

*Corresponding author: xuliang.jiang@serono.com

Drug Discovery Research: New Frontiers in the Post-Genomic Era, Edited by Ziwei Huang
Copyright © 2007 John Wiley & Sons, Inc.
favorable ADME/Tox profile (absorption, distribution, metabolism, excretion, and toxicity) and proven *in vivo* activity in an animal model. Therefore, many factors need to be considered at the virtual screening stage to put together a compound set that is most likely to deliver multiple hit/lead candidates. Clearly, activity against the target of interest is an absolute requirement. Computational methods attempt to deliver the most active compounds possible, but usually a positive from high-throughput screening (HTS) in the 1–10 \( \mu \text{M} \) range is acceptable, with the compound then being optimized by medicinal chemistry. Computational methods can be applied to deliver compounds that are different from what is already known and provide the chemists with a range of compounds from different chemical classes, thereby giving the opportunity to accept or reject hits based on other considerations such as chemical tractability, intellectual property, and toxicity. Thus, a desirable property of a computational method is to be able to “jump” classes—that is, use the information derived from one class to correctly predict compounds from another chemical class. The results then also give a more useful structure–activity relationship (SAR) for further computational analysis. In a large-scale screening campaign it is also desirable to include a nontargeted set of diverse compounds. The advantage of screening a diversity set is that the approach does not make assumptions about must-have properties or substructures in potential actives and thus invites discovery of unexpected active chemical types or modes of action. There is no universal agreement as to how much computational effort should be expended into predicting other druglike properties at the screening stage, such as those associated with ADME/Tox, but most agree that at least a rough measure should be applied because many compounds, especially from sources such as combinatorial libraries and natural products, have been found to violate the most basic requirements for druglikeness. Filters for clearly undesirable compounds are also usually applied.

Virtual screening originated in the 1970s when compound database searches were introduced using two-dimensional (2D) structural fragments (Blake et al., 1977; Dromey, 1978; Feldmann et al., 1977; Feldmann and Heller, 1972; Saxberg et al., 1976). Subsequently, a wide variety of diverse methodologies have been introduced, and the field is still rapidly evolving. These methods can be usefully divided into two categories: (1) those that predict activity for a particular target and (2) those computations not related to a target. While it is not possible in this review to describe all the methods, the following two sections will describe some of the major approaches in each category.

### 3.2 TARGET-DEPENDENT METHODS

This category relies on having some knowledge of the biological target or compounds that bind to the target. Therefore, this category can be further subdivided into those that are ligand-based and those that are receptor (protein)-based. When a set of active ligands is available, the methods that can be applied are those of QSAR, substructure analysis, pharmacophore analysis, or simpler methods such as using a measure of molecular similarity to query compound libraries for potential active compounds. When a crystal structure of the protein target is available (or a homology model can be built based on a closely related protein), then docking methods can be employed.
3.2.1 Molecular Similarity Searching

Similar chemical structures tend to have similar biological activities (Fig. 3.1), although exceptions to this generalization are frequent, because it is often observed that a small change in chemical structure, sometimes involving a single atom or functional group, can have a profound affect on activity for a particular target (Martin et al., 2003; Matter, 1997; Schuffenhauer et al., 2003). Conversely, molecules with seemingly unrelated structures (at least when looked at on paper in a 2D representation) can exhibit similar activities (Bajorath, 2002). In spite of these shortcomings, the hit rate from similarity-based searches is higher than that from random screening (Martin et al., 2003).

Small molecules are usually represented (both on paper and in chemical databases) as atoms connected by bonds, and as such they are difficult to compare computationally in this form. To overcome this difficulty, several other representations have been devised to facilitate quantitative molecular comparisons. The most common representation is to encode chemical features in a fixed-length binary bitstring, referred to as a “molecular fingerprint.” Examples include the MACCS keys (MDL Information Systems, Inc.), the UNITY fingerprint (Tripos, Inc.), and the Daylight fingerprint (Daylight Chemical Information Systems, Inc.). The length of these keys vary from hundreds to thousands or more, and each bit encodes the presence or absence of a chemical feature, which is usually an atom type, a functional group, or a connected fragment. The similarity of the fingerprints of two molecules is therefore a measure of overall chemical similarity. Several metrics have been proposed to compare bitstrings (Willett et al., 1998), each with its own advantages and disadvantages, but probably the most widely used is the Tanimoto coefficient, defined as

\[ T = \frac{N_{ab}}{N_a + N_b - N_{ab}} \]

where \( N_a \) and \( N_b \) are the number of bits set in bitstrings \( a \) and \( b \) respectively, and \( N_{ab} \) is the number of bits set in both bitstrings (i.e., applying the AND operator). The value varies

Aventis, US930720

Pfizer, EP19900418

Merrell Dow, EP303961

Lily, Prozac

Lily, US19910226

**Figure 3.1** The chemical structures of a number of 5-HT reuptake inhibitors showing clear similarities in structure, illustrating the basis for the similarity and substructure methods described in the text. The common substructures are highlighted. See color plates.
between 0.0, where two bitstrings have no bits in common, and 1.0, where they are identical. This method discriminates well for most compounds in the molecular weight (MW) range of 200–600.

Many attempts have been conducted to improve the performance of this approach. Salim et al. (2003) showed that a combination of different similarity coefficients can improve similarity searching performance, although they failed to identify a single optimum combination for all activity classes. Another improvement for similarity searching is fingerprint scaling, whereby the contributions of each bit to the Tanimoto coefficient are scaled up if they also occur in a consensus fingerprint derived from a set of molecules active for a target. Xue et al. (2001, 2003, 2004) demonstrated that this method can improve the hit rate for most biological active classes. Some research groups have incorporated reduced molecular graphs into the approach, in which molecule structures are represented by topological graphs, where each node represents a substructure feature such as a ring or functional group. With this simplified structure representation, they found the method was more effective and could find more diverse compounds (Gillet et al., 2003; Harper et al., 2004).

Similarity searches using chemical fingerprints based on 2D structures will most likely pull out only analogs of known compounds. Although this is useful for generating SAR around a particular scaffold and producing “me too” compounds, the novelty of selected compounds is often an issue. Similarity searches based on fingerprints that incorporate three-dimensional (3D) molecular information are more able, in principle, to generalize to new scaffolds. McGregor and Muskal (1999) showed improved performance over 2D methods when the information of 3D pharmacophore fingerprints was used, but earlier studies found that 2D descriptors are equally or even more effective than 3D descriptors in selecting representative subsets of bioactive compounds (Brown and Martin, 1996; Matter, 1997).

The scope of fingerprints in similarity searches can also be extended beyond single targets to protein families, based on the observation that ligands for different targets in the same protein family are often seen to contain similar features or property ranges (Deng et al., 2004; Schuffenhauer et al., 2003). Therefore, the known active chemical structures for a whole target family can be grouped together and used in similarity searching; the resulting screening set then has an enlarged chemical space, and the same set can be screened repeatedly against targets of the same class.

### 3.2.2. Privileged Substructure Search

This approach is based on the observation that specific structural motifs in small molecule ligands often complement features in a target protein or protein family (Bondensgaard et al., 2004; Muller, 2003). These features may be common to the ligands of different members of a protein family (this is often the case with kinases). Sometimes a particular motif is a necessary requirement for optimal binding for a particular target—for example, the phenol group in estrogen receptor ligands (Anstead et al., 1997; Minutolo et al., 2003). By comparing structures of actives against inactives, these recurrent substructures can be identified, and a screening set is then selected by substructure search based on the identified fragments.
Many such features have been discovered by eye, either by examining sets of active compounds or by looking at ligand–receptor interactions in crystal structures. However, when data sets become larger and/or the features become more subtle, an automated procedure is desired. Examples of such procedures include an early artificial intelligence approach (Klopman, 1984), structural commonality analysis (Lewell et al., 1998), and a retrosynthetic analysis (Hajduk et al., 2000; Shemetulskis et al., 1996). Another example of such a method is discrete substructure analysis (DSA), developed by Serono Pharmaceutical Research Institute (Merlot et al., 2002, 2003). This method automatically enumerates all possible substructures up to a given size (typically 10–16 atoms) in a set of small molecules in 2D representation (the molecule connection table). Given a training set of actives and inactives, DSA calculates statistics that evaluate the occurrence of each fragment in the active set relative to the inactive set. Using the DSA Visualizer, the resulting statistics can be viewed interactively on a computer screen (Fig. 3.2). Each compound is displayed as a point on a 2D plot, and the compound structure can be displayed by clicking on the point. Thousands of fragments can be displayed and explored this way. Probably the most instructive statistics to plot are the percent occurrence in the active set against the Fisher log score—a measure of the significance of the occurrence of a fragment in the active set compared to the inactive set. Clearly, the most discriminating fragments are the ones that have a highly significant Fisher log score (i.e., enriched in the active

![Figure 3.2](image_url)

**Figure 3.2** The DSA Visualizer for exploring substructure fragments in a set of compounds active for a particular target. This result was generated for a set of JNK kinase ligands. Each point is a fragment, which, when selected, is illustrated to the left. The horizontal axis is the number of compounds in the active set that contain the fragment. The vertical axis is the Fisher Log score—a measure of the significance of the occurrence of the fragment in the active set relative to an inactive set. The fragments are colored by size (number of atoms). The most significant fragments appear to the top right of the plot. The structures corresponding to the picked (highlighted) points are illustrated on the left. See color plates.
set) and also occur in a large percentage of actives. Fragments that are highly significant but of low occurrence are unlikely to pull out many actives in a test set (i.e., a prediction with many false negatives), whereas fragments that are of high occurrence but low significance are likely to pull out many inactives (false positives) in a test set (e.g., a benzene ring usually occurs as one of these fragments). There is still an element of subjectivity in selecting a small number of substructures to use to search a set of compounds for screening. Future work might include the ability to automatically generalize substructures by combining features of two or more of them, which presently is done manually.

The advantages of the DSA method are that it can rapidly be applied to large compound sets and many targets and that the results are in a chemically intuitive form. This method has been applied to the activity classes in the MDL drug data report (MDDR) and has been extended to predict toxicological outcomes. A disadvantage is a limited ability to cross chemical classes. Thus, if a substructure is considered to define a chemical class, then searching with this substructure cannot cross to another class. Also, the current method overlooks the potential requirement of multiple substructures instead of a single fragment in some targets. Nevertheless, the DSA method has been proven to be effective in the identification of novel hits and leads for several drug targets at Serono.

3.2.3. QSAR and Recursive Partitioning

Quantitative structure–activity relationship (QSAR) methods were some of the earliest developed to predict drug activity, where a relationship was found between drug activity and hydrophobicity as measured by $\log P$ (Hansch et al., 1986). The same methods have also been used to predict other properties, such as those associated with ADME/Tox, where the term used is quantitative structure–property relationships (QSPR) (Grover et al., 2000a, 2000b). The essence of all the methods is to derive a relationship between a set of molecular descriptors and a biological activity (or property) for a set of known molecules, giving a mathematical model that can then be used to predict untested compounds. A wide variety of molecular descriptors (Todeschini and Consonni, 2000) have been employed, derived from 2D and 3D representations of small molecules. Examples are topological state (Kier and Hall, 1989), E-state (Kier and Hall, 1990), binary fingerprints (McGregor and Muskal, 1999), and 3D molecular fields (Cramer et al., 1988; Cruciani et al., 2000).

An equally wide variety of methods of analysis has been used, derived mostly from the fields of applied statistics and machine learning. The most basic is linear regression, sometimes implemented as principal components regression (PCR) or partial least squares (PLS); the latter two methods are appropriate for dealing with the case in which there are more variables than compounds and/or when many variables are highly correlated, which often happens with molecular descriptors. These techniques work best when both the dependent and independent variables take on continuous values, are associated with small errors, and approximate to some typical kind of statistical distribution, usually Gaussian. However, it is not unusual to have molecular descriptors and activity data that are discontinuous, contain large errors, and/or do not belong to a typical distribution function. It is then more reasonable to
classify compounds into categories such as active/inactive or high/medium/low. This would be the case when analyzing HTS results, which are usually of low quality. Examples of methods more suited to such data are discriminant function analysis, neural networks, genetic algorithms, and recursive partitioning (RP). As an example, we will describe in more detail the methodology of RP.

RP is a statistical method that constructs a decision tree to distinguish different classes of data (Chen et al., 1998) (Fig. 3.3). In its application in a virtual screening process, RP identifies relationships between specific chemical structure features (e.g., continuous descriptors, binary fingerprints, and/or substructure fragments) and biological activities. The types of biological activities can be categorical data (e.g., an activity class such as a “protein tyrosine kinase inhibitor”) or numerical data (e.g., percentage inhibition data from HTS). The first step in the RP process is to generate descriptors or substructures for each compound in the collection with prior knowledge of biological activity. Numerical activities have to be converted into categories (e.g., good and bad or high, medium, and low). The RP algorithm evaluates each descriptor sequentially to determine the best descriptor at each node that separates actives from inactives, using a statistical test. As the RP decision tree grows, the whole data set is progressively split into smaller subsets until all compounds in the subsets are grouped into a homogeneous class or a parameter reaches its limit (e.g., the minimum number of compounds in each subset, or the maximum number of nodes). The descriptors associated with the detailed path by

![Figure 3.3](image-url)

**Figure 3.3** Example of a hypothetical recursive-partitioning tree. The nonterminal nodes are shown in black squares, and the terminal nodes are shown in colored squares. The reasons for the terminations are as follows: A homogeneous class is found (blue), the number of compounds is not sufficient (yellow; the minimum here is 25), and the maximum number of nodes is reached (red: the upper limit in this example is 15). The descriptors can be substructures or numerical properties (\(N_{\text{rotbonds}}\) is the number of rotatable bonds, and \(O_{\text{mN}}\) is the number of oxygen atoms minus nitrogen atoms). The statistical test here is simply the percentage of positives (Pos%). See color plates.
which the molecules are split give an indication about their contribution to the activity or inactivity of the molecules in the subset. A major advantage of the RP approach, as compared to typical QSAR approaches, is that it does not assume a particular type of distribution in the activity data or a specific type of function that relates activity to the descriptors and therefore accommodates multiple binding modes of actives or multiple reasons for inactivity (Hawkins et al., 1997). This is important because the data from prior HTS or multiple sources (e.g., MDDR, literature, or patents) are usually not homogeneous or accurate enough to yield a reliable parametric regression equation. Like other QSAR approaches, cross-validation is an important quality-control step in the RP method to optimize the parameters, such as the minimum number of observations per subset and the maximum depth of a tree, before a final model is built for use in a virtual screen. One of the cross-validation methods is leave-N-out. This method splits data into N groups, builds a model from all but one group, and then predicts the activity of the left-out group. The process is repeated for each group, and the optimal parameter setting is then determined by comparing the overall quality of the predictions from all runs.

### 3.2.4 Pharmacophore Analysis

Another commonly used computational technique when a set of active ligands is known is that of pharmacophore analysis. In this context a pharmacophore is regarded as a set of atom (or atom-group) types (features) in a 3D spatial arrangement that represent the interactions made in common by a set of small molecule ligands with a protein receptor site (Langer and Wolber, 2004). The types of pharmacophore features used are based on the kinds of interactions seen in crystal structures of ligand–protein complexes. These usually consist of hydrogen bonding groups (H-bond donor and acceptor), groups with formal charges (positive/basic and negative/acidic), and hydrophobic centers. Other features may also be included. Often an aromatic group is used, and sometimes others, such as a combination H-bond donor/acceptor, are used. These are assigned to atoms or atom groups in a small molecule based on structural context. For example, a carboxylic acid is negatively charged at neutral pH, so this would be assigned the negative charge type; aliphatic amines are examples of a positively charged group; oxygen or nitrogen atoms with an attached hydrogen are H-bond donors, ones without an attached hydrogen are H-bond acceptors, and so on; and hydrophobic types are usually carbon atoms, or groups of them, removed from polar or charged groups. In this way, a small molecule is represented as a set of potential interactions with a protein, and the underlying chemical class or scaffold is unimportant in itself, only serving to present these kinds of functionality in 3D space. Thus one of the claimed advantages of this approach is that it can discover new chemical classes that bind to the same receptor.

All the methods start with a set of compounds that are known, or assumed, to bind to the same protein receptor site in a similar way (i.e., using a number of common pharmacophore interactions), even if they share little 2D chemical similarity. It is the aim of this method to calculate what these interactions are and output them as a set of pharmacophore points in a 3D spatial arrangement. This pharmacophore can then be
used to computationally screen compounds of unknown activity on the assumption that ones that fit the pharmacophore are more likely to be active for that receptor.

Probably the best-known and most widely used program for pharmacophore analysis is Catalyst, available from Accelrys, Inc. (Barnum et al., 1996; Greene et al., 1994; Sprague, 1995). An algorithm for generating a set of 3D structures for a small molecule that optimally covers conformational space was a central part of the development of the methodology, as was an algorithm for deriving pharmacophores from a training set of molecules, termed “hypotheses” in Catalyst terminology. Several hypotheses are generated with relevant statistics, allowing the user some flexibility to choose which one(s) to use, based on other knowledge such as that derived from the crystal structure of the target. Input parameters can be tuned to control such things as the number of points to be generated. Molecule conformations with pharmacophore assignments for large numbers of compounds are stored in a Catalyst database that is then used in a search to fit each molecule to a hypothesis. Statistics are generated, and the results can be visualized. A recent development has been the ability to apply an excluded volume fit to a Catalyst search. This approach recognizes that ligand binding depends both on satisfying individual specific interactions and on overall shape complementarity with a binding site. Compounds too big to fit into the binding site, even though they contain the desired pharmacophore, are excluded.

Another more recent development is pharmacophore fingerprinting (Mason et al., 1999, 2001; McGregor and Muskal, 1999, 2000; Pickett et al., 1996). In this method, a basis set of three- or four-point pharmacophores is enumerated with a set of distance ranges. For example, the PharmPrint method (McGregor and Muskal, 1999, 2000) uses six pharmacophore types (H-bond donor/acceptor, positive/negative charge, hydrophobic, and aromatic) and six distance ranges (2–4.5, 4.5–7, 7–10, 10–14, 14–19, and 19–24 Å) (Fig. 3.4). When all possible combinations of three-point pharmacophores are enumerated, this gives a basis set of 10,549. Each pharmacophore is then either present or absent in a small-molecule 3D structure. Thus the fingerprint for the molecule can be represented as a binary bitstring. Fingerprints for different conformations can be either kept separate or combined. The pharmacophore fingerprints are then amenable to the same types of calculations (e.g., the Tanimoto coefficient) that are applied to bitstrings derived in other ways. QSAR models have been built using the pharmacophore fingerprint as a descriptor (McGregor and Muskal, 1999), and the mapping of biologically active space has been explored (McGregor and Muskal, 2000). It is also possible to calculate the fingerprint of a protein active site (Arnold et al., 2004; Mason et al., 1999). For example, if there is an H-bond donor on the protein in the binding site, then a ligand is favored if it contains an H-bond acceptor in a position that is a certain distance and direction relative to it. In this way, a fingerprint of the binding site can be built, and comparison to a small molecule fingerprint gives a measure of the goodness-of-fit in the site.

3.2.5 Docking

When the crystal structure of the target is available, there is an opportunity to employ docking methodologies, whereby in a computational simulation ligands are oriented
in the active site of a protein and a measure of the goodness-of-fit is calculated, thereby explicitly implementing the key–lock principle of ligand–protein interactions (Fig. 3.5). Unlike ligand-based approaches, a target-based method, such as docking, does not depend on information from known ligands, so compounds selected by this method are not necessarily biased toward known chemotypes, providing an

Figure 3.4 Pharmacophore fingerprinting: (a) The definition of a three-point pharmacophore, consisting of three pharmacophore points separated by three distance ranges; (b) a single three-point pharmacophore mapped to the estradiol molecule; and (c) the fingerprints that record the occurrence of every possible pharmacophore in a set of molecules.

Figure 3.5 An example of molecular docking using the Glide program. This result was produced by docking a set of ligands to the crystal structure of JNK kinase (PDB code 1PMN). Several different ligands are overlapped, and common interactions can be seen to emerge (for example, the hydrogen bond to the hinge region, shown as a yellow dotted line), even though there is variation in the contacts made by the small-molecule ligands. See color plates.
opportunity to select or generate compounds that are truly novel. In addition, docking results can provide mechanistic explanations of ligand–receptor interactions at the atomic level, offering an excellent opportunity to then optimize hit or lead molecules based on a “rational design” strategy. In principle, one might expect docking to be the complete solution to the design of small-molecule ligands. In practice, the method is limited in several fundamental respects, but these limitations have not prevented its widespread use, both for selecting screening sets and for hit or lead optimization. Many commercial and academic software packages for docking are available, such as AUTODOCK (Morris et al., 2004), DOCK (Ewing et al., 2001; Kuntz et al., 1982), FlexX (Rarey et al., 1996), GOLD (Jones et al., 1997), Glide (Friesner et al., 2004; Halgren et al., 2004), and ICM (Cavasotto and Abagyan 2004; Stigler et al., 1999). Among them, Glide and GOLD are highly regarded in the pharmaceutical industry (Bissantz et al., 2000; Kontoyianni et al., 2004; Perola et al., 2004). A requirement for docking is the availability of a high-resolution crystal structure of the target of interest, preferably with bound ligand, although docking to good-quality homology models can still produce positive results (Diller and Li, 2003; Oshiro et al., 2004).

Most docking algorithms proceed in two steps: (1) the orientation of a particular conformation of a ligand in the protein binding site, often referred to as the ligand “pose”; and (2) the calculation of a score that correctly ranks different poses of the same molecule and the best pose of different molecules. Small molecules in a bound state in a protein active site have been often found to adopt a conformation different from that at the energy minimum of the unbound state, either in solution or in vacuo (Perola and Charifson, 2004). As a result, a docking algorithm needs to generate multiple conformations of the ligand to achieve the first step. It is, however, a challenge to sample the vast conformational space with reasonable accuracy and speed for a high number of rotatable bonds. Some programs, such as Glide, sacrifice speed for a more thorough search of conformational space, while others, such as FlexX, attempt to avoid the need to generate many conformations of the entire ligand by breaking it into fragments, orienting the fragments, and then joining them back together.

An area of critical importance is the scoring function, which calculates the interaction energy (and/or other terms contributing to the goodness-of-fit) between the ligand and receptor. Analysis of docking results from various studies indicate that it is largely the issue of inaccurate scoring functions that give rise to high false positives and/or false negatives for many targets (Bissantz et al., 2000; Kontoyianni et al., 2004; Stahl and Rarey, 2001). Most early implementations used scoring functions based on the nonbonded term of the molecular mechanics force field, incorporating a Lennard-Jones potential for van der Waal interactions and a Coulomb term for electrostatics and sometimes including a supplemental hydrogen bond energy term (Kitchen et al., 2004). Examples of force-field-based scoring functions are DOCK (Kuntz et al., 1982), Goldscore (Jones et al., 1997) and AutoDock (Morris et al., 2004). The drawback here is that the molecular mechanics force field only accounts for the potential energy of interaction between ligand and receptor, whereas the factor that governs the binding affinity is the difference in the Gibbs free energy of the whole system for bound and unbound states, which includes additional solvent and entropy effects. Accurate calculation of the latter terms is intractable in the
context of high-throughput docking. This shortcoming has led to the addition of terms designed to take these effects into account and the development of empirical and knowledge-based scoring functions. The empirical scoring functions are usually the summations of various Gibbs energy changes parameterized by empirical values obtained from fitting the theoretical energy equations to experimental affinity data. LUDI (Bohm, 1994), FlexxScore (Rarey et al., 1996), ChemScore (Eldridge et al., 1997), and GlideScore (Friesner et al., 2004) belong to this type. The characteristics of knowledge-based scoring functions are that their potential energies are parameterized by the geometric distributions of atoms observed in known X-ray crystal structures: PMF (Muegge and Martin, 1999), SMoG (Dewitte et al., 1997), and DrugScore (Gohlke et al., 2000) are examples. These different scoring functions have been extensively evaluated (Perez and Ortiz, 2001; Stahl and Rarey, 2001; Terp et al., 2004; Wang and Wang, 2003), but no clear agreement has been reached as to which of the three types perform consistently better than the others. This situation has led some groups to implement a consensus scoring approach, in which the scores of several little-correlated scoring functions are combined, with reported enhancement of hit rates (Charifson et al., 1999; Jacobsson et al., 2003).

Most scoring functions have difficulty dealing with small steric clashes between ligand and receptor. A steric clash causes a large increase in interaction energy, which can lead to false-negative scores (Erickson et al., 2004). Steric clashes can be a result of errors in the conformation or pose of the ligand, or they can reflect a genuine incompatibility between ligand and receptor. Often steric clashes can be relieved by a change in the conformation of the receptor to accommodate the ligand, giving rise to the “induced fit” problem. If the change is major, involving movement of the main chain, then it is clearly unrealistic to expect current docking algorithms to deal with this problem. However, minor changes in side-chain conformation are very frequent, as seen when comparing crystal structures of the same protein with different ligands. Although soft docking techniques attempt to alleviate the problem of side-chain conformational changes (Jiang and Kim, 1991; Fernandez-Recio et al., 2002), it is difficult to devise a docking algorithm that can distinguish between the possible sources of steric clashes. An additional problem is that many ligands interact with their receptors through water-mediated hydrogen bonds, and it is also difficult for a docking algorithm to reproduce these kinds of interactions (Davis et al., 2003).

Some of these problems can be overcome with careful setup of the docking run by a user with knowledge of the specific target. For example, if it is known that there is a requirement for a bound water in the active site, this can be included as part of the receptor in the dock run. If it is seen that there is always a specific interaction or pharmacophore that occurs, this can be forced on the ligands in the dock run, either as part of the run if the software allows or in postprocessing of the results (Fradera et al., 2000; Good et al., 2003; Hindle et al., 2002; Verdonk et al., 2004). Kinase targets provide a good example, where there are known important interactions, such as a hydrogen bond to a main-chain NH group in the hinge region (Fig. 3.5) (Bossemeyer et al., 1993; Chuaqui et al., 2005; Toledo et al., 1999; Tong et al., 1997). If the ligands contain a privileged substructure known to bind in a certain place, then the docked poses can be filtered for RMS deviation to this substructure. Some commercially available docking packages, such as FlexX-Pharm (Hindle et al., 2002), GOLD, and
Glide, allow for the inclusion of these pharmacophore-type constraints in the docking stage. This greatly cuts down the search space of poses that needs to be covered. In an effort to overcome the induced fit problem, if multiple crystal structures are available, they can be used in parallel, and the final score can be either the best score (assuming this is the closest fit) or an average of all scores, possibly weighted by relevance of structure. Alternatively, a protein structure can be energy-minimized around an important ligand, if no crystal structure is available for that particular ligand, to bias the receptor toward ligands of that type. Moreover, scoring functions are known to scale poorly with some simple molecular descriptors, such as MW and the number of rotatable bonds. In our laboratory we have found value in normalizing docking scores for these kinds of descriptors in a postprocessing stage.

Despite the problems mentioned above, we anticipate that docking will continue to have a central role in computer-based drug discovery. Some of the problems are alleviated by increases in computer processing power and parallel computing, and the rapid increase in the number of experimentally determined crystal structures of proteins will extend this technology to a large proportion of pharmaceutical targets, such as G-protein-coupled receptors (GPCR) and ion channels, where docking technology currently depends on homology models with limited quality (Bissantz et al., 2003; Evers and Klabunde 2005; Zhorov et al., 2001).

3.3 TARGET-INDEPENDENT METHODS

The compounds selected from the second category of virtual screening methods are not dependent on the knowledge of any specific targets or their ligands, and this includes a wider set of considerations. One consideration is chemical diversity or coverage of biologically active space, the aim of which is to discover ligands from novel compound classes and maximize the chances of hitting a target for which there is little known information. Applying a measure of diversity also minimizes repetitive screening of the same or similar compounds. Another consideration is the calculation of ADME/Tox properties and the elimination of clearly undesirable compounds. The aim here is to screen molecules for properties that characterize a compound as druglike. These properties are independent of an individual target but may vary, depending on site of action or therapeutic area; for example, drugs acting on the central nervous system need to cross the blood–brain barrier and therefore need to have specialized permeability properties. Some of the methodologies may be common to the target-dependent and target-independent applications. For example, some of the same descriptors can be used in calculations of molecular similarity/diversity, ADME/Tox, and QSAR/QSPR.

3.3.1 Diversity Analysis

Screening a diversity set is relevant when there is no information available about a target structure or its ligands or when a compound collection is to be screened against many or unknown targets (such as in cell-based HTS); it also provides the opportunity to discover unexpected new chemotypes or modes of action. The concept of
molecular diversity is based on the idea that by using calculated properties, small molecules can be meaningfully mapped to a property space, where compounds with similar biological activities are close in the space and compounds with different activities are far apart (Willett, 2000). Ideally, the activity for a target will be confined to a limited continuous region of this space. If this region is unknown (e.g., if there is no information about a target or the same compound set is to be screened against different targets), then the best chance of hitting the target is with a set that covers the space as comprehensively as possible. The property space can be produced by calculating a set of numerical molecular descriptors, in which case the dimensionality of the space is the number of descriptors used. The dimensionality can be reduced, if needed, by techniques such as principal components analysis (PCA), enabling visualization in 2D on paper or in 3D on an interactive computer screen. Alternatively, the molecular similarity methods described above can be used to give a measure of the distance between molecules in property space, and algorithms for diversity analysis can operate on these distances directly (Trepalin et al., 2002; Willett, 2000).

Diversity analysis involves selecting a subset of compounds that optimally covers the chemical space of an entire set, which can be a corporate collection, commercially available compounds, a combinatorial library, a virtual library, or a combination. Many algorithms have been devised for this purpose (Blake, 2004; McGregor and Pallai, 1997; Sauer and Schwarz, 2003a, 2003b; Trepalin et al., 2002; Young and Ge, 2004), and Fig. 3.6 illustrates three of them. Some algorithms aim to select a subset of compounds that are representative of the entire set, while others aim to cover the underlying property space as widely as possible, regardless of the distribution of the entire compound set in this space. Clustering is based on the observation that compounds are not usually evenly distributed in a property space but group together in certain regions, similar to galaxies in outer space. If clusters can be identified within which compounds have similar biological activities, then only a single representative (or a small number) of each need be chosen for screening, usually the calculated “centroid.” Clustering has been applied to many kinds of

**Figure 3.6** Schematic illustration of three different diversity methods. (a) **Clustering:** The compound closest to the centroid of each cluster is selected to represent the whole cluster. Some clustering algorithms have difficulty dealing with unevenly distributed clusters, and large clusters can be underrepresented. (b) **Maximum dissimilarity selection:** The first compound is randomly selected; in successive steps the chosen compound is the one with the greatest distance to its nearest neighbor in the chosen set. (c) **Maximum similarity elimination (or sphere exclusion):** A compound is randomly chosen, and then all compounds within a predefined similarity threshold are excluded from further analysis; the selection proceeds until no compounds are left. The latter two methods produce a more even sampling of property space but have a tendency to select outliers.
pattern-recognition problems, and there are well-known difficulties. For example, with small molecules, one problem is that clusters vary greatly in size, which gives rise to uneven sampling. Also, chemical space often gives rise to clusters with irregular shapes known to be difficult for most algorithms to deal with (Subhash, 1996). Another method, maximum dissimilarity selection (Snarey et al., 1997; Trepalin et al., 2002), proceeds by randomly choosing the first compound and then selecting the compound with the least similarity to it and adding it to the chosen set. Then, the next least similar compound to the existing set is then added until a certain number of compounds is reached. Maximum similarity elimination (also known as sphere exclusion) (Gobbi and Lee 2003; Taylor, 1995) proceeds by randomly selecting a compound and then excluding all the compounds that have a similarity higher than a given threshold from further analysis. This process continues until no compounds are left. This method ensures that no two selected compounds are more similar to each other than the threshold. Other methods aim to optimize a function of the chosen set. For example, for each compound in the chosen set, the MaxMin function (Snarey et al., 1997) is a summation of the distance to its nearest neighbor. A numerical algorithm needs to be applied to optimize this function.

It has been found that many diversity algorithms tend to be biased toward outliers in chemical space, and those unusual compounds are less likely to be drug-like. Thus, the search space should be limited to the regions that are populated by compounds with calculated favorable profiles. In the end, however, selecting a diversity set is a problem for which there is no objectively correct answer (Kubinyi, 1998), and it can be difficult to devise experiments in a practical setting to make an in-depth comparison of the methods. Several commercially available software packages for molecular modeling have modules designed for diversity analysis. Examples are DiverseSolution (DVS) from Tripos, Inc. and C2-Diversity from Accelrys, Inc. Both have methods for reducing a high-dimensional space to 3D for visualizing with computer graphics and provide a choice of algorithms for diverse subset selection.

3.3.2 ADME/Tox Filters

The majority of drug candidates fail in clinical trials due to ADME/Tox problems (in vivo absorption, distribution, metabolism, excretion, and toxicity) (FDA, 2004; Kennedy, 1997; Venkatesh and Lipper, 2000). These problems can be hard to resolve at an advanced stage of development, which has given rise to the desire to address these issues as early as possible in the drug-development cycle, either experimentally or computationally. However, in vivo ADME/Tox and pharmacokinetic outcomes are notoriously difficult to predict, even with in vitro experiments, due to the complex nature of the interaction of chemicals with biological systems. Due to the importance of the problem, much attention has been focused on it, and software is available for prediction of various ADME/Tox endpoints. However, many of these software packages are not suited for high-throughput treatment of compounds and are best used at the lead optimization stage. Additionally, many QSAR/QSPR models on ADME/Tox properties are found to be limited to certain chemical classes, and the goal of a global model that can be applied to all chemical classes proves elusive. Fortunately, many ADME/Tox endpoints are related to basic physicochemical
properties of compounds, some of which are feasible to treat computationally at the screening stage. Examination of profiles of these properties for drug compounds show clear preferred ranges that can then be used to classify a new molecule as druglike or otherwise. The druglike structures are then assumed to be more likely to possess favorable ADME/Tox properties, without having to address individual ADME/Tox endpoints directly (Frimurer et al., 2000; Muegge et al., 2001; Sadowski and Kubinyi, 1998; Walters and Murcko, 2002; Xu and Stevenson, 2000). Some of these properties are the same molecular descriptors used in the target-dependent methods, such as QSAR. Some physiochemical properties, if unfavorable, can pose problems even at the screening stage; for example, poor solubility is known to compromise screening results as well as cell-based and in vivo measurements. The training sets for identifying druglike property profiles can be commercially available databases of bioactive compounds, such as the Comprehensive Medicinal Chemistry (CMC) or MDDR databases from MDL or the World Drug Index (WDI) database from Thomson Scientific.

A widely used virtual screening filter for compounds intended to be orally available drugs is the so-called “rule of five,” derived from an analysis of property profiles of compounds that have survived preclinical stages and Phase I safety evaluation (Lipinski et al., 1997). Several computed properties were selected based on their clear relationship with oral bioavailability: MW, $clogP$, and counts of hydrogen bonding groups. The rule of five states that compounds are likely to have poor absorption or permeability properties if (1) there are more than five H-bond donors, (2) MW is over 500, (3) $clogP$ is greater than 5, or (4) there are more than 10 H-bond acceptors. Compounds outside these ranges for two or more descriptors are predicted to be problematic, but the converse is not true; that is, compliance with the rule is no guarantee of good bioavailability. By comparing the properties of early-stage molecules with that of final drugs, Oprea et al. (2001) concluded that the properties of leadlike molecules often need to have more restrictive values: MW less than 450, $clogP$ between $-3.5$ and $4.5$, no more than four rings, no more than 10 nonterminal single bonds, no more than five H-bond donors, and no more than eight H-bond acceptors. The descriptors are easy to calculate and are often stored for every compound in a database, and compounds that violate the rule can be flagged. In addition to the use of the rule as a negative filter in compound selection, profiles of these kinds of descriptors can be examined for whole compound collections and adjustment made to the content to reflect the distributions found in druglike sets.

There may be an inconsistency between the aim of an in vitro HTS assay to generate potent compounds and the need for drug discovery to deliver bioactive molecules with oral availability. Most HTS assays are in vitro molecular assays. With potency as a major criterion, compounds that are larger, more hydrophobic, and less soluble tend to stand out at the HTS stage due to their propensity to be more potent. This can be compounded at the post-HTS optimization stage if medicinal chemists seek to improve in vitro activity by adding hydrophobic groups to compounds (Kubinyi, 2003; Lipinski et al., 1997). However, there is a clear relationship between, for example, high MW and poor intestinal and blood–brain barrier permeability (Navia and Chaturvedi, 1996; Pardridge, 1995). This highlights the need to address druglike properties early on in the drug-development pipeline.
Similarly, the in vitro nature of HTS technology does not favor compounds with optimum toxicity profiles. Consequently, toxicity problems are usually detected late in drug development. However, toxicity is a problem with many physiological outcomes, so there is no simple rule to filter toxic compounds. Many approaches have been explored, mostly aimed at identifying molecular substructures and/or descriptor patterns responsible for specific toxicity outcomes (Cronin et al., 2003; Enslein and Craig, 1982; Henry et al., 1982; Hodes, 1981; Mattioni et al., 2003; Pearl et al., 2001; Yuta and Jurs, 1981). Merlot et al. (2003) have extended their DSA method to analyze toxicological properties for large compound sets such as the MDL Toxicity database.

Other filters are also desired in order to enhance success rate in screening. Many false positives from an HTS program arise from compounds with the presence of a reactive functional group (such as an epoxide or an aldehyde) that is likely to interact covalently with proteins (Rishton, 1997, 2003) or “promiscuous inhibitors” that inhibit targets via aggregate forms (McGovern et al., 2002, 2003). Several computational approaches have been developed to exclude these types of “nuisance” compounds (Roche et al., 2002; Walters and Namchuk, 2003).

3.4 INTELLECTUAL PROPERTY

In the pharmaceutical industry, the novelty of compounds with regard to intellectual property (IP) is usually a crucial criterion in many aspects: hit/lead selection, compound library design, and evaluation of compound acquisition and maintenance. For a small number of molecules, the novelty evaluation is often performed manually against known data, for example, against the online information in the SciFinder database (Chemical Abstracts Service). This evaluation becomes harder for large numbers of compounds, but using the more efficient measures of molecular similarity discussed above it is possible to do a molecule-by-molecule comparison of a screening set against a large reference set to ensure a relatively high IP value for the screening compounds. Other methods, such as privileged substructure searches, can also be applied in a similar way. The reference sets are usually commercial databases installed in-house, such as the CMC, MDDR, and WDI. More specialized reference sets for individual target classes such as kinases and GPCRs are provided by Jubilant Biosys, Inc., Aureus Pharma, Inc., and Eidogen-Sertanty, Inc.

3.5 COMPOUND COLLECTION MAINTENANCE AND ACQUISITION

The success of a screening project, whether it is done virtually or experimentally, depends on the quality of the compound source. The physicochemical profile of hits in an organization tends to resemble that of the overall screened compounds, as observed by Lipinski at Pfizer, Inc. (Lipinski et al., 1997) and also by us at Serono. Many compounds in a corporate collection have been derived from hit to lead programs and are consequently close analogues and not diverse. This gives rise to
the need to maintain a favorable druglike profile and diversity of chemotypes in a corporate collection, using the methods described previously. Additional logistics issues may be encountered that impact a computational selection of compounds for screening. For example, in most corporate collections compounds are stored on plates ready for screening. Therefore computational methods need to be applied at the plate level and not the compound level, which can complicate calculations. If the opportunity arises, selection can be made at the plate creation stage.

Other considerations also require computational support in order to have a high-quality compound collection. Corporate collections of compounds need to be replenished on a regular basis for a number of reasons: compound samples stored over a number of years deteriorate, especially in solution; certain compound classes may no longer hold the IP position that they once did; old compounds may fail newly implemented methods for assessing druglikeness; or old compound classes may no longer have relevance to current targets or therapeutic areas. Some organizations decide to simply discard all compounds older than a certain date. Companies looking to add to their collections, or smaller or newer companies screening for the first time, need to rely on either an outside source or a technology that can rapidly produce large numbers of compounds. There exist many vendors of compound sets for screening, and most will freely supply electronic files or databases of their libraries for computational analysis. Some supply focused sets for target types (e.g., kinase, GPCR) that have been computationally designed, or they work with customers on a contract basis to fulfill their requirements using their own proprietary design technologies. Therefore the same considerations with regard to druglikeness, IP, and diversity should be applied to purchased compounds. An additional consideration is complementarity to the existing collection, where the similarity methods described above can be applied to compare possible new compounds against the existing set.

3.6 PERSPECTIVE

The sections above have outlined the desirable properties that a compound screening set should possess and the methods available to achieve this goal. The methods are many and varied, so in case a reader seeking guidance is overwhelmed by possibilities, we would like to conclude by making some remarks about how screening sets are actually chosen at Serono. The overall scheme is summarized and illustrated in Fig. 3.7. The selection is driven by the nature and goals of the project and by the information already available. A final screening set is usually put together from a target set, a target-family set, and a diversity set chosen from our corporate collection. The compounds in the target set are selected based on any information that is available on a particular protein target, and the target-family set is based on knowledge of the protein structural or functional family of which the target is a member (e.g., protein kinases, or GPCR family A). It is highly desirable to have a known set of active ligands to create a target set, by either the similarity or DSA methods described above or a more sophisticated QSAR or pharmacophore model. If there are no known active compounds, then modeling can be based on the actives for the closest targets for which information is available. Each of these sets may be a combination of smaller
sets chosen by different methods. In addition, other sets may be added—for example, a newly acquired compound collection from a commercial vendor. To some extent, the methods used reflect the experience and preferences of the computational chemist on the project. At Serono we make extensive use of methods such as DSA, PharmPrint, and diversity analysis that have been developed in-house or by the authors elsewhere. In addition, we use a combination of commercial and in-house tools for fingerprinting of compounds in similarity and other methods. The different sets are then combined according to weights reflecting the amount of information we know about the target and the project team’s desire to explore the possibility of discovering novel compound classes by serendipity. Because of the unpredictable nature of drug discovery, we rely on the combination of the sets and methods to avoid becoming too attached to one particular approach. Compounds with undesired properties in the combined set are excluded using the above-mentioned druglike and related filters. The remaining compounds are then subject to pairwise similarity analysis, and highly similar compounds are removed. The list may finally be subjected to inspection by eye (often by a medicinal chemist) before it is delivered to screening.

We have experienced mixed success with docking as a method for choosing screening compounds, where some small sets chosen by this method produced no positives in screening, an observation that is consistent with other findings; for example, docking does not always perform well in all target classes (Cummings

---

**Figure 3.7** Illustration of a way that compounds chosen by different methods are combined to create a final screening set. The actual procedures can vary from one implementation to another. For example, the exclusion of non-druglike compounds is often performed before docking.
et al., 2005). Consequently, we have not used high-throughput docking to a large extent in choosing screening sets, preferring to use more thorough (and therefore slower) docking methods at the post-HTS stage. However, we have found improvements in enrichment by implementing pharmacophore or substructure constraints in docking, either as part of the algorithm or in postprocessing. In addition, we have recently experimented with using methods such as DSA and PharmPrint as a filter to reduce the number of compounds before the docking stage. A more thorough docking search can then be implemented on a smaller set of compounds. This implements a “funnel” concept in which a large number of compounds are put through progressively more thorough methodology, reducing the number of compounds at each stage, to produce a small screening set highly enriched in actives. In a test case this produced positive results.

Despite the many unsolved problems, the appeal of docking remains. Because it is based on the key–lock principle of drug-target interaction, docking has the potential to be the ultimate solution to discovering a novel small-molecule binder to a protein. There have been many new docking methods developed recently, and much attention is being devoted to the subject, with consequent improvements in algorithms and results. In many cases, the correct binding modes of ligands to their targets have been successfully predicted, and some novel and highly active compounds have been identified (Desai et al., 2004; Li et al., 2004; Lyne et al., 2004). Structure-based design studies of molecules that eventually lead to clinical compounds have been well documented for their important roles in drug discovery (Davis et al., 2003).

ADME/Tox is then the other important issue involved in turning a small-molecule inhibitor into a drug candidate. There are many commercial software packages available for predicting a variety of endpoints (Cronin et al., 2003; Ekins, 2003). The accuracy of predicting some properties, such as metabolism and toxicity, is still an issue, at least at the high-throughput level. New methods as well as more training data are needed to improve the prediction quality. At present, it is only practical to calculate basic descriptors for large numbers of molecules, such as the rule-of-five descriptors, number of rotatable bonds, and polar surface area (Clark, 1999a,b; Ertl et al., 2000; Palm et al., 1996). The more time-consuming methods are only practical at the hit-to-lead stage.

Computational methods have become an indispensable tool in drug discovery. Accurate predictions of both high-affinity binding and a favorable ADME/Tox profile will move us closer to identifying drug candidates in silico prior to laboratory screening. With improved algorithms, we expect that docking and other methods will identify high-affinity binders with increasing frequency. In addition, improvements in ADME/Tox prediction accuracy and efficiency will enable successful prediction of the in vivo behavior of compounds on a large scale. A combination of these two technologies will be a powerful tool in the hands of the computational chemist. This progress will be aided by (1) the fact that information available from genomics and X-ray crystallography is growing at an unprecedented rate and (2) increases in computer power drive new kinds of computations that were intractable a few years before. While the prospects for performing a drug-discovery project completely in silico are still distant, we anticipate an increased role for computational methods in drug discovery and development.
ACKNOWLEDGMENTS

We would like to thank the members of the Design Technology Team at Serono for contributions to this work.

REFERENCES


4.1 INTRODUCTION

Even though many routes for the discovery of new drugs exist, most clinically used drugs were discovered accidentally or by screening against the enzymatic products, chemicals, and natural products. In the traditional drug discovery process, a lead compound might be obtained out of around 80,000 compounds, and then lead optimization is carried out by numerous efforts in chemical synthesis to improve its activity or to reduce the toxicity. A structure–activity relationship (SAR) might be achieved after a variety of analogues are synthesized, and then SAR could be used as a guideline for future synthesis and optimization. Sulfonamide drugs, penicillin, and 5-fluorouracil (5-FU) were found in this way. Their action mechanisms were explored afterward. For example, many derivatives of sulfonamides were prepared to evaluate the antibacterial activity, and this exhaustive synthetic work led to the discovery of other treatments such as diuretics, drugs for leprosy, antimalarial agents, anticoagulant agents, and antihypertensive drugs.

Advances in molecular biology and the Human Genome Project have revealed a superfluity of potential targets. The techniques of X-ray crystallography and nuclear magnetic resonance (NMR) enable the structural determination of many biological targets (enzymes or proteins). Then, pharmaceutical companies have invested heavily in lead discovery in target assessment technologies such as high-throughput screening.
Combinatorial synthesis makes large chemical libraries available, and HTS might afford more hits for finding potential lead compounds. However, problems of how to increase the probability of success in screening of vast chemical libraries still exist (Spencer, 1998; Bailey and Brown, 2001). In addition, screening large chemical databases remains an expensive and time-consuming process, especially for academic research. There are over 3000 structures deposited every year and more than 32,000 protein structures in total available today in the Protein Data Bank (PDB, Bernstein et al., 1977). Therefore, computational methods are the best alternatives to increase efficiency and reduce costs in the drug discovery process. In addition, three-dimensional (3D) molecular databases are available to implement the computational screening (Shadowski and Gasteiger, 1993; Manallack, 1996).

“Direct” and “indirect” are the two major strategies used in the current drug design process (Cohen et al., 1990). In relation to the availability of target (enzyme or protein) structures and ligands as shown in Fig. 4.1, drug discovery can be divided into four phases: no target 3D structure or ligands (phase 1), no target 3D structure but ligands available (phase 2), target 3D structure available but no ligands (phase 3), and both target 3D structure and ligands available (phase 4). In the early stage, no 3D target structures and little or no SAR are known, and therefore HTS and CombiChem are the two major approaches (phase 1). In the indirect strategy, the structural features and biological activity of inhibitors or ligands are taken into account for quantitative structure–activity relationship (QSAR) or pharmacophore-based design (phase 2). In the direct approach, the 3D features of the enzyme or receptor are directly considered for de novo design (phase 3) and structure-based design (phase 4). In this review, we will describe ligand- and structure-based design, employing computational techniques and their features.

### 4.2 LIGAND-BASED DRUG DESIGN

#### 4.2.1 Quantitative Structure–Activity Relationship

When ligands or inhibitors are available, QSAR was the first method developed for virtual screening and has been advanced from a two-dimensional (2D) to a 3D model. Quantitative structure–activity relationship analysis uses statistical methods to study the

![Figure 4.1 Strategies in drug design respective to known or unknown 3D structure of the target.](image-url)
correlation of biological activity, physicochemical properties, or toxicities to structural properties of candidate molecules. Computer modeling analyzes a series of chemical molecules possessing a diversity of binding affinities to a protein or an enzyme. A QSAR attempts to find consistent relationships between the variations in the values of molecular properties and the biological activity. In the classical or 2D QSAR techniques, traditional SAR methods such as the Hansch equation (Hansch and Fujita, 1964), Topliss tree (Topliss, 1972), and Craig plot (Craig, 1971) could be used to optimize the leads.

Modern 3D QSAR methods analyze the interaction fields around a molecule by calculating the interaction energy in a grid. This is repeated for all molecules in the series, with their alignment relative to each other being a crucial parameter. The interaction energies for each point in the grid are then subject to a QSAR across the set of molecules. The most well-known 3D QSAR technique is Comparative Molecular Field Analysis (CoMFA, http://www.tripos.com). The CoMFA approach calculates molecular properties such as steric, electronic, hydrogen bonding, and hydrophobic fields (Cramer et al., 1988).

4.2.2 Substructure Search

When a lead compound is known, the similarity or substructure search can be carried out to retrieve novel compounds, containing the main skeleton or group present in the original lead, from databases. Structural search is the first type of screening method broadly employed for screening chemical databases. Once important structural features or patterns are determined, they can be used to search databases for all molecules possessing the features. Then, these molecules can be verified in a biological assay. Recently, we discovered compound soluble blue as a hepatitis C virus (HCV) helicase inhibitor with an IC$_{50}$ value of 40 µM (Fig. 4.2). Then, we used the 2D substructure search approach to find compounds based on the structure of

![Substructure](image)

**Figure 4.2** Example of HCV helicase inhibitor derived from substructure search.
soluble blue. The fragment was specified as a query substructure in which the polar groups SO₂ and NH might be involved in the H-bonding interactions with the enzyme. The substructure search was conducted using the ISIS/Base search program and the MDL database Available Chemicals Directory-Screening Compounds (ACD-SC, MDL Information Systems Inc., San Leandro, CA, USA, http://www.mdli.com). A more potent HCV helicase inhibitor with an IC₅₀ value of 12.6 μM was identified (Chen et al., in preparation).

4.2.3 Pharmacophore-Based Design

The pharmacophore is an important concept in rational drug design. It represents a set of functional features in 3D geometry that can interact with a specific receptor to gain activity. In other words, a pharmacophore is the geometrical depiction of fragments or features believed to be responsible for the biological activity (Ghose and Wendoloski, 1998). In general, a pharmacophore is atom-based and defined as chemical properties such as acid, base, hydrophobe, or aromatic group. Sometimes, others such as plane and normal are included. When an unknown structure of a receptor or a membrane-associated protein is the target, the pharmacophore model becomes a useful tool for searching lead compounds since the structure-based virtual screening is not accessible. Pharmacophore-based virtual screening has been demonstrated as an efficient method and can reach hit rates of 1–20%.

Many approaches have been utilized in the generation of a pharmacophore. They can be divided into three categories: (1) similarity of a known agent against an enzyme or a ligand to a receptor: this method predicts the necessary shape of molecules in binding to a target protein through the shape-similarity of molecules; (2) prediction of bioactive molecules with different structures: this method analyzes a set of structurally different chemical molecules with diverse activity, finds the consistency and specificity, and determines the pharmacophore model of a protein or an enzyme; and (3) analysis of the target’s active site: The application of the pharmacophore model can extend to the structure-based design. By analyzing the active site of a target, one can find a ligand possessing specific chemical features and suitable binding orientation. Combined pharmacophore-structure-based virtual screening would be a potential virtual screening approach but is more time-consuming. The generation of a pharmacophore can be achieved by DISOtech and RECEPTOR (Tripos, St. Louis, Missouri, USA; http://www.tripos.com) and CATALYST (Accelrys, San Diego, California, USA; http://www.accelrys.com). Once the pharmacophore is determined, it can be used to screen chemical libraries for compounds fitting to the pharmacophore. Through the biological assay to validate the model, a potential novel lead compound might be generated. For example, in CATALYST, pharmacophore models are also called hypotheses and consist of 3D geometrical arrangements of chemical functions. The chemical functions are the features related in 3D geometry and represented by a set of spheres as shown in Fig. 4.3 (Chen et al., 2001). In general, a pharmacophore consists of three to five features. The volume of the sphere stands for the accessible space of the chemical functions. For organic molecules, different structures could have the same chemical function and therefore result in the same biological activity.
This approach is based on two fundamental assumptions: (1) a same binding mode for all compounds in the training set to a receptor and (2) the more binding interactions between compound and receptor, the higher the activity. The selection of the training set, a set of molecules, plays the essential role in determining the success of HypoGen. Compounds in a training set should contain a large variety of chemical structures and interact with the target via the same mechanism with a range of potency from highly active to inactive. There are two successive major steps in the generation of a pharmacophore. First, a conformational analysis is needed to determine all the possible low-energy conformations. The biological activity of a molecule is associated with a unique conformation, and it often is not the global minimum-energy conformation (Ghose et al., 1999). The search of the conformational model is the major task in this approach (Bausen and Shands, 1996). Second, all the conformations in the training set are used to calculate their stereoelectronic properties, such as steric, electronic, lipophilic, and hydrogen bonding. Then, a pharmacophore is generated on the basis of the stereoelectronic properties of the training set. A 3D database search is used to identify compounds mapped to the pharmacophore. For example, HypoGen (Sutter et al., 2000; Kurogi and Güner, 2001) in CATALYST analyzes the relationship of biological activity and structures in the training set. Then, it automatically generates a chemical functions–based pharmacophore model. Following a schematic process as shown in Fig. 4.4, we generated a pharmacophore for 5α-reductase inhibitors, which was used to screen the NCI database (Milne et al., 1994) for novel nonsteroidal lead compounds. A series of isoflavone derivatives were identified as 5α-reductase inhibitors as shown in Fig. 4.5 (Chen et al., 2001).

Many successful lead identifications using pharmacophore searches have been reported, and Fig. 4.6 depicts five successful case studies. Novobiocin (1) is a known inhibitor of bacterial enzyme DNA gyrase, and a pharmacophore was obtained based on the structure of 1 (Böhm et al., 2000). Seven classes of novel DNA gyrase inhibitors were recognized, and among them a 3D guided optimization provided a highly potent DNA gyrase inhibitor, 3,4-disubstituted indazole (2), which was 10 times more potent than 1. According to the X-ray structure of the thyroid hormone (3), a pharmacophore containing hydrophobic, hydrogen bond donor and acceptor features, and excluded volumes was defined (Greenidge et al., 1998). A new ligand (4) was subsequently shown to bind to THR-R with an IC$_{50}$ value of 69 μM. A series of farnesyl protein transferase (FPT) inhibitors exemplified by Sch44342 (5) were analyzed by CATALYST to generate a pharmacophore model containing five features.
of four hydrophobic and one hydrogen bond acceptor regions (Kaminski et al., 1997). A 3D search of the pharmacophore identified two peptides such as 6 and two dihydrobenzothiophenes such as 7 exhibiting in vitro FPT inhibitory activity with IC\(_{50}\) values below 5 \(\mu\)M. A Human Immunodeficiency Virus type 1 Integrase (HIV-1 IN) inhibitor (8) was taken as a lead to generate a pharmacophore, and after modification of the pharmacophore a 3D search led to highly potent inhibitors 9 and 10 (Nicklaus et al., 1997; Hong et al., 1997).

### 4.3 DE NOVO LIGAND DESIGN

With known 3D structure targets (enzymes or proteins), new inhibitors or ligands can be designed de novo. De novo design programs try to construct novel structures, using sets of predefined fragments, into an active site or onto a pharmacophore model. The combinatorial nature of de novo design leads to a large amount of chemical structures. GRID (Goodford, 1985) is the first de novo method. Then, de novo ligand design programs such as LEGEND (Nishibta and Itai, 1993), LeapFrog (http://www.tripos.com), LigBuilder (Wang et al., 2000), SPROUT (http://www.simbiosys.ca/sprout/index.html), CONCERTS (Pearlman and Murcko, 1996), PRO-LIGAND (Waszkowycz et al., 1994), TOPAS (Schneider et al., 2000), BUILDER (Roe and Kuntz, 1995), and LUDI (http://www.accelrys.com, Böhm, 1992a,b) were
Figure 4.5 Isoflavone derivatives as 5α-reductase inhibitors from pharmacophore screening. See color plates.
developed to build up structures according to the active site of the receptor. *De novo* design generates functional groups in suitable forms and suitable conformations to interact with the protein surface. More importantly, *de novo* design programs can also construct structures including novel skeletons. They show how suitable small fragments can be positioned into clefts of protein structures. For example, Ludi has two major processes: (1) It can be used to design inhibitors *de novo*, and (2) it can be used to suggest modifications for a known ligand. When run in link mode, Ludi fits fragments in a way that simultaneously exploits available hydrogen bonding and hydrophobic sites on the receptor and suggests how fragments might be bonded to the existing ligand. The processes of Ludi *de novo* design are the following: (1) On the basis of the 3D structure of the receptor, the program computes the points of interaction that will be fulfilled by a ligand within the protein’s active site. The interaction points are hydrogen bondings and hydrophobic sites. Four directional interaction sites are defined for portions of the binding site: lipophilic–aliphatic, lipophilic–aromatic, hydrogen donor, and hydrogen acceptor. (2) Fragments from a library of fragments are aligned with these interaction sites and are scored with respect to their degree of overlap. (3)
Fragments are then connected in a straightforward way (the two closest hydrogens between fragments are removed and a linker of the appropriate length is attached to each terminal atom in the fragments). (4) Conformational flexibility can be accounted for by including multiple conformations of a fragment in the library. Most de novo programs contain two strategies of ligand generation: (a) growth (Liu et al., 1999; Rotstein and Murcko, 1993) and (b) linkage (Böhm, 1996) as shown in Fig. 4.7.

De novo ligand design led to a novel series of factor Xa (f Xa) inhibitors incorporating an amidino 6,5-fused bicyclic moiety such as (11, SE170, Fig. 4.8) with a $K_i$ value of 0.3 nM and 350-fold selective for fXa over trypsin (Han et al., 2000). A second example is the ligand against human cannabinoid receptor 1 (CB1). The TOPAS program was used to design a library of ligands and led to a very potent ligand (12) with a $K_i$ value of 0.3 μM. A potent CDK4 inhibitor (13) with nanomolar inhibitory activity was developed by LEGEND.

4.4 STRUCTURE-BASED DESIGN

When the 3D structure of the target enzyme or protein is known, the lead identification and optimization can be efficiently processed with the aid of molecular modeling. It is especially beneficial to have the 3D structure of a protein–ligand complex so that the active or binding site can be explored.
4.4.1 Docking Techniques

Molecular docking techniques dock small molecules into the protein binding site. Success of the docking depends on the 3D target protein structure with or without a ligand bound. The 3D structures have to be detailed at atomic resolution. With the vast 3D protein structures available in the PDB, one should keep in mind that small changes in structure significantly influence the results of docking (Muegge, 1999), and crystal structures employed in docking should have resolution below 2.5 Å. Ligand–receptor docking has become the choice to accelerate structure-based drug design and has been successfully used in the discovery of drug candidates. Docking procedures try to identify correct poses of ligands in the binding site of a protein and to predict the binding affinity between the ligand and the protein. Comprehensive reviews on the docking techniques have been appeared in the literature (Boyd and Lipkowitz, 2001). Most methods available now treat the ligands flexibly and the proteins rigidly. Van Leeuwen has compiled some docking algorithms (http://www.bio.vu.nl/nvtb/Docking.html). Table 4.1 depicts some docking methods used for single molecule docking, and Table 4.2 shows the tools used only for single molecule docking. Among them, DOCK, FlexX (Rarey et al., 1997), and GOLD (Jones et al. 1995) are frequently used platforms in the docking approach.

4.4.1.1 Program DOCK

The computer program DOCK (Meng et al., 1992; Kuntz et al., 1994) is the most widely used computational docking method and provides a rapid method by using the 3D structure of a target receptor to identify putative ligands among commercially available compounds. When a ligand fits to a receptor well, it may bind to the receptor with high affinity, and it has the potential to be an inhibitor for a biological protein receptor. DOCK discovers how a ligand and a receptor fit together. DOCK has been utilized to discover leads in a number of diverse protein systems such as HIV-1 protease (Rutenber et al., 1993), thymidylate synthase (Shoichet et al., 1993), influenza hemagglutinin (Bodian et al., 1993), purine salvage enzymes in parasitic...
protozoa (Somoza et al., 1998), and macrophage migration inhibitory factor (Orita et al., 2001). Several successful examples in structure-based lead generation have been reported. This structure-based computational method, using the DOCK suite to screen a large database of small molecules, was shown to be a powerful tool in identifying drug leads.

In DOCK, spheres are created to map and fill the molecular surface of an active site. A set of spheres represents the ligand. In general, the docking process includes four primary tasks: (1) creating a negative image of the active site represented by a set of overlapping spheres whose centers are putative locations for ligand atoms, (2) matching sphere centers to atoms of small molecules from a database, (3) scoring and ranking the docked orientations of compounds on the basis of their complementarities to the protein surface, and (4) reviewing the docked orientations of the best-scoring compounds. The DOCK algorithm (Lightstone et al., 2000) was illustrated in Fig. 4.9.

DOCK is used to locate favorable binding orientations for a ligand molecule and a receptor molecule. The program evaluates the ligand orientations by a contact (shape-fitting) score, a forced field interaction energy, and a user-defined chemical

**TABLE 4.1  Software Available for Single Molecule Docking and Database Screening**

<table>
<thead>
<tr>
<th>Software</th>
<th>Docking Method(s)</th>
<th>Scoring Method(s)</th>
<th>Web Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>GLIDE</td>
<td>Rigid docking</td>
<td>Empirical scoring, force-field scoring</td>
<td><a href="http://www.schrodinger.com">www.schrodinger.com</a></td>
</tr>
<tr>
<td>DOCK</td>
<td>Flexible docking</td>
<td>Force-field scoring, chemical scoring, contact scoring</td>
<td><a href="http://www.cmpharm.ucsf.edu/kuntz/dock.html">www.cmpharm.ucsf.edu/kuntz/dock.html</a></td>
</tr>
<tr>
<td>FlexX</td>
<td>Flexible docking</td>
<td>Empirical scoring</td>
<td>cartan.gmd.de/FlexX</td>
</tr>
<tr>
<td>LigandFit</td>
<td>Monte Carlo</td>
<td>LIGSCORE, PLP, PMF, LUDI</td>
<td><a href="http://www.accelrys.com">www.accelrys.com</a></td>
</tr>
<tr>
<td>FRED</td>
<td>Rigid docking</td>
<td>Chemscore, PLP, Gaussian shape scoring, SCreenScore</td>
<td><a href="http://www.eyesopen.com/fred.html">www.eyesopen.com/fred.html</a></td>
</tr>
<tr>
<td>DockVision</td>
<td>Monte Carlo, genetic algorithm</td>
<td>Force-field scoring</td>
<td><a href="http://www.dockvision.com">www.dockvision.com</a></td>
</tr>
<tr>
<td>GOLD</td>
<td>Genetic algorithm</td>
<td>Van der Waals and hydrogen bond potentials</td>
<td><a href="http://www.ccdc.cam.ac.uk/prods/gold">www.ccdc.cam.ac.uk/prods/gold</a></td>
</tr>
</tbody>
</table>

**TABLE 4.2  Software Available Only for Single-Molecule Docking**

<table>
<thead>
<tr>
<th>Software</th>
<th>Docking Method</th>
<th>Scoring Method(s)</th>
<th>Web Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liaison</td>
<td>Flexible docking</td>
<td>Free–energy scoring</td>
<td><a href="http://www.schrodinger.com">www.schrodinger.com</a></td>
</tr>
<tr>
<td>Affinity</td>
<td>Flexible docking</td>
<td>Molecular mechanics, force-field scoring</td>
<td><a href="http://www.accelrys.com">www.accelrys.com</a></td>
</tr>
</tbody>
</table>
scoring scheme. The binding energy is taken to be the sum of the van der Waals attractive, van der Waals dispersive, and Coulombic electrostatic energies. The best-scoring orientation of each molecule is saved. The molecules in the list of best-scoring compounds may have the potential as new ligands for the receptor. The docking procedure is divided into four general stages: ligand preparation, site characterization, scoring grid calculation, and docking itself, as shown in Fig. 4.10 (http://www.cmpharm.ucsf.edu/kuntz/dock.html).

### 4.4.1.2 Rigid Docking

Early versions of the DOCK program (Kuntz et al., 1982) performed rigid-body docking based on the distance-compatible match searches. Later, chemical properties are taken to match spheres (Shoichet et al., 1993). In DOCK 4.0, different scoring functions (Meng et al., 1992, 1994; Gschwend and Kuntz, 1996) and solvation (Shoichet et al., 1999; Zou et al., 1999) are available. Rigid-body docking, also called orientation searching, is a process that considers the orientations of ligands. The conformations of ligands are kept. First, compounds are docked using rigid-body docking to search for favorable binding orientations based on the interaction force field scoring that included van der Waals and electrostatic terms. The number of orientations can be defined. Since this process is less time-consuming, it can be used to screen large-scale databases. This method can rank a larger amount of compounds that might bind to the receptors. This procedure can narrow down the size of databases to be in-house databases, which then can be used for flexible docking. Program FlexX employs a pose clustering technique (Rarey et al., 1996a,b).

### 4.4.1.3 Flexible Ligand Docking

In general, the biologically active conformations of ligands are not the conformations with the lowest energy (global minimum). Therefore, flexibility is considered for
ligands. The conformation of a flexible molecule is searched or relaxed in flexible ligand docking. Normally, it is approached by fragmentation, ensembles, or genetic algorithms. For example, in DOCK, only the torsion angles are modified, and the bond lengths and angles are not changed. Only the interactions between segments are considered. There are two processing stages involved in this procedure: identification of rigid segments and conformation search. A flexible molecule is considered as a collection of rigid segments. Each segment contains the largest set of adjacent atoms separated by nonrotatable bonds. Segments are separated by rotatable bonds. This procedure can predict the binding more accurately, but it is much more time- and memory-consuming. Therefore, it will be employed on smaller and in-house databases.

### 4.4.2 Scoring Function

The binding affinity of a noncovalent ligand–protein complex is governed by the change in enthalpy and entropy of the whole system. In fact, a docked complex involving ligands and proteins actually is associated with the changes of free energies. However, the binding of free energies is difficult to estimate. The docking result is estimated by the binding affinity, and the scoring functions range from computer-intensive theoretical calculations to simple empirical methods (Murcko and Murcko, 1995; Tame, 1999). Since docking programs need to screen large databases
efficiently, in practical terms the scoring functions favor speed over theoretical accuracy to get qualitative scores. As a result, most scoring functions estimate the binding affinity according to hydrophobic, surface area, hydrogen bonds, and solvent accessible area. Scoring functions describing protein–ligand interactions play a key role in identifying the correct binding geometry of a protein–ligand complex. Researchers have invested a great deal of effort to achieve accurate predictions of binding affinities for ranking different complexes (Guida, 1994; Colman, 1994; Kuntz, 1992; Lybrand, 1995; Rosenfeld et al., 1995; Lengauer and Rarey, 1996).

In general, scoring functions are divided into four categories: (1) force-field-derived scoring functions considering nonbonded interactions (Jones et al., 1997; Weiner et al., 1984, 1986) or solvations (Zhang et al., 1997; Shoichet et al., 1999); (2) regression-based scoring functions (Horton and Lewis, 1992; Böhm, 1998; Eldridge et al., 1997; Head et al., 1996); (3) knowledge-based scoring functions (Gohlke et al., 2000; Muegge and Martin, 1999; Verkhivker et al., 1995); and (4) other scoring functions including chemical scores, contact scores, or shape scores (Shoichet et al., 1992; Nicklaus et al., 1995; Stouch and Jurs, 1986; Wang et al., 1998; Flower, 1998). FK506 binding protein (FKBP) ligand 14 (Burkhard et al., 1999), farnesyl transferase inhibitor 15 (Perola et al., 2000), HIV-1 RNATAR inhibitor 16 (Filkov et al., 2000), and Aldose reductase inhibitors 17 (Iwata et al., 2001) are successful cases from docking (Fig. 4.11).

### 4.5 CONCLUSION

Pharmacophoric models and de novo ligand design provide potential leads that would possess novel chemical structures, and virtual screening of 3D pharmaceutical or chemical libraries would give bioactive compounds commercially available or prepared feasibly according to known methods. As shown in Fig. 11.12, ligand- and receptor-based virtual screening can be used in combination with each other and can
be broadly employed in the discovery of bioactive compounds. Indisputably, one needs to come back to bench work for synthesis and biological assay, even though in silico drug design and virtual screening can accelerate the process of drug discovery. Currently, molecular modeling including ligand- and structure-based drug designs play significant roles in drug discovery. Especially, the advances in high-speed computing platforms have made multiprocessor workstations and PC clusters affordable to the majority of researchers. With the improvement of computing methods and scoring functions, virtual screening is becoming important and may be indispensable in the drug discovery.

Figure 4.12 Schematic illustration of virtual screening.

Figure 4.13 Iterative process in structure-based drug design.
As a whole, the available 3D structures of proteins or protein–ligand complexes enable lead generation by means of molecular modeling to increase efficiency. With the help of QSAR, pharmacophore, and synthesis, new secondary leads can be designed. In cooperation with experimental biological assays and with further rounds of an iterative process (Fig. 4.13), the concept of structure-based drug design would provide potent and specific compounds more productively in the drug discovery process.

REFERENCES


*J Med Chem* 36, 1700–1710.

Rutenber, E., Fauman, E. B., Keenan, R. J., Fong, S., Furth, P. S., Ortiz de Montellano, P. R., Meng, E., 

Sadowski, J., and Gasteiger, J. (1993). From atoms and bonds to three-dimensional atomic coordinates: 


*Proteins Struct Funct Genet* 34, 4–16.

Somoza, J. R., Skillman, A. G., Munagala, N. R., Oshiro, C. M., Knecht, R. M., A., Mpoke, S., Fletterick, 
salvage in a parasitic protozoan. *Biochemistry* 37, 5344–5348.


the volumes and shapes of chemical compounds. *J Chem Inf Comput Sci* 26, 4–12.

Sutter, J., Güner, O., Hoffmann, R., Li, H., and Waldman, M. (2000). Pharmacophore perception, development, 
and use in drug design. In International University Line, La Jolla, CA, USA.


applied to the prediction of human immunodeficiency virus 1 protease binding affinity. *Protein Eng* 8, 677–691.


Zhang et al. (1997).

5.1 INTRODUCTION

The chemical space was estimated to have chemical structures in the order of magnitude of up to $10^{100}$ (Walters et al., 1998). This inconceivably large number gives hope that every therapeutic target would have an appropriate chemical as its ligand and that there might exist several different molecules exhibiting isofunctional activity (Schneider, 2002). Therefore, the task of drug discovery is to find such molecules that specifically bind to a certain therapeutic target relevant to a disease. There are many approaches developed for the task, one of which is target structure-based drug discovery (Lyne, 2002), which is now an important technique in cases where the three-dimensional (3D) structure of a therapeutic target is available. Numerous excellent reviews have been published on this approach (Lyne, 2002;
Jansen and Martin, 2004; Alvarez, 2004; Kitchen et al., 2004; Beavers and Chen, 2004; Erlanson, 2004). Computationally, there are at least three different target structure-based methodologies established for identifying or creating chemicals as potential drug leads, namely, docking-based virtual screening, *de novo* drug design, and combinatorial library design.

### 5.1.1 Docking-Based Virtual Screening

Virtual screening can be defined as a process of reducing a library containing a great number of compounds to a limited number of potentially promising compounds for a specific therapeutic target by means of computational methods. The screening consists of positioning in a target a series of small molecules contained in a chemical database ("docking") followed by scoring. Thus, molecular recognition is the fundamental basis for virtual screening (Alvarez, 2004). This action mechanism was firstly proposed by Emil Fischer in 1894 as the "lock-and-key" principle whereby steric and electrostatic complementarity between proteins and their ligands drive the complexation process (Fischer, 1894). Docking-based virtual screening algorithms attempt to generate and identify the most complementary match between a ligand from a ligand database and its macromolecular target (Brooijmans et al., 2003; Glen and Allen, 2003; Halperin et al., 2002). Scoring functions are used in the approach for ranking both the different ligands and the different orientations of the same ligand. There have been many publications describing the identification of novel ligands by docking-based virtual screening methods (Alvarez, 2004). The most remarkable characteristic of this approach is that the potential ligands discovered are usually structurally diverse.

When high-performance computing is utilized, the virtual screening of compound libraries is very time-efficient and cost-effective relative to real high-throughput-screening (HTS) (Lengauer et al., 2004). Therefore, it is a promising tool as a preliminary step toward discovering drug hits or leads. Furthermore, it also provides the interaction details between a target and its ligands for further lead optimization. Thus, docking-based virtual screening has become part of today’s lead discovery arsenal for the identification of more diverse hits and has been established as a complement to real HTS. However, as the screening is performed in compound databases, the discovered potential leads are limited within the available compounds in the databases. In practical terms, there are true ligands missed due to algorithmic limitations in conformational sampling (Brooijmans et al., 2003) and imperfect scoring functions (Wang et al., 2003), no matter which docking programs are employed. On the other hand, the user-defined parameters and selected algorithms throughout the docking process might result in different results by different users for a same project.

### 5.1.2 De novo Drug Design

*De novo* drug design does not start from a database of complete molecules but aims at building a complete molecule from molecular bricks ("building blocks") to chemically fill the binding sites of a target molecule (Rotstein and Murcko, 1993). The complete chemical entries could be constructed through linking the "building
blocks” together, or by growing from an “embryo” molecule with the guidance of evaluation of binding affinity. The “building blocks” could be either atoms or fragments (functional groups or small molecules). However, using atoms as “building blocks” is thought to be inefficient; therefore, this approach is seldom used nowadays. In the fragment-linking approach, the binding site is mapped to identify the possible anchor points for functional groups. These groups are then linked together and form a complete molecule. In the sequential-growing approach, the molecule grows in the binding site controlled by an appropriate search algorithm that evaluates each growing possibility with a scoring function.

Different from docking-based virtual screening, fragment-based de novo drug design can perform sampling in whole compound space, obtaining novel structures that are not limited in available databases. However, the quality of a growing step strongly depends on the previous steps. Any step chemically growing incorrectly would lead to an unacceptable result. For the fragment-linking approach, there is the problem of how to choose linkers to connect fragments together as complete structures. The most remarkable drawback of this approach might be the synthetic accessibility of the designed structures.

### 5.1.3 Combinatorial Library Design

Generally speaking, combinatorial chemistry is a collection of methods that allow the simultaneous chemical synthesis of large numbers of compounds using a variety of starting materials (Tropsha and Zheng, 2002). The primary benefit that combinatorial chemistry brings to drug discovery is speed and the large size of the designed libraries. For instance, one billion compounds can be synthesized from a molecular scaffold with three different substitution positions if each of the positions has 1000 different substituents. Therefore, combinatorial library provides a large number of structures for docking-based virtual screening. Actually, it is one of the most exciting developments in medicinal chemistry in the last decade (Gallop et al., 1994; Gordon et al., 1994; Terrett et al., 1995). Coupled with automation technologies and HTS, it offers great potential for discovering structurally novel drug leads. However, many products in the huge library are redundant. In addition, it was found that the large numbers of compounds synthesized did not result in a remarkable increase in drug candidates, although the number of compounds synthesized and screened has increased by several orders of magnitude (Oprea, 2002). It also does not make sense to validate and assay billions or millions of compounds.

### 5.2 FOCUSED LIBRARY DESIGN BASED ON HIT AND TARGET STRUCTURES

#### 5.2.1 Strategy

Because success in new drug development is directly tied to identifying drugable molecules, one of the biggest challenges is how to efficiently discover the molecules with novel structures, high affinities to therapeutic targets, ideal ADMET profiles, and
good manufacturing features. Different approaches for receptor structure-based drug discovery have different characteristics as discussed above. An effective way of discovering druglike molecules can be undertaken by using the advantages of different approaches. Moreover, 3D structural information and properties of a studied therapeutic target must be focused for specificity in discovering drug leads. Here we introduce (1) a focused library design method based on hit and target structures and (2) its application (Chen et al., 2005).

The focused library is built with fragments and geared toward one particular molecular target, which is actually a representative sample of the full chemical space. Starting with the structures of hits and a therapeutic target, the overall skeleton of potential ligands is schematically split into several fractions, or “fragments,” based on the interaction mechanism and the physicochemical properties of the binding site. An individual library, called a fragment library, is constructed for each fragment, taking into account the binding features of the fragments to the binding site. Finally, focused libraries on the studied target are constructed with the judgments from structural diversity, druglikeness, ADMET profiles, and binding affinities (Chen et al., 2005).

Some stochastic methods for exploring the whole chemical space are applied, of which methods a genetic algorithm (GA) is generally used because of its high efficiency in searching large combinatorial spaces (Gillet et al., 2002; Weber et al., 1998; Sheridan and Kearsley, 1995). There is no guarantee that the globally best solution will be found. However, good results are usually found much more quickly than a purely random search or a systematic search (Holland, 1975). A software package called LD1.0 was developed based on the approach as shown in Fig. 5.1 (Chen et al., 2005). The critical challenges are, first, to select sets of fragments that have the best potential to be parts of new drug leads for a given target and, second, to set up proper criteria for product judgment, including structural diversity, druglikeness, ADMET, and binding affinities.

![Flowchart of the software LD1.0 for focused library design](image)

**Figure 5.1** Flowchart of the software LD1.0 for focused library design (Chen et al., 2005).
5.2.2 Fragments or Building Blocks

The quality of fragment libraries has a critical effect on the final focused library. There are at least three ways to construct the fragment libraries (Chen et al., 2005).

1. From Known Drugs or Inhibitors: Extracting fragments from known drugs or ligands (inhibitors or activators) of the studied target is an effective approach for collecting building blocks. A specific fragment library should be constructed based on available structural information of its ligands and the binding mechanism with the target.

2. From Existing Drug Fragment Libraries: Some fragment libraries, such as the fragment library in the Ludi module of the InsightII software, are often used for de novo drug design (InsightII 2000.1, 2002). The LD1.0 software has a default fragment database for those targets with little information about their ligands.

3. From Inhibitors of Similar Targets: Homology proteins usually share similar structural features and characteristics, especially at the binding site or active site. Therefore, the ligands for different targets belonging to the same family should share some common fragments in their inhibitors. Thus, the fragment libraries for a target could be constructed by referring to the structures of the ligands of its homology proteins.

5.2.3 Diversity of the Focused Library

Initially, the focus in combinatorial library design was on selecting diverse sets of compounds on the assumption that maximizing diversity would result in a broad coverage of bioactivity space and hence would maximize the chances of finding drug leads (Jorgensen and Pedersen, 2001). Suggestions and assumptions on how to assess diversity have been studied during the last decade (Willett and Winterman, 1986; Alexander, 2000; Brown and Martin, 1996; Matter and Potter, 1999; Flower, 1998b; Ashton et al., 1996; Warr, 1997; Pearlman and Smith, 1998; Gillet et al., 1999). It was reported that two-dimensional (2D) structural descriptions are fast and accurate for calculating molecular diversity (Flower, 1998b; Brown, 1997). The descriptors employed in LD1.0 are a set of the 39 best structural descriptors proposed by Flower (1998a). Distance method is selected to normalize each descriptor. Then, a Euclidean space is described by all descriptors with their weights. So the difference (or similarity), $d_{ij}$, of two molecules can be represented by the distance between the two molecular points in the space, calculated according to equation (5.1) (Chen et al., 2005),

$$ d_{ij} = \sqrt{\sum_k (\hat{x}^j_k - \hat{x}^i_k)^2} $$

(5.1)

where $\hat{x}^j_k$ is the normalized value of the $k$th descriptor for the $j$th molecule, which can be calculated using equation (5.2),

$$ \hat{x}^j_k = w_k \left( \frac{x^j_k - \bar{x}_k}{\sigma_k} \right) $$

(5.2)
where $x^j_k$ is the value of the $k$th descriptor for the $j$th molecule, $\bar{x}_k$ is the mean value of the descriptor $k$, $\sigma_k$ is variance difference of the descriptor $k$, and $w_k$ is weight of the descriptor $k$. Then, a library’s diversity, $D_k$, can be calculated using equation (5.3),

$$D_k = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{j<i} d_{ij}$$

(5.3)

where $n$ is the number of molecules in the library. So $D_k$ is a sum of the distances between any two molecules in a library. The final scores of all libraries in a certain GA generation are normalized using equation (5.4),

$$D_{k,Nor} = \frac{D_k - D_{min}}{D_{max} - D_{min}}$$

(5.4)

where $D_{k,Nor}$ is the final score of the $k$th focused library. $D_{max}$ and $D_{min}$ are the maximum and minimum $D_k$ values of a specific generation, respectively. Thus, the highest score of library diversity is 1 and the lowest is 0.

The reasonability of the descriptor set for calculating molecular diversity was verified by two libraries, No. 1 and No. 2, which were constructed from two sets of molecules randomly selected from the MACCS-II Drug Data Report (MDDR) with molecular weight less than 350. The library No. 1 has 23 molecules that are composed of only three elements, C, H, and O. The library No. 2, also randomly selected from MDDR, has 23 molecules as well, but without any restraint in their structures. Therefore, library No. 1 should possess lower diversity than library No. 2 (Chen et al., 2005). Thus, $D_{k,Nor}$ for libraries No. 1 and No. 2 should equal to 0.0 and 1.0, respectively. Meanwhile, two more libraries, No. 3 and No. 4, were constructed by mixing some molecules from library No. 1 with some molecules from library No. 2 (Table 5.1). Indeed, the calculation result using the molecular diversity descriptor set shows that the library diversity correlates with the number of the molecules from library No. 2, demonstrating that the structural diversity descriptor set is reasonably good (Table 5.1).

### 5.2.4 Druglikeness (Zheng et al., 2005)

Druglikeness is another key factor that should be considered during library design. It has been estimated that $\sim 40\%$ of compounds fail to be developed into drugs due to their poor pharmacokinetic properties (Stanton et al., 2000). Therefore, initial strategies toward this goal should be involved in use of computational filters to
remove compounds deemed to be chemically unsuitable for drug development. This approach has already been applied by several research groups (Gallop et al., 1994; Gillet et al., 1999; Lewell et al., 1998). In 1997, a set of assumptions about necessary features for a “good” drug candidate was suggested and embodied in a so-called “Rule of Five” by Lipinski et al. (1997). Then, other scientists developed new methods to improve prediction performance of the “Rule of Five” (Galvez et al., 2001; Ajay et al., 1998; Sadowski and Kubinyi, 1998; Ghose and Crippen, 1987; Xu and Stevenson, 2000). However, most of the methods have limitations in virtual screening, such as being time-consuming and performing poorly. Therefore, how to discriminate a druglike compound from a non-druglike one is still a great challenge in library design. Considering the advantages and disadvantages of the available computational methods of predicting druglikeness (SSKEYS; Viswanadhan et al., 1989; Gillet et al., 1998; The SPRESI database; Wagener et al., 2000; Muegge et al., 2001), a new chemistry space filter was developed and embedded in the software package LD1.0 (Chen et al., 2005; Zheng et al., 2005).

Two new descriptors, UNSATP and NO_C3, which are ratios of the molecular physicochemical properties, are defined in Table 5.2. The new descriptors are molecular size or molecular weight (MW) and database independent. Table 5.3 shows the initial values and distributions for the new descriptors.

To further improve the descriptors’ performance and robustness, the UNSATP is combined with NO_C3 to create a new chemistry space filter (Table 5.3).

### TABLE 5.2 Definitions of the Newly Developed Descriptors (Zheng et al., 2005)

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>UNSATP</td>
<td>UNSAT/BDUH</td>
</tr>
<tr>
<td>NO_C3</td>
<td>(N + O)/C3</td>
</tr>
<tr>
<td>C3</td>
<td>Number of (sp^3) hybridized C atoms</td>
</tr>
<tr>
<td>N</td>
<td>Number nitrogen atoms</td>
</tr>
<tr>
<td>O</td>
<td>Number of oxygen atoms</td>
</tr>
<tr>
<td>BDUH</td>
<td>Number of bonds that don’t contain H and X atoms</td>
</tr>
<tr>
<td>UNSAT</td>
<td>RNG + BD2 + 2*BD3 + (BDAR + 1)/2</td>
</tr>
</tbody>
</table>

### TABLE 5.3 Initial Values and Distributions for the New Descriptors and the New Chemistry Space Filter in Different Libraries (Chen et al., 2005)

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>Initial Value</th>
<th>ACD(^a)</th>
<th>MDDR</th>
<th>CMC(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>UNSATP</td>
<td>0.15 &lt; UNSATP ≤ 0.40</td>
<td>31.76%</td>
<td>57.17%</td>
<td>62.73%</td>
</tr>
<tr>
<td>NO_C3</td>
<td>0.2 &lt; NO_C3 ≤ 1.2</td>
<td>39.35%</td>
<td>68.09%</td>
<td>63.38%</td>
</tr>
<tr>
<td>Chemistry Space filter</td>
<td>0 ≤ UNSATP ≤ 0.43</td>
<td>39.09%</td>
<td>69.05%</td>
<td>70.39%</td>
</tr>
<tr>
<td></td>
<td>0.10 ≤ NO_C3 ≤ 1.8</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) The Available Chemicals Directory (ACD).

\(^b\) CMC, comprehensive medicinal chemistry.
Considering that ACD is not a pure non-druglike database, the final value for the new descriptors was reset and optimized to get maximum capability of the chemistry space filter in differentiating druglike and non-druglike compound databases (Table 5.3). The result indicates that this chemistry space filter can distinguish the MDDR and CMC more efficiently than individual NO_C3 or UNSATP. Therefore, this filter is suitable as a chemistry space filter to discriminate druglikeness and non-druglikeness (Zheng et al., 2005).

As the new filter is atomic type-based, it shows a favorable computational speed. The evaluation of drug-likeness can be completed within 2 hours for the MDDR database (~80,000 compounds) and within 4 hours for the ACD database (~300,000 compounds) on a computer with a CPU of 733MHz and RAM of 256M.

5.2.5 Activity Assessment

Molecular activity could be defined by its binding affinity to target protein. Nowadays, molecular docking is the most commonly used method to evaluate binding strength of a ligand to its target(s). One of the most popular molecular docking programs is DOCK4.0 (Kuntz et al., 1982; Oshiro and Kuntz, 1995; Ewing and Kuntz, 1997). The best energy score from DOCK 4.0 was used in LD1.0 for assessing molecular bioactivity. The score could be very different in different systems, suggesting that the ordinary methods for normalization cannot be used in the case of the activity score. To solve this problem, the sigmoid function, equation (5.5), was introduced into LD1.0 for dealing with the normalization,

\[
y = \frac{1 - e^{ax}}{1 + e^{ax}}
\]

where \(a\) is a constant, \(x\) is the binding energy from the output file of DOCK 4.0, and \(y\) is the normalized activity score; for the binding energy in the output file of DOCK4.0 is always no greater than 0.00 kcal/mol, and \(y\) is always between 0 and 1. Therefore, sigmoid function is a good choice for the normalization of the bioactivity score. Figure 5.2 demonstrates the plot of sigmoid function against binding energy when \(a\) is

![Figure 5.2](image)

**Figure 5.2** The plot of sigmoid function \((y)\) against binding score from DOCK 4.0.
equal to 0.05. If the value of $x$ is between 0.00 and $-80.00 \text{ kcal/mol}$, the normalized activity score, $y$, is between 0.0000 and 0.9640 (Fig. 5.2).

### 5.2.6 Test Cases for LD1.0

To verify the reliability of the software package LD1.0, two tests were carried out on building COX-2-focused and PPAR-γ-focused libraries (Chen et al., 2005).

#### 5.2.6.1 Focused Library Design for COX-2 Inhibitors

Cyclooxygenase (COX) is a key enzyme associated with arachidonic acid (AA) metabolism. COX has two isoforms, namely, COX-1 and COX-2 (Marnett, 2000; Garavito and Dewitt, 1999; O’Banion, 1999). It is believed that the inhibition of COX-1 causes the side effects seen with nonsteroidal anti-inflammatory drugs (NSAIDs). Therefore, the selective inhibitors of COX-2 would constitute a novel approach to the treatment of inflammation with fewer side effects (Masferrer et al., 1994). The LD1.0 software was used to build a COX-2-focused library to see whether the optimized focused library contains the marketed and published COX-2 selective drugs or inhibitors.

**Fragment Libraries**

SC58635, an FDA-approved drug for anti-inflammation, is one of the NSAIDs (Graul et al., 1997), which could be divided into three parts as shown in Fig. 5.3 based on its interaction characteristic with the binding site of COX-2. The head of SC58635 is part A, $p$-amino-phenyl-sulfo-phenyl; the tail is part C, $p$-methyl-phenyl; and the body is part B, 3-trifluoro-$1H$-pyrazol-1-yl, which acts as a linkage between parts A and C. Three fragment libraries were constructed for parts A, B, and C with 12, 16, and 4 fragments, respectively.

**Focused Library**

These fragments from different parts could react with each other to produce 768 compounds. Starting from these fragments and based on the

![Figure 5.3](image_url)  
*Figure 5.3* Structure of SC58635 and fragment division for COX-2 inhibitors. See color plates.
X-ray crystallographic structure of COX-2 (PDB entry 6COX), a COX-2-focused library with a population of $3 \times 3 \times 3$ was built. The program terminated normally after running for 28 generations. It took one CPU on an SGI R12000 Origin3800 computer 5.42 hours. Based on the optimized fragments, a focused library containing 27 compounds has been constructed. A preliminary literature survey showed that six of them are the reported COX-2-selective inhibitors (Fig. 5.4) (Marot et al., 2000; Huff et al., 1995). For comparison, a pure molecular docking was performed using DOCK4.0 for the 768 structures, costing more than 40.4 hours. The docking result shows that none of the six experimentally tested COX-2 inhibitors is found via the pure docking. Accordingly, LD1.0 should be a fast and reliable tool for building up a focused library against a specific target structure.

5.2.6.2 Focused Library Design for PPARγ Agonist

The peroxisome proliferator-activated receptors (PPARs) form a subfamily of the nuclear receptor superfamily. Three isoforms, encoded by separate genes, have been identified: PPARγ, PPARα, and PPARδ (also named PPARβ). PPARγ is an important target for the treatment of type II diabetes, and its agonists are therefore expected to be novel drugs for the disease (Michalik and Wahl, 1999; Kliwer et al., 1992). Indeed, thiazolidinediones (TZDs)—for example, Rosiglitazone and Pioglitazone (glitazones), approved by FDA as drugs for type II diabetes—are high-affinity PPARγ agonists. To further verify LD1.0, a test on designing a focused library of PPARγ agonists was performed based on the X-ray crystallographic structure of PPARγ (PDB entry 2PRG) by using LD1.0 (Chen et al., 2005).
It was noticed that most of the PPAR\(\gamma\) agonists can be divided into three parts as shown in Fig. 5.5. Part A is a hydrophilic head, part C is a hydrophobic tail, and part B is a linker between parts A and C. After analyzing a large number of PPAR agonists and referring to other known drug fragment libraries, three fragment libraries were constructed for parts A, B, and C with 118, 88, and 98 fragments, respectively. These fragments could react with each other to give about \(10^6\) structures. Taking into account the feasibility of fast bioassay, a PPAR\(\gamma\)-focused library was designed with a population of \(10 \times 10 \times 10\) by \(3 \times 10\) building blocks.

The calculation conditions and parameters were set the same as those for COX-2-library construction. After running 434 generations, the program LD1.0 terminated normally, and the optimized fragments for parts A, B, and C were reported. Indeed, the parts A and B of TZDs were found among the optimized fragments for parts A and B. However, the tail of TZDs was not found in the optimized building blocks for part C. This is in agreement with the high structural flexibility of the entrance of the PPAR\(\gamma\) binding site. On the other hand, it also hints that there might be some compounds in the library that are more active than the usual TZDs. Indeed, a bioassay experiment on some of the compounds showed that some compounds in the designed focused library are more powerful as agonists than those FDA-approved TZD drugs (unpublished data). This testing example demonstrates again that, for a focused library based on hit and target structures, the LD1.0 software as well as the newly modified descriptor sets for druglikeness, structural diversity, and bioactivity can optimize the building fragments, reduce the library size, and increase the hit rate.

### 5.3 APPLICATION: DISCOVERING NOVEL CHEMICAL INHIBITORS OF CYCLOPHILIN A

Cyclophilin A (hCypA) is one of the most important members of the 15 known human cyclophilins, which play a critical role in a variety of biological processes, including enhancing the rate of folding/unfolding of proteins via its peptidyl-prolyl isomerase (PPIase) activity (Li et al., 2006a,b; Handschumacher et al., 1984; Galat, 2003;
Dornan et al., 2003), binding to the HIV-1 Gag polyprotein for facilitating viral replication (Luban et al., 1993). Therefore, CypA inhibitors are of therapeutic significance. However, inhibitors of CypA are mainly derived from the nature sources (Handschtumacher et al., 1984; Siekierka et al., 1989; Clane et al., 1989; Sedrani et al., 2003) and peptide analogues (Li et al., 2000), which are all large molecules. In order to discover potent small chemical ligands for CypA, a CypA-focused library was constructed by using LD1.0 (Chen et al., 2005).

5.3.1 Focused Library Design

Thirty-five binders of CypA with binding affinities under submicromolar or micromolar level were discovered by docking-based virtual screening targeting the 3D structure of CypA (PDB entry 1NMK), and four of them showed high CypA PPIase inhibition activities with IC_{50}s of 2.5-6.2 μM (Li et al., 2006a,b). According to the binding models of these binders, the skeleton of potential CypA inhibitors is divided into three parts: Part A interacting with the small pocket of CypA (site A), part B is located in the large pocket (site B), and part L interacting with the “saddle” is a linker between A and B (Fig. 5.6). Five fragments for part A, 17 fragments for part B, and three linkers for part L are selected. These fragments from different parts could react with each other to generate 255 molecules. Starting from these fragments and the X-ray structure of the hCypA, the LD1.0 software with default parameters was employed to build an optimized CypA-focused library with a population of \(4 \times 3 \times 12\) (Li et al., 2006).

5.3.2 Chemical Synthesis and Biological Assay

With the guidance of the designed hCypA-focused library, 16 of the 144 designed compounds in the library were synthesized. Surface plasmon resonance (SPR) measurements were used for determining the binding affinity of the 16 molecules to CypA, which reveal that the compounds can bind to CypA in vitro at the micromolar range (\(K_D = 0.076-41.0\) μM, Table 5.4). Thus, the hit rate of the focused library approach is \(~100\%\) in regard to SPR binding. The PPIase activity assessment
TABLE 5.4 Chemical Structures, SPR Binding Affinities, and PPIase Inhibitory Activities of the 5 Compounds Selected from the Focused Library (Li et al., 2006a,b)

<table>
<thead>
<tr>
<th>Compound</th>
<th>Structure</th>
<th>$K_D$ ($\mu$M)</th>
<th>IC$_{50}$ ($\mu$M)$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4a</td>
<td><img src="image" alt="Structure 4a" /></td>
<td>8.19</td>
<td>0.25</td>
</tr>
<tr>
<td>4c</td>
<td><img src="image" alt="Structure 4c" /></td>
<td>41.0</td>
<td>6.43</td>
</tr>
<tr>
<td>4e</td>
<td><img src="image" alt="Structure 4e" /></td>
<td>38.7</td>
<td>2.86</td>
</tr>
<tr>
<td>4f</td>
<td><img src="image" alt="Structure 4f" /></td>
<td>22.8</td>
<td>0.96</td>
</tr>
<tr>
<td>4g</td>
<td><img src="image" alt="Structure 4g" /></td>
<td>35.3</td>
<td>0.78</td>
</tr>
</tbody>
</table>

$^a$Data obtained from SPR-based binding affinity assay. $K_D$ represents the dissociation constant.

$^b$Data are means of three independent experiments.
shows that five compounds (i.e., 4a, 4c, and 4e–g in Table 5.4) have the IC\textsubscript{50} values of 0.25, 6.43, 2.86, 0.96, and 0.78 \(\mu\text{M}\), respectively, demonstrating that the hit rate with the enzymatic inhibitory activity at the micromolar level is as high as \(\sim 31.25\%\). Remarkably, both the binding affinities and inhibitory activities of the compounds from focused library design increase on an average \(\sim 2\) times that of the compounds discovered previously by using virtual screening. This demonstrates the effectiveness of the focused library design approach LD1.0 (Li et al., 2006a,b).

### 5.4 SUMMARY

It could be concluded that the focused library based on hit and target structures is an approach with high efficiency in discovering potent novel drug leads with high specificity against a therapeutic target. The software package LD1.0 with calculation methods integrated for druglikeness, structural diversity, bioactivity, and favorable ADMET properties seems to be powerful in practical terms for new drug development.

### ACKNOWLEDGMENTS

This work was supported by grants from the State Key Program of Basic Research of China (2002CB512802, 2004CB518901) and the Key Program of Science and Technology (2005BA711A04).

### REFERENCES


SSKEYS, MDL Information Systems Inc., San Leandro, CA.


The SPRESI database is produced by the All-Union Institute of Scientific and Technical Information of the Academy of Science of the USSR (VINITI) in Moscow and by the Central Information Processing for Chemistry (ZIC) in Berlin. This database consists of data extracted from 1000 journals and patents, books, and other sources from 1975 to 1990. SPRESI is distributed by Daylight Chemical Information Systems, Inc., Mission Viejo, CA.


PART II

CHEMICAL AND SYNTHETIC APPROACHES IN DRUG DISCOVERY
6.1 INTRODUCTION

Historically, humans have looked to nature as a source of medicinal remedies. In the more recent past, many biologically active natural products were clinically advanced without structural modification because they were too difficult to modify or too little was known about how modifications would improve performance. While such unmodified natural products have occasionally succeeded as drugs (e.g., paclitaxel), it is clear that many natural leads have failed because they are not “designed” for therapeutic use. The increasing effectiveness of synthesis coupled with our growing understanding of therapeutic pathways—advanced by the sequencing of the human genome—is changing this situation. It has now become possible to examine a natural lead with an eye toward synthesis...
and biological function and design a potentially superior clinical candidate that can be synthesized in a practical fashion (function-oriented synthesis) (Wender et al., 2004). Thus, while many natural products will continue to be beyond the reach of practical synthesis, insightful design could allow synthetic access to targets that possess similar, if not superior, activity. The focus of this chapter is to provide an overview on how one such natural product, bryostatin 1, provided the starting point for the design of new leads that exhibit activity superior to the natural product. We report on the background of this effort, the evolution of our function-oriented design strategy, and studies that provide an important approach for evaluating the functional activity of these analogues.

In 1968, extracts from bugula neritina, a marine bryozoan, were found to have potent anticancer activity (Pettit et al., 1970). This activity was ultimately traced to bryostatin 1, whose complex structure was finally determined in 1982 (Fig. 6.1) (Pettit et al., 1982). Subsequently, 18 other naturally occurring bryostatins were identified and characterized. These structures differ primarily at the C7 and C20 positions (Mutter and Wills, 2000). Originally collected in the Gulf of Mexico, bryostatins have since been isolated from bryozoa collected from the Pacific coast of California, the Gulf of California, and the Gulf of Japan. There is also evidence that the bryostatins are produced by some Atlantic bryozoa (McGovern and Hellberg 2003). It has been postulated that the bryostatins are natively produced by symbiotic bacteria in order to protect host larvae from predation (Lopanik et al., 2004).

![Figure 6.1](image_url)  
**Figure 6.1** The natural bryostatin family.
Bryostatin 1 exhibits activity against murine P388 lymphocytic leukemia both \textit{in vitro} and \textit{in vivo} (Table 6.1). While bryostatin 1 is only moderately potent \textit{in vitro}, \textit{in vivo} dosing of bryostatin 1 results in significant life extension with a remarkably low dose of only 10–70 \( \mu \text{g/kg} \). The reason for the discrepancy between the \textit{in vitro} and \textit{in vivo} potency is not understood but might involve other cellular systems and their resulting interactions that are not seen with a single cell line \textit{in vitro}. Bryostatin 1 is currently in Phase I and II clinical trials (descriptions of NIH supported trials can be found at http://www.clinicaltrials.gov). As of this writing, the NIH database shows that 14 clinical trials involving bryostatin are underway, and five new trials are recruiting patients. Although bryostatin has not excelled as a single agent therapeutic, it has been shown to increase the potency and activity of other antineoplastic agents (Mutter and Wills, 2000). Examples include agents that act upon microtubules, such as dolastatin, and auristatin (Mohammad et al., 1998b), the microtubule depolymerizer vincristine (Al-Katib et al., 1998), and the microtubule stabilizer paclitaxel (Wang et al., 1998). Bryostatin also synergizes with DNA replication inhibitors such as arabinosylcytosine (ara-C) (Wang et al., 2002) and 2-chlorodeoxyadenosine (2-CdA) (Beck et al., 2004), DNA cross-linking agents such as cisplatin (Basu and Lazo, 1992), and the estrogen receptor agonist tamoxifen (McGown et al., 1998). In a landmark clinical trial, treatment of cancer of the esophagus with a combination of bryostatin 1 and paclitaxel resulted in a 60% response rate as opposed to the 17% response seen with paclitaxel alone (Marsa, 2002). It has also been observed that the order of dosing is critical in obtaining a synergistic effect, which might lead to new insights about the mechanism of action. In recent clinical trials, bryostatin has performed synergistically with ara-C (Cragg et al., 2002) and vincristine (Dowlati et al., 2003).

Since synergy is observed with a variety of agents whose mechanisms of action differ, it is thought that bryostatin might predispose cells toward apoptosis, resulting in an increased response to various therapeutic strategies (Farrow et al., 2002; Wall

### Table 6.1  \textit{In vivo} and \textit{in vitro} Potency of the Bryostatins Against Murine P388 Lymphocytic Leukemia

<table>
<thead>
<tr>
<th>Bryostatin</th>
<th>( \text{GI}_{50} ) (P388 cells) (( \mu \text{g/mL} ))</th>
<th>Life Extension \textit{in vivo} (Dose)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.89</td>
<td>52–96% (10–70 ( \mu \text{g/kg} ))</td>
</tr>
<tr>
<td>2</td>
<td>N/A</td>
<td>62% (46 ( \mu \text{g/kg} ))</td>
</tr>
<tr>
<td>4</td>
<td>( 10^{-3} )–( 10^{-4} )</td>
<td>62% (46 ( \mu \text{g/kg} ))</td>
</tr>
<tr>
<td>5</td>
<td>( 1.3 \times 10^{-3} )–( 2.6 \times 10^{-4} )</td>
<td>88% (185 ( \mu \text{g/kg} ))</td>
</tr>
<tr>
<td>6</td>
<td>( 3.0 \times 10^{-3} )</td>
<td>82% (185 ( \mu \text{g/kg} ))</td>
</tr>
<tr>
<td>7</td>
<td>( 2.6 \times 10^{-5} )</td>
<td>77% (92 ( \mu \text{g/kg} ))</td>
</tr>
<tr>
<td>10</td>
<td>( 7.6 \times 10^{-4} )</td>
<td>64% (10 ( \mu \text{g/kg} ))</td>
</tr>
<tr>
<td>11</td>
<td>( 1.8 \times 10^{-5} )</td>
<td>64% (92.5 ( \mu \text{g/kg} ))</td>
</tr>
<tr>
<td>12</td>
<td>( 1.4 \times 10^{-2} )</td>
<td>47–68% (30–50 ( \mu \text{g/kg} ))</td>
</tr>
<tr>
<td>13</td>
<td>( 5.4 \times 10^{-3} )</td>
<td>NA\textsuperscript{a}</td>
</tr>
<tr>
<td>16</td>
<td>( 9.3 \times 10^{-3} )</td>
<td>NA</td>
</tr>
<tr>
<td>17</td>
<td>( 1.9 \times 10^{-2} )</td>
<td>NA</td>
</tr>
<tr>
<td>18</td>
<td>( 3.3 \times 10^{-3} )</td>
<td>NA</td>
</tr>
</tbody>
</table>

\textsuperscript{a}NA, not available.
et al., 1999; Mohammad et al., 1998a). Mechanistic studies have shown that bryostatin increases the amount of the proapoptotic protein Bax relative to the antiapoptotic protein Bcl-2. This imbalance results in cytochrome c release, which induces the onset of apoptosis. Bryostatin has also been shown to promote the release of tumor necrosis factor alpha (TNF-α), which can induce apoptosis through an extrinsic pathway (Wang et al., 2003; Cartee et al., 2003).

Another way in which bryostatin potentiates the activity of anticancer agents is in its ability to reverse multidrug resistance (MDR) (Elgie et al., 1998). In HeLa cells, bryostatin 1 is able to inhibit the action of mutant p-glycoprotein (P-gp), an efflux pump responsible for resistance to many cancer therapeutics (Spitaler et al., 1998). An additional study showed reduced expression of P-gp in cells treated with bryostatin (Al-Katib et al., 1998). Since P-gp is overexpressed in resistant cells, it should be possible to increase selectivity in targeting malignant cells.

Bryostatin is also able to differentiate cancer cells into new potentially treatable cell types. Resistance to differentiation is overcome with bryostatin in acute myeloid leukemia cells, allowing differentiation into dendritic-like leukemia cells (Matsui et al., 2005; Roddie et al., 2002).

A drawback to most anticancer therapies is the suppression of immune system activity. In contrast, bryostatin stimulates the immune system in vitro and in vivo (Rabah et al., 2001; Curiel et al., 2001; Oz et al., 2000; Scheid et al., 1994). In various studies it has been shown that bryostatin promotes T-cell production, antigen presentation (Kudinov et al., 2003), dendritic cell activation, and differentiation (Do et al., 2004). Bryostatin has also been effective in adoptive immunotherapy, wherein antigens are generated ex vivo (Parviz et al., 2003; Bear et al., 2001).

The recommended dose of bryostatin from Phase I clinical trials is nothing short of spectacular. Bryostatin is effective when dosed at 25 μg/m² per week or approximately 50 μg per week in an average adult. In comparison, antineoplastic agents that require milligram doses per day are considered potent. For example, an 8-week treatment cycle would require 2.6 g of paclitaxel, whereas a similar treatment would require only 1.1 mg of bryostatin. The dose-limiting toxicity of bryostatin 1 is myalgia (muscle pain); consequently, higher amounts of bryostatin for use as a therapeutic have not been explored.

While the mode of action of the bryostatins is still under investigation, bryostatin’s activity is thought to be mediated through binding to protein kinase C (PKC; \( K_i = 1.4 \text{nM} \)) (Hennings et al., 1987; Wender et al., 1998b), an enzyme family that plays a critical role in cell signaling. PKC is involved in many cellular processes, such as apoptosis, cell-cycle regulation, vesicle transport, gene expression, cellular motility, ion channel regulation, neuronal growth, and cellular differentiation (Dempsey et al., 2000). Abnormal PKC signaling has been identified in many cancer types (Gavielides et al., 2004; Lahn et al., 2004; Mackay and Twelves, 2003; da Rocha et al., 2002; Caponigro et al., 1997). PKC is also a target for the potent tumor-promoting phorbol esters as well as endogenous diacylglycerols (DAGs) (Gomez-Fernandez et al., 2004; Silinsky and Searl, 2003; de Vries et al., 1988). While bryostatins induce a subset of phorbol responses, including activation and subsequent downregulation of PKC, they antagonize other responses, notably tumor promotion (Kraft et al., 1986; Berkow and Kraft, 1985; Ramsdell et al., 1986; Hennings et al., 1987).
PKC isozymes can be grouped into three classes: the calcium-dependent conventional isozymes (\(\alpha, \beta I, \beta II, \gamma\)), the calcium-independent novel isozymes (\(\delta, \varepsilon, \eta, \theta\)), and atypical isozymes (\(\zeta, \iota/\lambda\)). Only the conventional and novel families bind DAG, the bryostatins, and phorbol esters at the C1 domain (Fig. 6.2). This binding results in the modulation of PKC activity and distinguishes bryostatin from many other kinase targeting agents that bind at the ATP binding site or the catalytic site. Ligands that bind to the C1 domain offer a selectivity advantage over ligands for the ATP binding domain of kinases. All kinases (\(\sim 500\)) (Manning et al., 2002) have an ATP binding site, whereas approximately 50 proteins possess a C1 domain.

Although bryostatin holds great promise as a uniquely potent therapeutic agent, its limited supply presents a currently insurmountable obstacle to its widespread therapeutic use. The bryostatins have been isolated from only one source, *Bugula neritina*, in extremely low yields. In the most substantial effort, which has supplied all the material used in clinical trials, 18 g of bryostatin 1 was isolated from 14 tons of sponge, a yield of only \(1.4 \times 10^{-7}\%\) (Schaufelberger et al., 1991). While aquaculture has been explored as a method to overcome difficulties in collection, it does not circumvent the issue of yield (Mendola, 2003). Efforts are underway to identify the genetic mechanisms responsible for bryostatin production with the goal of transplanting such a mechanism to another host for overexpression and subsequent harvesting of bryostatin (Hildebrand et al., 2004). These efforts are important for accessing bryostatin and its natural derivatives. However, it is abundantly clear that bryostatin is not produced for human therapy. Like many natural products, it is an important therapeutic lead but not an optimized therapeutic agent.

**Figure 6.2** The conserved subunits of conventional, novel, and atypical PKCs.

PKC isozymes can be grouped into three classes: the calcium-dependent conventional isozymes (\(\alpha, \beta I, \beta II, \gamma\)), the calcium-independent novel isozymes (\(\delta, \varepsilon, \eta, \theta\)), and atypical isozymes (\(\zeta, \iota/\lambda\)). Only the conventional and novel families bind DAG, the bryostatins, and phorbol esters at the C1 domain (Fig. 6.2). This binding results in the modulation of PKC activity and distinguishes bryostatin from many other kinase targeting agents that bind at the ATP binding site or the catalytic site. Ligands that bind to the C1 domain offer a selectivity advantage over ligands for the ATP binding domain of kinases. All kinases (\(\sim 500\)) (Manning et al., 2002) have an ATP binding site, whereas approximately 50 proteins possess a C1 domain.

Although bryostatin holds great promise as a uniquely potent therapeutic agent, its limited supply presents a currently insurmountable obstacle to its widespread therapeutic use. The bryostatins have been isolated from only one source, *Bugula neritina*, in extremely low yields. In the most substantial effort, which has supplied all the material used in clinical trials, 18 g of bryostatin 1 was isolated from 14 tons of sponge, a yield of only \(1.4 \times 10^{-7}\%\) (Schaufelberger et al., 1991). While aquaculture has been explored as a method to overcome difficulties in collection, it does not circumvent the issue of yield (Mendola, 2003). Efforts are underway to identify the genetic mechanisms responsible for bryostatin production with the goal of transplanting such a mechanism to another host for overexpression and subsequent harvesting of bryostatin (Hildebrand et al., 2004). These efforts are important for accessing bryostatin and its natural derivatives. However, it is abundantly clear that bryostatin is not produced for human therapy. Like many natural products, it is an important therapeutic lead but not an optimized therapeutic agent.
The complexity of the bryostatins has challenged synthetic chemists since the structure of bryostatin 1 was first reported in 1982. Over two decades of work has resulted in the total syntheses of bryostatins 2, 3, and 7 by Evans, Yamamura, and Masamune respectively (Evans et al., 1999; Ohmori et al., 2000; Kageyama et al., 1990). While impressive in content, each synthesis requires over 70 steps and is consequently unlikely to be a cost-effective source. Of course, the main justification for these efforts is to access bryostatin, or a functional equivalent, to pursue its therapeutic value. The identification of functionally equivalent or superior analogues of bryostatin that can be produced cost-effectively represents a solution to the low abundance of the natural product.

### 6.2 BRYOSTATIN ANALOGUE DESIGN

A major intellectual challenge and a key to solving the bryostatin supply and clinical performance problems is to identify those functionalities in bryostatin that are responsible for its activity and to design a molecule with these features that could be synthesized in a practical fashion. Along with bryostatin, several other natural products are known to bind competitively to the C1 domains of PKC. These natural products include the phorbol esters, the endogenous activator DAG, gnidimacrin, aplysia toxin, ingenol, and teleocidin B-4 (Fig. 6.3).

![Figure 6.3 Ligands that bind to the C1 regulatory domains of PKC.](image-url)
Because these ligands all bind to PKC, it was supposed that they might share part or all of a common pharmacophore, a similar three-dimensional orientation of a subset of atoms in each molecule. While this is not a requirement, it provides a working hypothesis that allows for a rational approach to the design of PKC activators and inhibitors. Comparison of the interatomic distances of all possible heteroatomic triads of phorbol and other PKC ligands revealed that the best triad overlay corresponded to the C4, C9, and C20 oxygens of the phorbol esters (Wender et al., 1986). The second-best correlation was with the C3, C9, and C20 triad. The acyl chains, also shown to be required for activity, were spatially oriented in a similar fashion as the lipophilic moieties of the other species. These findings are in agreement with known phorbol structure–activity relationships that include a long acyl chain at C12 or C13, a C20 alcohol, and a C3 ketone and/or C4 hydroxyl group being required for activity (Hecker, 1978).

When the structures employed in elucidating the phorbol pharmacophore were compared to the heteroatomic triads in bryostatin 1, it was found that the C1, C19, and C26 heteroatomic triad of bryostatin had the best overlay with the C4, C9, and C20 heteroatoms of phorbol (RMSD = 0.16 Å) (Fig. 6.4) (Wender et al., 1998b).

![Figure 6.4](image-url)  
*Figure 6.4* The preferred pharmacophoric correlation of 1,2-diacyl-sn-glycerol, the phorbol esters, and bryostatin 1. See color plates.
Moreover, when the correlated atoms were compared, their lone pair electrons or attached hydrogen atoms assumed a similar orientation. Structure–activity relationships established using the naturally occurring bryostatins substantiated this hypothesis, as well as additional information concerning which elements of the molecular architecture are essential for potency. These studies revealed that acylation of C26 resulted in loss of potency (Wender et al., 1998b) and suggests that the C3 hydroxyl is involved in a transannular hydrogen bonding network that holds the macrocycle in an active conformation (Pettit et al., 1982, 1991; Kamano et al., 1996). Also, while hydrogenation of the C13–C30, C2–C3, and C4–C5 olefins did not affect potency, hydrogenation of the C21–C34 olefin resulted in a significant potency loss (Pettit et al., 1992).

Using the pharmacophore hypothesis, it became possible to propose structurally simplified analogues of bryostatin that should have similar activity. Given that alterations to the A- and B-rings of bryostatin had little effect on potency (Pettit et al., 1991; Kamano et al., 1996), it was proposed that these rings could be simplified while maintaining the rigidity of the macrocycle. In addition, saturation of the C20 side-chain was proposed based on the previously discussed hydrogenation of the natural product. Finally, introduction of an additional B-ring oxygen was anticipated to simplify the synthesis and make it convergent. The ability to simultaneously design for function and select for ease of synthesis was key to advancing this effort. Structure 1 emerged from this analysis (Scheme 6.1) (Wender et al., 1998b). Its synthesis was planned to arise from the assembly and coupling of two fragments, the spacer and recognition domains, followed by a proposed macro-transacetalization process. Completion of the synthesis allowed for a major test of the pharmacophore hypothesis. Of great significance to this approach, compound 1 binds to PKC ($K_i = 3.4 \text{nM}$) with an affinity comparable to bryostatin 1.

The synthesis of analogue 1 allowed for the preparation of additional compounds (2–6, Fig. 6.5) to probe the structural requirements for analog binding. It is noteworthy that the C3 hydroxyl plays a significant role in PKC binding. NMR studies suggest that it forms a bifurcated transannular hydrogen-bond array, presumably preorganizing the macrocycle for binding. Deletion of the C3 hydroxyl or inversion of the C3 stereochemistry reduced affinity by two orders of magnitude. Acylation or oxidation of the C26 alcohol resulted in a complete loss of PKC binding. Deletion of the macrocycle also resulted in total abrogation of potency (Wender et al., 1999).
6.3 THE DESIGN AND SYNTHESIS OF A C26 DES-METHYL ANALOGUE

When compared to other PKC ligands, the hydroxyl at C26 in bryostatin corresponds to the hydroxyl groups at C20 in phorbol, C3 in DAG, and C24 in teleocidin (Fig. 6.3). However, these counterparts to the C26 hydroxyl are primary rather than secondary alcohols. An analogue possessing a C26 primary alcohol was anticipated to retain biological activity while reducing molecular complexity by removal of a stereocenter. Molecular modeling of the proposed C26 des-methyl analogue (15, Scheme 6.2) suggested a global minimum that closely resembled the crystal structures of the bryostatins. Analogue 15 was prepared from an advanced intermediate in the synthesis of analogue 1 (Scheme 6.2) (Lippa, 1999). Starting from intermediate 7, hydrogenation followed by oxidative cleavage and olefination of the resultant aldehyde afforded terminal olefin 9. This three-step sequence successfully deleted the C26 methyl group. Following conversion of the C17 silyl ether to aldehyde 11, the C25–C26 diol was reintroduced via Sharpless asymmetric dihydroxylation. Protection as the bis-TES ether provided intermediate 12. After a four-step homologation of the C17 aldehyde to enal 13, treatment with aqueous HF simultaneously cleaved the silyl ethers while hydrolyzing the methyl ketal. The C26 primary hydroxyl group was selectively protected as the TBS ether. The completed recognition domain 14 was coupled to the completed spacer domain, 21 (vide infra), via a Yamaguchi esterification. Finally, treatment with HF-pyridine effected a macrotransacetalization to close the macrocycle and set the C15 stereocenter while simultaneously cleaving the two silyl ethers to provide the completed analogue 15.

In agreement with prediction, the C26 des-methyl analogue displays picomolar affinity for PKC with a $K_i$ of 250 pM. When tested in vitro against various human

![Figure 6.5](image-url)
15 displayed greater potency than bryostatin 1 in 24 of 35 cases. In some cell lines, such as MOLT-4 and NCI-H460, 15 is three orders of magnitude more potent than bryostatin 1 at inhibiting cell growth (see Section 6.6). The exceptional potency of 15 prompted the development of a more efficient synthesis to address the supply problem and provide access to potentially superior clinical candidates (Wender et al., 2002). Building on the existing synthesis of the C1–C13 spacer domain, 21, the synthesis began with the ozonolysis of 16, which, after in situ reduction, produced pentane-1,3,5-triol (Scheme 6.3). Desymmetrization of the resulting triol via acetal formation with \((/C0)-menthone\) (Harada et al., 1993; Moune et al., 1997) and subsequent oxidation yielded aldehyde 17. A hetero Diels–Alder cycloaddition between 17 and Danishefsky’s diene using Jacobsen’s tridentate Cr(III) catalyst (Joly and Jacobsen, 2002) provided pyranone 18 with exceptional selectivity. Luche reduction, vinyl ether formation, and subsequent Claisen rearrangement provided the unsaturated aldehyde 19. Hydrogenation of 19 followed by asymmetric allylation of the aldehyde and silylation of the resultant alcohol afforded silyl ether 20. Oxidative cleavage of the terminal olefin provided spacer domain 21. The overall conversion of 16 to 21 required 11 steps and resulted in an overall yield of 11%.

A shorter, more selective synthesis of the spacer domain was subsequently developed to provide 30. In this approach, the \((-)-\)menthione cyclic ketal was replaced with an acetone-derived ketal, and the C3 TBS ether was replaced with a TBDPS ether (Scheme 6.4) (Wender et al., 2003b). The synthesis of 30 began with the...
Scheme 6.3  An efficient synthesis of spacer domain 21. (a) i. O₃, MeOH, −78°C, ii. NaN₄, −78°C → rt, 90%; (b) (−)-menthone, pTsOH·H₂O, CH(OE)₃, Et₂O, rt, 71%; (c) (COCl)₂, DMSO, Et₃N, CH₂Cl₂, −78°C, 87% (β:α = 1.6:1); (d) i. Danishefsky’s diene, Jacobsen’s tridentate Cr(III) catalyst, 4 Å MS, acetone, rt, ii. TFA, rt, 88%; (e) NaBH₄, CeCl₃·7H₂O, MeOH, −40°C, 92%; (f) Hg(OAc)₂, isobutylvinyl ether, rt; (g) decane, 150°C, 83% over two steps; (h) H₂, Pd(OH)₂/C, EtOAc, rt, 85%; (i) (−)-(Ipc)₂BOMe, allyl-MgBr, CH₂Cl₂, −78°C → rt; (j) TBSCI, imidazole, THF, rt, 69% over two steps; (k) KMnO₄, NaIO₄, 1:1 t-BuOH:pH 7 buffer, rt, 84%.

Scheme 6.4  Selective synthesis of spacer domain 30. (a) i. LDA, THF, 23, −78°C, ii. 22, 68%; (b) Ru-(S)-BINAPCl₂, MeOH, H₂ (95 atm), 30°C, 92%; (c) silica, PhMe, reflux, 95%; (d) TBDPSCl, imidazole, DMF, rt, 85%; (e) ethylacetocetate, LDA, THF, −78°C; (f) Et₃SiH, TFA, −30°C, 70% over two steps; (g) Ru-(R)-BINAPCl₂, EtOH, H₂ (78 atm), rt, 91%; (h) i. H₂ (1 atm), Pd(OH)₂/C, Et₂O, rt, ii. LiBH₄, rt, 96%; (i) i. 2,2-dimethoxypropane, pTsOH, DMF, rt, ii. silica, CH₂Cl₂, 93%; (j) TEMPO, NaOCl, NaClO₂, CH₃CN, 50°C, 92%.
acylation of 4-benzyloxy-2-butanone 23 with acid chloride 22 to provide diketone 24. A Noyori asymmetric hydrogenation was employed to set the C3 and C5 stereocenters in excellent yield and in greater than 95:5 enantioselectivity to provide diol 25 (Kitamura et al., 1988; Noyori et al., 1987). The secondary alcohols were then differentiated through lactonization of the C5 alcohol with the proximate ester. Protection of the remaining alcohol gave lactone 26. The remainder of the carbon backbone (C10–C13) of the spacer domain was then introduced by addition of the dianion of ethyl acetooacetate to lactone 26. The resulting lactol was reduced to set the C9 stereocenter, and a second asymmetric reduction converted β-ketoester 27 to alcohol 28. The C1 benzyl ether was cleaved and the ethyl ester reduced in one operation to give the corresponding triol. The 1,3-diol was protected as acetonide 29 and the primary alcohol oxidized to the carboxylic acid (Zhao et al., 1999), yielding spacer domain 30 in 10 steps and 25% overall yield.

Coordinated with the synthesis of the spacer domain, the synthesis of the C15–C26 recognition domain 14 began with the monoprotection and oxidation of inexpensive diol 31 to generate aldehyde 32 (Scheme 6.5). Reaction of 32 with the Grignard reagent derived from 4-chloro-1-butanol (Cahiez et al., 1978) followed by oxidation and asymmetric Keck allylation (Keck and Krishnamurthy, 1998; Yu et al., 1997) provided homoallylic alcohol 33 and set the C23 stereocenter. Dehydrative cyclization of 33 followed by epoxidation and in situ methanolysis yielded a mixture of diastereomers, the major of which was oxidized to ketone 34. Conversion of ketone 34 to enoate 35 was readily accomplished in one step. A highly diastereoselective reduction of 35 followed by esterification afforded ester 37. Deprotection of 37 followed by oxidation gave aldehyde 38. In contrast to our original four-step sequence, the transformation of aldehyde 38 to enal 39 was achieved in only one step using a vinyl zincate reagent. Sharpless asymmetric dihydroxylation followed by hydrolysis of the C19 ketal and selective protection of the C26 alcohol provided recognition domain 14 in 17 steps and 3% overall yield.

Synthesis of analogue 15 was completed by coupling 14 and 21 using the Yamaguchi esterification protocol followed by macrotransacetalization (Scheme 6.2). This transacetalization procedure closed the macrocycle, set the C15 stereocenter under thermodynamic control, and allowed for global deprotection, providing analogue 15 in 19 steps (longest linear sequence) and 2% yield. The spacer domain 30 was also coupled to fragment 14 using a PyBroP-mediated esterification to provide the corresponding ester, which was treated with HF·pyridine to again close the macrocycle via a transacetalization and to cleave the protecting groups. Both syntheses provide novel strategies for the practical and scalable synthesis of 15. Additionally, the convergent nature of these routes is ideal for the rapid and efficient synthesis of related analogues.

### 6.4 BRYOSTATIN ANALOGUES WITH MODIFIED SPACER DOMAINS

While investigating conditions to diversify the completed analogs, it was found that treatment of 15 with dimethyldioxirane (DMDO) resulted in selective oxygenation at
C9 (Scheme 6.6) (Wender et al., 2005b). This highly selective C–H insertion reaction effectively reinstalled the C9 hemiketal present in the natural product without affecting other sensitive functionality in the molecule such as the C16–C17 olefin, the C15 and C19 acetals, or the C26 alcohol. Similar reactivity was observed when 1 was subjected to the same reaction conditions. However, partial oxidation of the C26 alcohol to the corresponding ketones 5 and 44 was observed with this substrate. This novel transformation produced new bryostatin analogues in a single step and represents one of the most complex intermolecular C–H activations reported to date. According to the proposed pharmacophore, the hemiketal functionality at C9 is not essential for PKC affinity (Wender et al., 1998b). This hypothesis is supported by the observation that analogues 1 and 15, which lack a C9–OH group, bind PKC as well or better than bryostatin 1. Analogues 42 and 43, along with their nonhydroxylated variants, 1 and 15, allowed the effect of the C9 hemiketal on affinity for PKC to

![Chemical structures and reactions](image-url)
be examined. The binding constants for new analogues 42 and 43 are 2 nM and 4 nM, respectively. Consistent with the proposed pharmacophore, the potencies of these compounds are similar to those found for 1 and 15. These data serve as a negative control further supporting the pharmacophore hypothesis.

The potent binding affinities of analogues 1 and 15 corroborate the hypothesis that the top portion serves as a lipophilic spacer domain and holds the recognition elements in the proper orientation for binding to PKC. These agents demonstrate that it is possible to dramatically reduce complexity in the spacer domain without adversely affecting binding affinity. These structural simplifications provide a significant increase in the synthetic accessibility of the analogues compared to the natural product.

While the existing spacer domains 21 and 30 do not contain much of the functionality present in the natural product, they still possess four stereocenters and require 10 or 11 steps to synthesize. Consequently, further simplifications of the spacer domain were explored. The A-ring tetrahydropyran, devoid of pharmacophoric elements, was targeted for further structural modification.

The greatest simplification of the A-ring is the removal of carbons C6 to C8, simplifying the structure dramatically by removal of two stereocenters. However, this approach was expected to increase the conformational flexibility of the spacer domain. In order to confer a greater level of conformational rigidity, a sterically demanding substituent was introduced at the C9 position in place of the A-ring. These design concepts were realized in the synthesis of four new simplified spacer domains (Fig. 6.6).

The synthesis of these spacer domains began with oxidation of alcohol 49 to aldehyde 17β (Wender et al., 1998a, 2004, Wender 2005a). Addition of the desired nucleophile followed by installation of the correct stereochemistry gave
alcohols 51 and 53. Preparation of intermediate 55 started with an asymmetric Brown’s allylation of 17β followed by TBS protection to give silyl ether 54. Cross metathesis with 4-bromostyrene and subsequent reduction of the olefin produced silyl ether 55.

After desilylation of 55, each of the secondary alcohols was carried independently through a parallel synthetic sequence to complete the individual spacer domains (Scheme 6.8). Alkylation with allyl bromide gave terminal olefins 56–59. Hydroboration followed by oxidation gave aldehydes 60–63. Asymmetric allylation and subsequent TBS protection allowed for isolation of the silyl ethers 64, 65, 66, and 68. Oxidative cleavage of the terminal olefins gave the completed spacer domains 45–48.

Scheme 6.7 The synthesis of intermediates 51, 53, and 55. (a) Dess–Martin periodinane, CH₂Cl₂, rt, 90%; (b) t-BuLi, Et₂O, −78°C, 58% (1:1 mixture of 50 and 51); (c) TPAP, NMO, 4 Å MS, CH₂Cl₂, rt; (d) NaBH₄, CeCl₃•7H₂O, MeOH, −78°C → rt, 87% (over two steps; 9:1 of 51 and 50); (e) PhMgBr, CH₂Cl₂, −78°C → rt; (f) Dess–Martin periodinane, CH₂Cl₂, rt, 84% (over two steps); (g) NaBH₄, CeCl₃•7H₂O, MeOH, −78°C → rt, 78%; (h) (−)-(Ipc)₂BOMe, allylMgBr, Et₂O, −78°C → rt, 94%; (i) TBSCI, imidazole, DMF, rt, 91%; (j) Grubbs’ second generation catalyst, 4-bromostyrene, 1,2-DCE, 40°C, 59%; (k) 5% Rh/Al₂O₃, H₂ (1 atm), EtOH, rt, 73%.
A second strategy toward simplified spacer domains focused on synthetic accessibility and structural simplification. Given that the syntheses of spacer domains 21 and 45–48 were linear sequences of 10–16 steps, it would be beneficial to develop a route where two or more complex pieces could be coupled together to complete the molecule. It was envisioned that replacing the ether linkage in the A-ring-modified spacer domains with an ester linkage would increase the convergency of the synthesis (Scheme 6.9) (Wender and Lippa, 2000).

The preparation of this family of spacer domains began with synthesis of secondary alcohol 71. Prenylation of 17β gave a mixture of alcohols, which was recycled through an oxidation–reduction sequence to give 72β. Reductive ozonolysis

Scheme 6.8 Completion of spacer domains 45–48. (a, 55 only) TBAF, THF, rt; (b) t-BuOK, allyl bromide, THF, rt; (c) i. 9-BBN, THF, rt, ii. NaOH, H2O2; (d) Dess–Martin periodinane, CH2Cl2, rt; (e) (–)-(Ipc)2BOMe, allylMgBr, CH2Cl2, 0°C → rt; (f) TBSCl, imidazole, CH2Cl2 or DMF, rt; (g) TBAF, THF, rt; (h) TBSCl, imidazole, CH2Cl2, rt; (i) NaIO4, KMnO4, 1:1 t-BuOH:pH 7 buffer, rt.

Scheme 6.9 The design of ester-linked A-ring spacer domains.
of the terminal olefin followed immediately by esterification with chloroacetic anhydride yielded the selectively protected derivative 71. The chloroacetate group was chosen as a temporary protecting group that could be later substituted with a variety of functionally diverse groups. The secondary alcohols 71 and 51 were then independently taken through the remaining two steps of the sequence. Yamaguchi esterification with known acid 73 followed by cleavage of the chiral auxiliary afforded the completed spacer domains 69 and 70 (Scheme 6.10).

The six A-ring-modified spacer domains reviewed in this chapter were synthesized over the course of several years. A consequence of this was that some of these spacer domains were coupled to the C26 methyl recognition domain 76, used to construct analogue 1, while those synthesized at a later time were coupled to the C26 des-methyl recognition domain 14 to complete the final analogues.

Of the spacer domains with the ether linkage, only 45 and 46 were coupled to 76 (Scheme 6.11) (Wender et al., 1998a). The coupling of these domains was accomplished using Yamaguchi’s esterification protocol. The C3 silyl group was removed and the macrocyclic acetal was closed under acidic conditions. Hydrogenolysis of the C26 benzyl ether gave the completed bryostatin analogues 79 and 80.

Spacer domains 69 and 70, with the ester linkage, were also coupled with recognition domain 76 in a similar manner (Scheme 6.12). The chloroacetate of 84 was then removed by treatment with excess thiourea.

The C26 benzyl ethers were deprotected using the standard conditions to provide three new analogues: 86, 87, and 88 (Scheme 6.13).

Intermediate 85 was further diversified before removing the benzyl group to provide access to various esters in the A-ring region of the molecule (Scheme 6.14).
Scheme 6.11  Completion of C26 methyl analogues 79 and 80. (a) 2,4,6-Trichlorobenzoyl chloride, Et₃N, DMAP, CH₂Cl₂; (b) HF·pyridine, THF; (c) Amberlyst-15 resin, CH₂Cl₂; (d) Pd(OH)₂/C, H₂ (1 atm), EtOAc.

Scheme 6.12  Coupling and closure steps for ester-linked analogues. (a) 2,4,6-Trichlorobenzoyl chloride, Et₃N, DMAP, MePh; (b) excess HF·pyridine, THF; (c) excess thiourea, THF, 4 days.

Scheme 6.13  Completion of C26 methyl analogues 86, 87, and 88.
this manner the myristate (89) and pivalate (90) analogues were generated. The 2,4,6-trichlorophenyl ester analogue 91 was generated as a byproduct in the second reaction.

All compounds exhibit high affinity for PKC (Fig. 6.7). Gratifyingly, most of the compounds show only slightly diminished activity compared with the parent compound 1 (3.4 nM) or the natural product (1.4 nM).

Attempts to couple the ester-based spacer domain 70 to the picolog recognition domain 14 to generate the des-methyl version of analogue 86 were hampered by decomposition of the material during purification. However, no stability issues were

Figure 6.7 Binding affinities of C26-methyl analogues for a PKC isozyme mixture.
observed when spacer domains 45–48 were used in place of 70. The spacer domains were coupled to the recognition domain and the macrocycle closed in the manner described previously to generate four new C26 des-methyl bryostatin analogues (Scheme 6.15).

All analogues exhibit single-digit nanomolar affinity for PKC (Fig. 6.8). While these A-ring modified analogues are roughly an order of magnitude less potent than the parent compound 15 in this assay, the activities are still very high. It is

Figure 6.8  Binding affinity of C26 des-methyl analogues for PKC isozyme mixture.
interesting to note that removal of the C26 methyl group from analogue 1 resulted in a 10-fold increase in binding affinity. Similarly, removing the C26 methyl from compound 79 resulted in a six-fold increase in potency. However, removal of the same methyl group from compound 80 only resulted in a negligible change in binding affinity.

6.5 BRYOSTATIN ANALOGUES WITH MODIFIED RECOGNITION DOMAINS

The fundamental understanding that the key pharmacophoric substructures of bryostatin are contained in the recognition domain has led to a focus on spacer domain modifications while leaving most of the original recognition domain of bryostatin 1 intact, in order to tune potency and function. However, early structure–activity relationship studies of the natural products and closely related derivatives (Wender et al., 1998b; Pettit et al., 1992) suggest strategies for the design of recognition domain analogues that incorporate more significant changes. Comparison of the PKC binding affinities of bryostatins 1, 2, 4–11, and 18 (Fig. 6.1) shows that variations in the C20 ester have only a modest effect on potency. A derivative in which both the C13 and C21 enoates were hydrogenated shows greatly reduced potency as compared to a derivative in which only the C13 enoate was hydrogenated, suggesting that unsaturation in the C-ring is important for PKC affinity. Bryostatins lacking the C19 hydroxyl group and the C20 side chain (bryostatins 16 and 17) show decreased potency. These initial comparisons, while crucial to early directions in analogue design, do not address a number of issues such as the role of the C21 substituent and the range of tolerated C20 side chains. A hypothesis for the necessity of the C21 enoate for potency is that it serves as a conformational constraint on the C-ring. Consequently, a synthetically simplified recognition domain should either retain unsaturation at C21 or contain an alternative conformational constraint. In analogue 104, a C20–C21 carbonate (Scheme 6.16) is such an alternative constraint.

The synthesis of the desired analogue started from pyranone 96 (Wender et al., 2003a; Koehler, 1999), which was diastereoselectively converted into diol 97 (Scheme 6.16). Treatment of 97 with triphosgene generated the carbonate 98. Oxidative deprotection of the PMB groups followed by allylic oxidation provided enal 100. Hydrolysis of the mixed ketal provided recognition domain 101. Coupling to spacer domain 21 followed by TBS removal, macrotransacetalization, and hydrogenation of the benzyl group provided completed analogue 104. This analogue shows substantially reduced affinity for PKC ($K_i > 10 \mu M$), suggesting that unsaturation might be required for potency.

Subsequent studies on analogues with simplified recognition domains focused on preserving unsaturation at C21 while modifying the C34 ester functionality. Analogues without the ester moiety could prove to be more stable in vivo. Furthermore, it was envisioned that modifications at C34 could allow for rapid diversification at a late stage. With these issues in mind, analogues 105 (incorporating an aryl group) and 106 (incorporating a benzyl ester that could be further elaborated) were targeted (Fig. 6.9) (Pattabiramen, 2004).
Analogue 105 was synthesized starting from ketone 34 (Scheme 6.17), an advanced intermediate in the synthesis of 15. Sharpless asymmetric dihydroxylation followed by protection as an acetonide provided 108. Condensation with benzaldehyde provided enone 109, which was reduced and esterified to provide 110. Elaboration to the completed recognition domain 113 using conditions developed for the synthesis of 15 proceeded smoothly. Coupling of 113 with spacer domain 21 following the standard conditions provided the completed analogue 105. This new analogue has a binding affinity of 100 nM.

Synthesis of analogue 106 began from ketone 34 (Scheme 6.18). Condensation with glyoxylate followed by esterification with benzyl chloroformate gave enone 114. Using conditions developed for the synthesis of 15, 114 was elaborated to recognition domain 120. Coupling of recognition domain 120 with spacer domain 21 followed by closure of the macrocycle provided the completed analogue 106. This new analogue exhibits a PKC affinity of 120 nM.

These analogues demonstrate the importance of the C21 enoate. Removal (104) or modification (105 and 106) of this functionality results in a significant loss of PKC.
Figure 6.9  Benzylidene and benzoate analogues.

Scheme 6.17  Synthesis of a C21 benzylidene analogue. (a) (DHQD)$_2$PYR, K$_2$OsO$_2$(OH)$_4$, K$_3$Fe(CN)$_6$, K$_2$CO$_3$, $t$-BuOH/H$_2$O, 4°C; (b) dimethoxypropane, PPTS, DMF, rt, 70% (over two steps); (C) PhCHO, K$_2$CO$_3$, MeOH, 55%; (d) NaBH$_4$, CeCl$_3$·7H$_2$O, MeOH, −30°C; (e) octanoic acid, DIC, DMAP, CH$_2$Cl$_2$, rt, 75% (over two steps); (f) 3HF·NEt$_3$, THF, rt; (g) DMP, Na$_2$CO$_3$, CH$_2$Cl$_2$, 0°C → rt, 90% over two steps; (h) (Z)-1 bromo-2-ethoxyethene, $t$-BuLi, Me$_2$Zn, Et$_2$O, −78°C; (i) pTsOH, MeCN/H$_2$O, rt, 55% (over two steps); (j) TBSCl, imidazole, CH$_2$Cl$_2$, rt, 85%; (k) 21, PyBroP, DIEA, DMAP, CH$_2$Cl$_2$; (l) HF·pyridine, THF, rt, 80% over two steps.
affinity. While bulky C34 substituents are detrimental, smaller substituents in this region might be better tolerated.

The structure–activity analysis of the natural bryostatins (Fig. 6.1) and their semisynthetic analogues suggests that changes could be made to the C20 substituent without sacrificing efficacy in vivo. Additionally, computer modeling and pharmacophore analysis of bryostatin and synthetic analogues suggest that the C20 substituent might correspond to the sn-1 substituent of DAG and the C12 substituent of the phorbol esters. These substituents have been shown to affect the selectivity of PKC activation for both DAG and the phorbol esters (Madani et al., 2001; Wada et al., 2002; Wang et al., 1999). A systematic investigation of C20 substituents would enhance the ability to tune the biological activity of analogues.

To explore possible correlations between the carbon content of the C20 ester and affinity to PKC, functionalization of the C20 position was studied (Wender and Hinkle, 2000). Following protection of enal 121 as a dimethyl acetal, the octanoate ester was hydrolyzed to the corresponding alcohol (Scheme 6.19). Acylation followed by PMB deprotection and cleavage of both acetals provided recognition domain 123. Coupling to spacer domain 21 followed by silyl deprotection, macrotransacetalization,

Scheme 6.18  Synthesis of a C21 benzoate analogue. (a) Glyoxylic acid, NaOH, THF/H2O, rt; (b) benzyl chloroformate, DMAP, CH2Cl2, rt, 82% over two steps; (c) NaBH4, CeCl3 · 7H2O, MeOH, −30°C; (d) octanoic acid, DIC, DMAP, CH2Cl2, rt, 69% over two steps; (e) 3HF · NEt3, THF; (f) DMP, Na2CO3, CH2Cl2, rt, 70% over two steps; (g) (Z)-1-bromo-2-ethoxyethene, t-BuLi, Me2Zn, Et2O, −78°C to rt, 63%; (h) (DHQD)2PYR, K2OsO2(OH)4, K3Fe(CN)6, K2CO3, t-BuOH/H2O, 4°C, 61%; (i) pTsOH, MeCN/H2O, rt, 64%; (j) TBSCl, imidazole, CH2Cl2, rt, 76%; (k) 21, PyBroP, DIEA, DMAP, CH2Cl2, rt; (l) HF · pyridine, THF, rt, 80% over two steps.
and hydrogenation of the C26 benzyl protecting group provided analogue 125, containing a C20 acetate.

Analogues 126 and 127 (Fig. 6.10) were prepared using a similar sequence. They have affinities for PKC similar to that of 1. Within this series, it is evident that there is a positive correlation between carbon content of the C20 ester and PKC binding affinity. However, replacing the octanoate ester with a benzoate ester does not affect binding affinity, suggesting that diverse structural types might be tolerated.

The success of these efforts prompted the further investigation of modifications to the C20 position (Wender and Baryza, 2005). To efficiently explore variations at the

![Scheme 6.19 Synthesis of C20 acetate analogue 125. (a) CH(OMe)3, PPTS, MeOH, rt; (b) K2CO3, MeOH, rt; (c) Ac2O, DMAP, CH2Cl2, rt, 82% for three steps; (d) DDQ, CH2Cl2/H2O, rt, 91%; (e) aq. HF, CH3CN, rt, 89%; (f) 2,4,6-trichlorobenzoyl chloride, Et3N, DMAP, then 21, CH2Cl2, rt, 63%; (g) HF-pyridine, CH3CN, rt, 82%; (h) amberlyst-15 resin, CH2Cl2, rt, 83%; (i) Pd(OH)2/C, H2 (1 atm), EtOAc, 93%.](image)

![Figure 6.10 PKC binding constants for first-generation C20 ester analogues.](image)
C20 position, a functional handle was incorporated into the molecule that could be selectively diversified as late as the final analogue. An aniline, which could be revealed by selective reduction of an aryl nitro group, was chosen as the diversification handle (134, Scheme 6.20). The 3-nitro-benzoate was introduced by

Scheme 6.20 Preparation of aryl amine for late-stage diversification. (a) NaBH₄, CeCl₃·7H₂O, MeOH, −30°C; (b) 3-(NO₂)C₆H₄CO₂H, DIC, DMAP, CH₂Cl₂, rt, 75% over two steps; (c) 3HF·NEt₃, CH₂Cl₂, rt, 58%; (d) DMP, NaHCO₃, CH₂Cl₂, rt, 86%; (e) K₂OsO₂(OH)₄, (DHQD)₂PYR, K₃Fe(CN)₆, K₂CO₃, H₂O, t-BuOH, 4°C, 67% (2:1 β:α); (f) Et₃SiCl, imidazole, CH₂Cl₂, rt, 98%; (g) Et₂BOMe, allylMgCl, Et₂O, rt; (h) Ac₂O, DMAP, CH₂Cl₂, rt, 67% over two steps; (i) OsO₄, NMO, H₂O, THF, rt; (j) i. Pb(OAc)₄, PhMe, ii. DBU, rt, 60% over two steps; (k) pTsOH·H₂O, H₂O, MeCN, rt; (l) TBSCI, imidazole, CH₂Cl₂, rt, 43% over two steps; (m) 21, PyBroP, i-Pr₂NEt, DMAP, CH₂Cl₂, rt, 75%; (n) HF·pyridine, THF, rt, 80%; (o) H₂ (1 atm), 10% Pd/C, EtOAc, 99%.
esterification to provide 128. This compound was elaborated to the corresponding recognition domain through a sequence similar to that used for the synthesis of 15. Coupling with spacer domain 21 and subsequent one-step deprotection/macrotransacetalization gave nitroaryl analogue 133, which was selectively hydrogenated to aniline 134.

Reaction of aniline 134 with a variety of anhydrides resulted in exclusive acylation of the aniline moiety. Using this strategy, an initial set of analogues was prepared (Scheme 6.21). All these analogues have low nanomolar affinity for PKC. The general trend for increased potency with increasing carbon count seems to hold, while many of the more polar groups are well-tolerated.

### 6.6 BIOLOGICAL STUDIES ON ANALOGUES OF BRYOSTATIN

It is important that designed analogues of bryostatin possess not only the affinity, but also the function of the natural product. Determination of binding affinity is an important first screen to eliminate compounds that are unlikely to act on the target of interest. However, binding potency in a cell-free environment does not necessarily correlate to cellular or \textit{in vivo} activity.

An important biological activity of bryostatin is its anticancer activity; therefore, analogues 1 and 15 were sent to the NCI for testing against the 60 cell-line panel used for profiling new anticancer agents. In many cell lines, both analogue 1 and

![Scheme 6.21](image-url)
analogue 15 outperformed bryostatin 1 at inhibiting the growth of cancer cells, in some cases by up to three orders of magnitude (Fig. 6.11) (Wender et al., 2002, Wender et al., 2004).

Bryostatin 1 is known to activate PKC, an activity likely to be responsible for at least some of the therapeutic effects it displays. Other C1 domain containing protein families are also likely to be involved (Kazanietz, 2002; Brose et al., 2004), but the activation and translocation of PKC to cellular membranes is a specific cellular effect of bryostatin treatment. Therefore, a PKC translocation assay was implemented as a way to measure potency and activity of bryostatin analogues in living cells.

Activation of PKC is tightly coupled to the location of the protein. Inactive PKC is kept in various intracellular locations. Activating cofactors increase the membrane affinity of the kinase and induce its translocation to cellular membranes. Membrane association drives the release of an autoinhibitory sequence from the kinase active site, which turns on signaling by the kinase. This union of translocation and activation has led to the widespread use of translocation as a surrogate measure for PKC activation.

For the measurement of PKC translocation, plasmids encoding for a fusion protein of PKC and green fluorescent protein (GFP) are expressed in living cells (Fig. 6.12). By treating these cells with bryostatin or bryostatin analogues and observing the localization of the fluorescence with confocal microscopy, the response of the protein can be observed in real time. Analysis of the images allows for quantification of the rate and degree of response as well as the final localization of the protein. As PKC translocates to the membrane, fluorescence in
the cytosol decreases, providing plots of the type shown in Fig. 6.12. The use of isoyme specific constructs allows for the response of individual PKC isoymes to be measured (Baryza et al., 2004).

Comparison of the response of PKCδ to bryostatin and to analogues 1 and 15 confirmed that the analogues are more potent than the natural product at inducing translocation of this isoyme. While bryostatin 1 failed to induce translocation of PKCδ at doses below 5 nM, analogue 1 still possessed activity at 1 nM, and analogue 15 maintained the ability to induce translocation even at 0.10 nM concentration (Fig. 6.13).

The ability of the analogues to induce translocation of PKCδ contrasts with their binding affinities for PKC. While this is not unexpected, the findings have led to an increased interest in testing a larger set of compounds against a wider range of PKC isoymes.

Over the last several years, aided by the sequencing of the human genome, other families of proteins containing C1 domains have been identified, including RasGRP (Caloca et al., 2003), the chimaerins (Hall et al., 2005), the Munc proteins (Betz et al., 1998), diacylglycerol kinase (DGK) (van Blitterswijk and Houssa, 2000), and protein kinase D (PKD) (Rozengurt et al., 2005). Where tested, bryostatin has been effective at activating these proteins as well (Lorenzo et al., 2000; Rhee et al., 2002; Shindo et al., 2003; Matthews et al., 1997). It is likely that some of the effects previously ascribed to PKC are in fact caused by these proteins.

Activation of RasGRP appears to be responsible for some of the immune system effects previously attributed to PKC. RasGRP is a protein family that is activated by translocation to the plasma membrane, where it activates Ras. RasGRPl appears to play a critical role in T-cell receptor signaling. When analogues 1 and 15 were tested in assays measuring RasGRPl activation, the compounds displayed the
same activity as bryostatin. The analogues, as well as bryostatin, induced RasGRP1 activation, (as measured by both RasGRP1 translocation and ERK phosphorylation), CD69 expression (a marker for T-cell activation), and thymocyte aggregation (Stone et al., 2004).

Bryostatin and the phorbol esters both have immunostimulatory properties. In vivo, immune system activation is possibly responsible for some of the anticancer activity of bryostatin. In a study of immune cell activation, it was shown that phorbol-12-myristate-13-acetate (PMA) increased the immunogenicity of chronic lymphocytic leukemia (CLL) cells taken from treated and untreated patients. In the same study, analogue 15 was tested and found to display the same effects on CLL cells as PMA, increasing the clinical potential of this finding. This could lead to immunotherapy-based treatments for CLL and other hematopoetic cancers based on bryostatin analogues (Hammond et al., 2005).

The phorbol esters have been widely used in biochemical studies to measure the consequences of activating C1 domain containing proteins. Due to their tumor-promoting effects, there is great reluctance to use them in therapeutic applications, although there have been some rare instances of testing the therapeutic effects of phorbol treatment in humans. With the exception of their effects on tumor growth, bryostatin and the phorbol esters share many biological effects (Sako et al., 1987). Bryostatin has been approved for therapeutic testing in humans, so in cases where the phorbol treatment induces a potentially beneficial effect in biochemical studies, it is natural to switch to bryostatin for in vivo testing and development. The limited availability and cost of bryostatin effectively prohibit its use in place of phorbol esters. Bryostatin analogues are therefore attractive as agents in these applications.

Figure 6.13 Translocation of PKCδ by bryostatin, analogue 1 and analogue 15.
6.7 CONCLUSION AND OUTLOOK

The established and growing links between PKC signaling and human diseases suggest that the need for and therapeutic potential of agents that act on PKC will grow. The most widely used agents, the phorbol esters, are tainted by their tumor-promoting effects. Bryostatin has shown anticancer activity, but the low isolation yield and structural complexity of the natural product effectively prohibit its widespread use. Therefore, functional analogues of bryostatin, which possess bryostatin-like activity but can be prepared in reliable quantities and tuned for function, are compounds with enormous potential for the treatment of human disease. More generally, these studies collectively illustrate the emerging role of rational drug design coupled to synthesis in the generation and advancement of new therapeutic leads. Whereas traditionally the focus of synthesis has been on making natural products themselves, the growing realization that few natural products are optimal for human use and the creative challenge of designing more effective but simpler leads collectively create new and highly significant opportunities for synthesis. While there is great satisfaction and value in achieving the synthesis of a known natural product, there is often even greater satisfaction and impact in the design and synthesis of previously unknown agents, especially as in the case of the bryologs, when these agents have activities and promise beyond natural products.

ACKNOWLEDGMENTS

Financial support for aspects of the work described herein was provided by grants from the National Institutes of Health (CA 31845) and GPC Biotech. This work is a product of many exceptionally talented, creative, and dedicated investigators whose names appear as coauthors and in the references provided.

REFERENCES


7.1 INTRODUCTION TO PARP AND PARG

Driven by DNA damage, the poly(ADP-ribose) polymerases (PARPs) utilize NAD\(^+\) to create poly(ADP-ribose) (PAR) on the glutamic acid residues of proteins, drastically altering the overall charge and size of the modified protein and ultimately initiating DNA repair (Bürkle, 2001). PAR biopolymers consist of up to 200 ADP-ribose units, and polymer levels may increase more than 100-fold in minutes (Meyer-Ficca et al., 2004). This poly(ADP-ribosylation) of proteins is transient, because the enzyme poly(ADP-ribose)glycohydrolase (PARG) operates in both an endo- and exoglycosidic fashion to break down PAR into ADP-ribose monomers (Davidovic et al., 2001). If more severe DNA damage occurs, PARP-1 becomes cleaved by caspases into 89- and

*Corresponding author: hergenro@uiuc.edu

Drug Discovery Research: New Frontiers in the Post-Genomic Era, Edited by Ziwei Huang
Copyright © 2007 John Wiley & Sons, Inc.
24-kDa subunits (Scovassi and Poirier, 1999; Lazenik et al., 1994), which separates the DNA binding domain from the automodification and catalytic domains of PARP, thus inactivating the enzyme. Most likely this PARP-1 cleavage plays a role in preserving cellular energy for apoptosis by preventing futile cycles of DNA repair (Putt et al., 2004). Finally, in cases of extreme DNA damage, PARP-1 is overactivated, which depletes the cell’s valuable energy resources, thus leading to death by necrosis (Pieper et al., 1999). The cellular activity and processing of PARP-1 in response to these three levels of DNA damage is depicted in Fig. 7.1. While PAR has been shown to play a role in DNA repair and replication (Dantzer et al., 2000), it has also been implicated in control of chromosome migration during cell division (Shall and de Murcia, 2000), activation of transcription after DNA damage (Hassa and Hottiger, 1999), activation of the proteasome (Ulrich et al., 1999), and regulation of telomere maintenance (Smith et al., 1998) and is required for spindle assembly and function (Chang et al., 2004). Failure to dynamically regulate PAR metabolism causes increased cytotoxic sensitivity and leads to early embryonic lethality (Koh et al., 2004).

### 7.1.1 Potential Applications of PARP and PARG Inhibitors

Due to these key roles of PAR in the cell, the PARP isozymes and PARG are potential drug targets. Overactivation of PARP caused by severe DNA damage quickly leads to the depletion of NAD$^+$, and this loss of energy inside the cell usually culminates in cell death through necrosis (Pieper et al., 1999). In cases of ischemic injury, when oxygen deprivation alone drastically decreases cellular energy output, overstimulation of PARP can also cause a drop of NAD$^+$ and ATP to less than 20% of the normal level (Berger, 1985). Inhibition of either PARP or PARG in such cases could potentially maintain cellular energy at acceptable levels, thus decreasing necrotic cell death. Conversely, when utilized in the context of tumor cells, PARP or PARG inhibitors might enhance the cytotoxicity of cancer therapies already in use (Li et al., 2001). It is presumed that the addition of such an inhibitor prevents the recovery of tumor cells from lethal DNA damage after radiation or other cancer therapies. The recently illuminated requirement for PAR during spindle assembly (Chang et al., 2004) also hints at the possible utility of spindle-associated PARP and PARG as cancer drug targets. Finally, both PARP and PARG are envisioned as potential targets for neuroprotection. Just as inhibition of these enzymes could alleviate damage in cases of ischemic injury, inhibition of PARP and PARG may potentially reduce neuronal death (Ying et al., 2001). Additionally, although it is unknown if abnormal accumulation of PAR is found in humans, a number of neurodegenerative disorders such as Alzheimer’s and Parkinson’s diseases are associated with the accrual of protein aggregates, and study of PAR accumulation in cases of PARG inhibition might be useful in understanding these conditions (Hanai et al., 2004). Clearly, PAR’s pivotal role in the cell indicates that both PARP and PARG inhibitors have considerable medicinal potential versus a variety of disease states (Fig. 7.1).

### 7.1.2 PARP Inhibitors and Their Drawbacks

Although they have limited cellular uptake and low potency and specificity, the most commonly used PARP inhibitors are nicotinamide, the endogenous inhibitor of
Figure 7.1  (a) In response to mild DNA damage, PARP-1 binds to DNA and catalyzes the formation of poly(ADP-ribose) polymers onto protein acceptors including itself. Because of electrostatic repulsion between the damaged DNA and the newly formed polymer, PARP-1 releases the DNA, which is then primed to recruit DNA repair enzymes such as XRCC1 and DNA ligase. PARP-1 remains in an inactive state until the enzyme PARG degrades the poly(ADP-ribose) polymer, which allows it to once again bind damaged DNA. (b) Upon more severe DNA damage, the cell initiates the apoptotic cascade. Release of cytochrome c ultimately leads to cleavage of PARP-1, thus preventing unnecessary cycles of DNA repair to save cellular energy for apoptosis. (c) Extreme DNA damage causes PARP-1 overactivation, which results in a loss of cellular energy. With NAD\(^+\) severely depleted, the cell cannot make ATP, and this leads to inflammation and death through necrosis. See color plates.
PARP, and 3-aminobenzamide (Virag and Szabo, 2002; Wilson et al., 2004; Contoni et al., 1987; Szabo et al., 1998). Many other PARP inhibitory compounds fall into two classes: monoaryl amides and bi-, tri-, or tetracyclic lactams, which often contain either an aromatic carboxamide or a carbamoyl group attached to a polyaromatic heterocyclic skeleton (Virag and Szabo, 2002). To date, the PARP family is thought to comprise at least 18 isozymes (Pion et al., 2003), and due to the highly conserved NAD$^+$ binding site of these PARPs, it is likely that isozyme-specific inhibitors will be difficult to develop (Smith, 2001). Although some preliminary studies have been completed that compare the inhibitory effects of phenanthridinones on PARP-1 versus PARP-2 (Perkins et al., 2001), the issue of isozyme specificity has not been fully examined. While it cannot be denied that PARP inhibitors have been shown to have a beneficial effect in many in vivo studies (Virag and Szabo, 2002; Tentori et al., 2002; Southan and Szabo, 2003), before they can be utilized in the clinic, some important safety aspects must first be considered. Although PARP-deficient mice do not show an increased level of spontaneous tumors, they do display an increased number of chemically induced tumors in comparison to wild-type ones (Tsutsumi et al., 2001), and even the most well-documented PARP inhibitors display undesired metabolic side effects by knocking out the normal physiologic functions of PAR (Lu et al., 2003; Milam and Cleaver, 1984). Long-term studies have not yet been completed, but due to PARP’s active roles in DNA repair and preservation of genomic integrity (Dantzer et al., 2000; Shall and de Murcia, 2000; Hassa and Hottiger, 1999; Smith et al., 1998; Chang et al., 2004), extended PARP inhibition could increase mutation rates and ultimately lead to cancer (Virag and Szabo, 2002).

### 7.1.3 PARG Inhibition: An Attractive Alternative to PARP Inhibitors

Obstruction of DNA repair mechanisms by PARP inhibition might be avoided by inhibiting PARG instead of PARP. It has been suggested that while the concentration of PARG in the cell is 13–50 times less than that of PARP (it has been reported that the number of PARG molecules present per cell may be as low as 2000), inhibitors of PARG should be just as useful as any previously demonstrated PARP inhibitors; in fact, the specific catalytic activity of PARG is 50- to 70-fold higher than that of PARP (Hatakeyama et al., 1986; Rapizzi et al., 2004). While PARP inhibition prevents the formation of PAR and consequently disrupts its normal metabolic functions, PARG inhibition indirectly slows PARP activity by locking PARP in the fully automodified (with PAR) and inactive state. PARG inhibition could possibly allow PAR to maintain its normal function within the cell (Lu et al., 2003; Milam and Cleaver, 1984). Finally, in stark contrast to the burgeoning PARP family of enzymes, there has only been one PARG identified to date (Pion et al., 2003), and thus its inhibition would not be complicated by the need for isozyme specificity. By searching for direct inhibitors of PARG enzymatic activity, it is possible that potent, specific, and pharmacologically acceptable drugs will be discovered.
7.2 PROPERTIES OF PARG

As mentioned earlier, there are at least 18 proposed members of the PARP family (Pion et al., 2003), but only one PARG has been identified to date, and its sequence is highly conserved among mammals (Davidovic et al., 2001). Three PARG isoforms are known: the 110-kDa nuclear form, and the more abundant 103- and 99-kDa cytoplasmic forms (Meyer-Ficca et al., 2004; Bonicalzi et al., 2003). Homology searches have indicated that the PARG sequence contains both a nuclear localization signal (NLS) (Lin et al., 1997) and a leucine-rich nuclear export signal (NES) (Shimokawa et al., 1999), suggesting that PARG shuttles between the cytoplasm and nucleus to maintain tight control of PAR catabolism (Fig. 7.2) (Bonicalzi et al., 2003). Interestingly, PARG is a substrate for caspase-3, but the functional implications of this cleavage are unclear to date (Affar et al., 2001). Although it is stable between pH 5 and 9, PARG’s optimum pH for enzymatic activity has been found to be 7.0 to 7.5. The enzyme exhibits substantial thermostability, and its activity increases up to temperatures of 50°C; thereafter, PARG activity decreases sharply (Hatakeyama et al., 1986).

7.2.1 Biochemistry of PARG

As highlighted earlier, although PARG has a low cellular concentration, it is constitutively active and has a high specific activity (5400 turnovers per minute)

Figure 7.2 Human PARG110 has a two-domain structure, with both a regulatory domain (1–417) and a catalytic domain (418–976). Along with a nuclear export signal from 126–134, a bipartite-type nuclear localization signal is found from 421–426, along with putative classical-type NLS signals from 10–16 (Meyer-Ficca et al., 2004) and 838–844 (Masutani et al., 2003). PARG 103 lacks exon 1, and PARG 99 lacks exons 1 and 2, which encode for the first NLS. See color plates.
(Hatakeyama et al., 1986); as such, the in vivo half-life of PAR in DNA-damaged cells is only ~1 minute (Alvarez-Gonzalez and Althaus, 1989). The nuclear PAR concentration is approximately one to two times the $K_M$ of PARG (0.3 μM) (Davidovic et al., 2001; Hatakeyama et al., 1986), and it has been found that the rates of PAR hydrolysis are a function of both concentration and structure (Brochu et al., 1994a). Long-chain polymers are catabolized more quickly, leaving behind 20- to 35mers,

![Figure 7.3](image.png)

**Figure 7.3** PARP catalyzes the formation of poly(ADP-ribose) polymers (from NAD$^+$) on glutamic acid residues of protein substrates. This process can occur from either the 2’ position (elongation) or 2” position (branching). PARG catabolizes the ADP-ribose polymers in both an exo- and endo-glycosidic fashion. ADP-ribose monomers are directly produced from exoglycosidic activity, whereas the endoglycosidic mode of action gives smaller ADP-ribose polymers that can be converted into ADP-ribose through further processing.
which are poorer substrates for PARG (Lin et al., 1997; Burzio et al., 1976). In early stages of PAR catabolism, the enzyme degrades larger polymers in both an endo- and exoglycosidic fashion (Fig. 7.3), leaving behind both ADP-ribose monomers and smaller polymers, usually less than half the size of the original polymer. Conversely, smaller PAR are degraded at a rate that is 20-fold slower (for larger PAR the $K_M$ is approximately 100-fold lower than that of small polymers), suggesting that PARG differentially controls the levels of large and small PAR (Hatakeyama et al., 1986). Interestingly, it is known that binding of photolabeled adenosine diphosphate (hydroxymethyl)pyrrolidinediol (ADP-HPD), an ADP-ribose analogue inhibitor of PARG, competes with the binding of short PAR (5- to 15mer) but not with the longer, highly branched polymer (Ramsinghani et al., 1998). For PARG, substrate identification and binding appears to be a complex event, with possible multiple attachment points or subsites for polymer recognition. It also must be noted that these striking differences in substrate processing may be attributed to higher-order structures of large PAR; those polymers with a chain length of more than nine ADP-ribose units exhibit a secondary structure, while smaller polymers do not (Minaga and Kun, 1983).

7.2.2 Structure of PARG

No crystal structure for PARG has yet been reported. However, from studies of the catalytic fragment of recombinant bovine PARG, it is believed that the active site consists of at least some residues in the Leu-771 to Arg-801 region. Photodervitization studies with a known PARG inhibitor indicate Tyr-796 (Y796), which is a highly conserved residue in PARGs across a wide range of organisms, as a possible amino acid important for initial binding of PAR. Site-directed mutagenesis to replace Y796 with an alanine residue (Y796A) leads to an eight-fold decrease in catalytic efficiency ($K_{cat}/K_M$), a difference attributed to the ring-stacking ability of the Y796 residue with the adenine of the PAR substrate (Koh et al., 2003a,b). From homology studies it is known that Gly-262 and Gly-263 in Arabidopsis PARG are found in all mammalian PARGs, and a third Gly-264 is highly conserved (Panda et al., 2002). When Gly-262 is replaced with a glutamic acid residue to create the G262E mutant PARG enzyme, Arabidopsis seedlings accumulate much higher levels of PAR as compared to wild-type. These data, together with the high degree of amino acid conservation in this region of the PARG sequence, suggest that this GGG motif may play an important role in the regulation of PARG activity (Panda et al., 2002).

7.2.3 Genetic Studies with PARG

Mutation in the PARG gene, tej, in plants has been demonstrated to influence the transcription of genes involved in circadian rhythm regulation (Panda et al., 2002). Using a cycling bioluminescence reporter phenotype, tej mutants with abnormal circadian oscillations were discovered in Arabidopsis thaliana. While the effects of this recessive tej mutation are independent of light input, tej plants exhibit a period length 2 hours longer than the wild-type and always flower earlier under...
both long- and short-day conditions. Interestingly, PARP inhibitors have been shown to rescue the mutant tej phenotype. By slowing or stopping the rate of PAR synthesis, the known PARP inhibitor 3-aminobenzamide causes tej mutants to exhibit a wild-type circadian rhythm, thus indicating that poly(ADP-ribosyl)ation establishes a period length of the Arabidopsis circadian oscillator (Panda et al., 2002).

To date few studies have been completed with PARG knockout organisms. Loss of the catalytic domain of PARG causes lethality under normal conditions for development (25°C) in the larval stages of Drosophila melanogaster, but if the incubation temperature is increased to 29°C, approximately 25% of the PARG knockouts survive to the adult stage (Hanai et al., 2004). These organisms lacking PARG are associated with excessive accumulation of PAR in the central nervous system (CNS), progressive neurodegeneration, reduced locomotor activity, and shortened lifespan. Accumulation of PAR in the CNS could be explained by any of the following: PAR cannot be diluted in the CNS because neuronal cells in the CNS do not divide; PARP is possibly more active in the brain, and therefore polymer builds up more quickly there (Menegazzi et al., 1991); or the other enzymes responsible for degrading PAR are less active in the brain than in other organs of D. melanogaster. Whatever the reason for the excessive amounts of PAR found in the CNS of these PARG knockouts, these studies clearly suggest that PAR metabolism is a necessary part of normal neuronal function.

Through gene targeting in embryonic stem cells and in mice, deletion of the 110-kDa PARG protein, which is normally found in the nucleus (Meyer-Ficca et al., 2004), has been thoroughly studied. While no full-length PARG mRNA is detected in the PARG knockout embryonic fibroblasts, a 60-kDa truncated version of PARG remains present (Cortes et al., 2004). Mice deficient in the full-length PARG do not show any obvious phenotype, are fertile, and do not exhibit excessive buildup of PAR in tissue as might first be expected. This is attributed to the fact that the 60-kDa PARG remains unaffected in such cases and can still be observed by Western blot analysis. On the other hand, specific PARG activity is significantly lower in cells obtained from 110-kDa PARG-deficient animals, and they are hypersensitive to DNA alkylating agents, radiation, lipopolysaccharide-induced septic shock, and streptozotocin-induced diabetes (Cortes et al., 2004).

Most recently, in vivo studies have indicated that catabolism of PAR by PARG is necessary for genomic stability in higher-order eukaryotes (Koh et al., 2004). In contrast to the above study in which PARG activity remained in the 110-kDa PARG knockout mice due to the presence of alternatively spliced isoforms, complete PARG knockouts are lethal. While PARG +/− mice are viable and fertile, homozygous null mice (PARG −/−) arrest before gastrulation at embryonic day 3.5. Under conditions of PARP inhibition (treatment with the PARP inhibitor benzamide), it is possible to culture PARG −/− embryonic trophoblast stem (TS) cells, although removal of the PARP inhibitor causes abnormal morphology, and after 3 days approximately 60% of the PARG −/− TS cells undergo apoptosis (Koh et al., 2004). Cells remaining after removal of the benzamide PARP inhibitor also exhibit slow growth rates, PAR accumulation, and increased sensitivity to the cytotoxic agent MNNG. Due to the lethality of the knockouts, it has been difficult to study PARG via traditional genetic approaches.
7.3 PARG PURIFICATION AND ASSAYS

Although provocative data indicate PARG inhibitors could be useful for a range of maladies, the lack of detailed structural knowledge has hampered their development. As a result, substrate analogues and compounds identified through random screening are the only known PARG inhibitors. However, recent advances in both PARG production and high-throughput assays should enable the discovery of a next generation of PARG inhibitors. Described in this section are the methods used to obtain the PARG enzyme, assay protocols, and a comprehensive look at known PARG inhibitors.

7.3.1 PARG Purification and Recombinant Protein Expression

Due to the very low content of PARG in various tissues and its instability during purification, difficulties have been encountered in obtaining pure quantities of this enzyme. Since its initial discovery in 1971 (Miwa and Sugimura, 1971), PARG has been purified and characterized from a wide range of cell types, which has allowed for full enzymatic characterization. Initial attempts at obtaining PARG focused on its extensive purification from calf thymus and other tissues, a process that can involve up to six purification steps and produces low yields of purified enzyme (Hatakeyama et al., 1986; Minaga and Kun, 1983; Brochu et al., 1994b; Thomassin et al., 1990; Shah et al., 1995). Additionally, although PARG is known to have a molecular weight of about 110 kDa (Lin et al., 1997; Winstall et al., 1999), during purification it is easily proteolyzed to 59- and 74-kDa fragments (Affar et al., 2001), which originally caused uncertainty as to whether different types of PARG protein exist in the cell (Maruta et al., 1991). However, full-length PARG was purified in 1997 (Lin et al., 1997) utilizing recombinant expression from the 4.1-kb PARG cDNA in Escherichia coli to produce the 110-kDa protein, thus eliminating the controversy surrounding PARG’s size.

7.3.2 PARG Enzymatic Activity Assays

Many assays for measuring PARG enzymatic activity are limited by the fact that they are neither convenient nor high-throughput. The most common technique used to evaluate PARG activity first involves the preparation of $^{32}$P-labeled PAR using PARP and $^{32}$P-labeled NAD$^+$. Then, following purification and reaction of the radiolabeled PAR with PARG, either a thin-layer chromatography (TLC) technique or high-resolution polyacrylamide gel electrophoresis is utilized to separate the catabolized ADP-ribose from PAR, and radioactivity is measured (Ménard and Poirier, 1987; Pacheco-Rodriguez and Alvarez-Gonzalez, 1999). Although both these radiometric assays have been helpful in determination of IC$_{50}$ values for PARG inhibitors (Koh et al., 2003a; Masutani et al., 2002; Slama et al., 1995) and in studies of PARG kinetics (Koh et al., 2003a; Maruta et al., 1997), neither is particularly useful in screening large libraries of compounds in a high-throughput manner. Additionally, PARG works in an endo- and exoglycosidic fashion, and these TLC-based methods are only able to quantitate monomeric ADP-ribose; the direct products of endoglycosidic cleavage cannot be detected (Putt and Hergenrother, 2004).
More recently, a nonradiometric assay involving the conversion of the ADP-ribose product of PARG activity into a fluorescent molecule has been developed. Taking advantage of the fact that reducing sugars are known to react with aromatic amidines to give fluorescent products (Kai et al., 1985; Coquet et al., 1991; Kakita et al., 2002), this assay involves the reaction of benzamidine with ADP-ribose (a reducing sugar) to create a fluorophore (Fig. 7.4) (Putt and Hergenrother, 2004). A simple fluorescence reading can be used to measure the amount of free ADP-ribose hemiacetal, thus quantifying PARG activity. One limitation of this assay is that it cannot be used in cases in which potential PARG inhibitors contain aldehydes or reducing sugars. Unlike the aforementioned PARG assays, this fluorescence-based method is capable of detecting both exo- and endoglycosidic cleavage of PAR, and this method is highly sensitive. Additionally, development of an assay based on fluorescence rather than radioactivity greatly reduces both cost and time expenditures and allows compounds to be screened in a high-throughput manner.

7.4 PARG INHIBITORS

From a drug-development perspective, the perfect PARG inhibitors would be highly potent and specific compounds of modest size (molecular weight <500). Although the current classes of inhibitors do not necessarily fit into all the categories of an ideal pharmaceutical, they have certainly offered insights into the structure and function of PARG. These compounds can be broadly grouped into the three categories of DNA intercalators, tannins, and substrate analogues. Below, all classes of PARG inhibitors along with their in vivo effects will be discussed.

7.4.1 DNA Intercalators

Of the few classes of PARG inhibitors known, the ones that potentially offer the best cell permeability are polyaromatic compounds typically thought of as DNA intercalators, depicted in Fig. 7.5. These compounds were initially evaluated as a consequence of their known DNA binding properties and the known role of PARG in the DNA damage and repair process.

Using the $^{32}$P-PAR assay, the $K_i$ for ethacridine was estimated at $<7$ µM, the $K_i$ for Tilorone was $\sim 7$ µM, and the $K_i$ for profavine was 36 µM, while another well-known DNA intercalator, ethidium bromide, is not inhibitory (Tavassoli et al., 1985).
These were all relative to a $K_M$ for the PAR substrate that was calculated as 1.7 μM in this particular assay system. In general, these inhibitors appear to give competitive inhibition, and there is some evidence that this inhibition may be due to the compounds forming a complex with the PAR polymer, thus blocking substrate binding to the enzyme; this seems to be particularly true for ethacridine (Tavassoli et al., 1985; Bernardi et al., 1997). Further support of this hypothesis is found in the fact that ethacridine and other intercalators not only inhibit the degradation of PAR by PARG, but also prevent degradation of the polymer by snake venom phosphodiesterase and hinder ethanol/acetate precipitation of PAR (Tavassoli et al., 1985). Recent work has focused on the design and synthesis of symmetrically disubstituted aromatic compounds that exhibit PARG-inhibitory properties (Li et al., 2001). Representative members of two of the most potent classes, 2,7-substituted fluorenes and xanthen-9-ones, which are structurally very similar to the DNA intercalator Tilorone, are also depicted in Fig. 7.5. In conclusion, while these compounds are modestly potent inhibitors of PARG and have been used in several in vivo studies (see below), questions about their selectivity for PARG inhibition versus DNA

**Figure 7.5** Intercalator and intercalator-like inhibitors of PARG and their IC$_{50}$ values.

These were all relative to a $K_M$ for the PAR substrate that was calculated as 1.7 μM in this particular assay system. In general, these inhibitors appear to give competitive inhibition, and there is some evidence that this inhibition may be due to the compounds forming a complex with the PAR polymer, thus blocking substrate binding to the enzyme; this seems to be particularly true for ethacridine (Tavassoli et al., 1985; Bernardi et al., 1997). Further support of this hypothesis is found in the fact that ethacridine and other intercalators not only inhibit the degradation of PAR by PARG, but also prevent degradation of the polymer by snake venom phosphodiesterase and hinder ethanol/acetate precipitation of PAR (Tavassoli et al., 1985). Recent work has focused on the design and synthesis of symmetrically disubstituted aromatic compounds that exhibit PARG-inhibitory properties (Li et al., 2001). Representative members of two of the most potent classes, 2,7-substituted fluorenes and xanthen-9-ones, which are structurally very similar to the DNA intercalator Tilorone, are also depicted in Fig. 7.5. In conclusion, while these compounds are modestly potent inhibitors of PARG and have been used in several *in vivo* studies (see below), questions about their selectivity for PARG inhibition versus DNA
binding are likely to doom their use as therapeutic agents or chemical tools to extensively investigate PARG biology.

### 7.4.2 Tannins

A survey of the literature shows that the naturally occurring polymethoxy-phenolic compounds called tannins are by far the most well-studied PARG inhibitors (Fig. 7.6). Initially extracted from green tea leaves and pinecone fractions, it has been found that the oligomeric ellagitannins such as nobotanin B, E, and K are the most potent tannin inhibitors of PARG, with *in vitro* IC$_{50}$ values ranging from 0.44 μM for nobotanin K to

![Nobotanin B](image1.png)

Nobotanin B

IC$_{50}$ 4.8 μM

![Nobotanin E](image2.png)

Nobotanin E

IC$_{50}$ 1.4 μM

**Figure 7.6** Tannins and their representative IC$_{50}$ values.
4.8 μM for nobotanin B (Tsai et al., 1992). In comparison, the circular dimeric ellagitannin Oenothein B (Oen B) inhibits with an IC₅₀ of 3.8 μM (Maruta et al., 1997). Inhibitory activity of oligomeric ellagitannins increases with the number of monomeric residues (dimeric < trimeric < tetrameric), with condensed tannins having little effect on PARG activity even at very high concentrations, possibly indicating that differences in inhibitory effects are related to variations in secondary
structures of the tannins (Aoki et al., 1993). Only hydrolyzable tannins (gallotannins and ellagitannins) were able to inhibit PARG. While tannins of the trimeric and tetrameric variety exhibit mixed-type inhibition, nobotanin B, (a dimer) has been found to competitively inhibit PARG with respect to PAR (Tsai et al., 1992).

Several studies with cell extracts have confirmed the inhibitory properties of these phenolic phytochemicals against PARG. When 200 μM tannin was added to HeLa nuclear extracts, a significant 40-fold increase of the PAR level was observed, and it was confirmed that this increase was indeed due to inhibition of the catalytic activity of PARG (Keil et al., 2004).

### 7.4.3 Substrate Analogues

As expected, substrate analogues of PAR have been synthesized and are shown to be both potent and specific inhibitors of PARG (Fig. 7.7). Though ADP-ribose monomer

<table>
<thead>
<tr>
<th>Entry</th>
<th>Name</th>
<th>IC&lt;sub&gt;50&lt;/sub&gt; (μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ADP-ribose</td>
<td>120</td>
</tr>
<tr>
<td>2</td>
<td>APD-HPD</td>
<td>0.12</td>
</tr>
<tr>
<td>3</td>
<td>APD-HPM</td>
<td>3.07</td>
</tr>
<tr>
<td>4</td>
<td>APD-HP</td>
<td>19.2</td>
</tr>
<tr>
<td>5</td>
<td>8-N&lt;sub&gt;3&lt;/sub&gt;-APD-HPD</td>
<td>0.39</td>
</tr>
<tr>
<td>6</td>
<td>2-N&lt;sub&gt;3&lt;/sub&gt;-APD-HPD</td>
<td>290</td>
</tr>
<tr>
<td>7</td>
<td>APD-GPD</td>
<td>&gt;1000</td>
</tr>
</tbody>
</table>

*Figure 7.7* Substrate analogue PARG inhibitors and their comparative IC<sub>50</sub> values.
units inhibit PARG with an IC$_{50}$ of 120 μM, ADP-HPD, an NH-analogue of ADP-ribose, has been shown to inhibit the action of PARG with a 1000-fold lower IC$_{50}$ value (entry 2) (Slama et al., 1995). Structure–activity analysis utilizing ADP-HPD analogues containing various alterations to the purine base has revealed the importance of the adenine moiety for optimal inhibition. It was discovered that substitution at the 2-position of adenine negatively affects PARG inhibition (entry 7), while substitution at the 8-position has no adverse effects (entry 5). Both the pyrroline cis-hydroxyls of ADP-HPD are necessary for inhibitor binding; removal of one increases the IC$_{50}$ value to 3.07 μM (entry 3), and complete elimination of these hydroxyls results in a 160-fold increase in IC$_{50}$ value as compared to ADP-HPD (entry 4) (Koh et al., 2003a).

**7.4.4 Other Inhibitors**

Additionally, through random screening it has been found that some cyclic peptides are capable of inhibiting PARG activity. One cyclic hexadepsipeptide (known as PD 124,966) isolated from the fermentation products of actinomycetes was found to have an IC$_{50}$ value of 40 μg/ml, while pargamicin, which is purified from the fermentation broth of *Amicolatopsis* exhibits an IC$_{50}$ value of 28 μg/ml (Masutani et al., 2002). Interestingly, linear peptides containing the same piperazine residues as PD 124,966 and pargamicin appear to have no effect on PARG activity.

**7.5 IN VIVO STUDIES**

**7.5.1 Studies with DNA Intercalators**

The second-generation intercalator-like PARG inhibitors are ideal for *in vivo* studies due to their low molecular weights and good cell permeability properties. One member of the Tilorone family of PARG inhibitors, GPI 16552 or 2,7-bis [N-(3-phenyl-propyl)carbamoyl]-9-oxo-9H-fluorene, has exhibited neuroprotective properties in a rat model of focal cerebral ischemia. Both a 30-minute pre-ischemia treatment or a 1- or 2-hour post-ischemia treatment with GPI 16552 reduces the total infarct volume by 47–53%, with greater protection observed in the cortical areas (43–59%) than in the subcortical areas (28–40%) of the brain (Lu et al., 2003). Furthermore, application of the novel PARG inhibitor GPI 18214 substantially reduces the septic-shock-like syndrome caused by zymosan in mice (Genovese et al., 2004). Zymosan is a nonbacterial, nondotoxie agent that produces liver, intestine, lung, and kidney failure in test animals. In mice treated with 40 mg/kg of GPI 18214, which has an IC$_{50}$ value of 3.0 μM, all signs of pancreatic, renal, and liver damage were abolished, and lung injury was significantly reduced after induction of zymosan (Genovese et al., 2004). The protective effects of GPI 18214 correlate with reduced neutrophil and PMN lysosome infiltration into the intestine and lung, which have been known to augment tissue damage through inflammatory effects (Hassa and Hottiger, 2002; Lefer and Lefer,
1993). It is hypothesized that these PARG inhibitors might reduce infarct volumes not only by stopping the vicious cycle of NAD\(^+\) depletion, but also by their anti-inflammatory effects, as described below (Lu et al., 2003; Rapizzi et al., 2004; Genovese et al., 2004).

### 7.5.2 Studies with Tannins

The *in vivo* effects of tannins have been thoroughly studied. While the more potent PARG inhibitors nobotanin E (trimer) and K (tetramer) have little inhibitory effect on the catabolism of PAR in intact 34I cells, it has been found that nobotanin B causes concentration-dependent inhibition of PAR catabolism (Tsai et al., 1992). The inability of nobotanin E and K to inhibit PARG is due to the poor penetration of these compounds into plasma membranes. When cell extracts incubated with various tannins were analyzed by HPLC, it was found that only nobotanin B could be taken up by the cells in a significant level (Tsai et al., 1992), thus hinting at the major downfall of this high-molecular-weight class of compounds.

PARG enzymatic activity also appears to directly aid the process of DNA damage repair (Oei and Ziegler, 2000). Using the specific competitive inhibitor Oen B, the role of PAR catabolism by PARG in synchronized HeLa S3 cells at G1 phase has been studied. Because the process of excision repair synthesis of DNA requires large amounts of energy, it has been hypothesized that the energy produced during PAR catabolism by PARG is directly involved in this DNA repair. In fact, when \(^3\)H-labeled thymidine 5'-triphosphate (\(^3\)H-dTTP) was added to poly(ADP-ribosyl)ated nuclei from G1-synchronized HeLa cells that were subsequently treated with the DNA-damaging agent MNNG, \(^3\)H-dTMP incorporation into the DNA began after a short 3-minute lag (Maruta et al., 1997). On the other hand, in the presence of Oen B very little incorporation of \(^3\)H-dTMP was observed, suggesting that energy in the form of ATP produced by PAR catabolism serves as a direct energy source for repair synthesis. Furthermore, in cases of moderate DNA damage, such as that caused by NO donors such as spermine nonoate (SNO), addition of the PARG inhibitor gallotannin greatly decreases the ability of rat germinal cells to recover (Di Meglio et al., 2004). In this case, the important role of PARG in protection from DNA damage might be twofold. First, PARP-1 seeks out and repairs DNA strand breaks while suppressing transcription to prevent the expression of damaged genes (Masson et al., 1998; Ziegler and Oei, 1999). If PARP-1 is inhibited by its own auto-poly(ADP-ribosylation (Lautier et al., 1993; D’Amours et al., 1999; Zahradka and Ebisuzaki, 1982), PARG works to restore PARP to its original active status; in short, PARG inhibition ultimately prevents PARP’s DNA repair mechanism from functioning properly. Additionally, as explained above, by metabolizing PAR into ADP-ribose monomers and thus adding to the NAD\(^+\) energy stores, PARG possibly supplies energy needed for the DNA ligation step in base excision repair (Oei and Ziegler, 2000).

On the other hand, in cases of severe DNA damage, PARG inhibition can reduce the amount of cell death observed. A significant decrease in H\(_2\)O\(_2\)-induced cell death is observed when cultured murine astrocytes are preincubated with as little as 100 nM gallotannin (Ying and Swanson, 2000). Amazingly, gallotannin is 10-fold more...
potent than the PARP inhibitor benzamide in preventing oxidative cell death. Later studies have shown that nobotanin B, a PARG inhibitor isolated from the plant *Tibouchina semidecandra*, also offers neuroprotection against both oxidative (induced by both H\textsubscript{2}O\textsubscript{2} and N-methyl-D-aspartate) and excitotoxic (induced by MNNG) cell death (Ying et al., 2001). Attempts to discover the mechanism of tannin cytoprotection led to the discovery that while gallotannin increases the amount of PAR accumulation in astrocyte cultures, the known PARP inhibitor benzamide actually decreases protein poly(ADP-ribosyl)ation (Ying et al., 2001). This indicates that PARG inhibitors do not necessarily inhibit PARP-1 directly but instead slow the turnover of PAR and therefore stop the vicious cycle of NAD\textsuperscript{+} consumption that ultimately leads to cell death through necrosis. While the above explanation of PARG inhibitors involving prevention of an energy crisis in the cell seems to be the most accepted, other mechanisms could possibly contribute to the known protective effects of PARG inhibition in cases of severe DNA damage. Since PARP-1 is inhibited by its own automodification by PAR (D’Amours et al., 1999; Zahradka and Ebisuzaki, 1982), PARG inhibitors could work to indirectly knock out PARP by preventing catabolism of the inhibitory polymers. Another enzyme, Ca\textsuperscript{2+}/Mg\textsuperscript{2+}-dependent endonuclease, which is responsible for DNA fragmentation, is also inhibited by poly(ADP-ribosyl)ation and could be indirectly inhibited by PARG in a similar fashion (Tanaka et al., 1984; Yakovlev et al., 2000). Although the exact mechanism of action is not totally clear, PARG inhibition by tannin-related drugs has potential as a neuroprotective tool.

*In vivo* it has been found that PAR metabolism, which is controlled by the concerted action of PARP and PARG, plays a role in the transfer of metabolites, proteins, and drugs across cellular membranes through ATP-binding cassette (ABC) transporters (Dumitriu et al., 2004). UVB irradiation causes oxidative stress in cells due to the generation of peroxynitrite (PN) and inhibits the activity of ABC transporters (Deliconstantinos et al., 1996). Application of either 3,4-dihydro-5(4-(1-piperindinyl)butoxy)-1(2H)-isoquinoline (DPQ), a PARP inhibitor, or the known PARG inhibitor gallotannin restores function of ABC transporters following irradiation, most likely because ADP-ribose itself actually serves as an ABC transport inhibitor (Dumitriu et al., 2004). Multiple-drug resistance of tumor cells has often been linked to activity of ABC transporters (Cole et al., 1992), and this discovery highlights yet another potential application of PARG to cancer therapy.

Remarkably, PARG has also been linked to the regulation of transcription and consequently the expression of proinflammatory genes. When a PARG inhibitor such as gallotannin is incubated with macrophages, PAR accumulates, which in turn causes expression of inducible nitric-oxide synthase (iNOS) and cyclooxygenase-2 (COX-2), but not interleukin-1\textbeta or tumor necrosis factor-\textalpha (Rapizzi et al., 2004). This discovery is further supported by the fact that silencing of PARG with small interfering RNA (siRNA) prevents gallotannin-mediated expression of iNOS and COX-2 (Rapizzi et al., 2004). At this point, it is necessary to mention that chromatin condensation, which is one of the markers of gene down-regulation, has been shown to be tied to PARG and its associated PAR catabolism (de Murica et al 1986). Conversely, inhibition of PARG and the consequent buildup of
PAR is associated with chromatin loosening (de Murcia et al., 1988), which allows for gene transcription. Taken together, these data indicate that PARG inhibition might increase immune cell activation and therefore have major impacts on inflammation.

Through the use of the tannins as putatively selective inhibitors of PARG, the neuroprotective effects of PARG inhibition have been demonstrated. However, recently it has been suggested that while gallotannin may act as a PARG inhibitor, this compound is promiscuous and ineffective in vivo (Falsig et al., 2004). Initial questions concerning the specificity of gallotannin for PARG first arose upon close examination of its properties as provided by its manufacturer, namely that gallotannin is known to form insoluble complexes with proteins and acts as a fixative/stain. In models of H₂O₂- and MNNG-mediated cytotoxicity, it was found that while gallotannin prevents cell death in oxidative models as expected, in the studies involving nonoxidative DNA damage this polyphenol has no effect. Additionally, two other polyphenols, quercetin and catechin, were shown to alleviate cell death in the same model at concentrations similar to gallotannin, although they have never been reported as PARG inhibitors. One would also expect that PARG inhibition by gallotannin in living cells would enhance PAR production, but it has actually been found to reduce PAR formation, an effect attributed to its reactive oxygen species scavenging ability (Falsig et al., 2004). The strong H₂O₂-scavenging properties of tannins at low concentrations have been well-documented (Yokozawa et al., 1998), and the major component of gallotannin, tetra-galloyl-glucose, contains 13 phenolic hydroxyl groups, and 12 are in ortho positions and primed to react with hydroxyl radicals. While gallotannin does act as a PARG inhibitor in cell-free assays, it also has antioxidant properties and can protect cells from oxidative stress without PARG inhibition. Therefore, because gallotannin is not ideal for the evaluation of PARG in cellular death models, all results obtained previously using this compound must be carefully reviewed (Falsig et al., 2004).

### 7.5.3 Studies with Substrate Analogues

Along with an array of microtubules and their associated proteins, which act to align and separate chromosomes, it has been found that PAR is a nonproteinaceous, nonchromosomal component of the spindle necessary for bipolar spindle assembly and function (Chang et al., 2004). Enriched at both the spindle poles and kinetochores, cellular PAR concentration markedly increases during metaphase and anaphase–telophase due to heightened PARP activity (Bakondi et al., 2002), which suggests that PARG itself might be important in regulating spindle disassembly/assembly as well. After addition of 100 μg/ml of PARG to Xenopus egg extract spindles, a rapid breakdown of spindle structure is observed. Because of this hydrolytic activity of PARG, which can indeed be inhibited by the addition of ADP–HPD, microtubules were observed to turn outward from the spindle center, causing disconnection of the two half-spindles. Along with PARP and PAR, PARG is found to colocalize with spindle microtubules in BSC1 and HeLa cells, once again suggesting that it actively works to regulate PAR levels in the spindle (Chang et al., 2004).
As highlighted above, three major classes of PARG inhibitors exist to date: ADP-ribose analogues, DNA intercalators, and tannins. While the last two classes of compounds show considerable potency, neither one is ideal because each class has high toxicity, high molecular weight, and an inability to act as specific inhibitors. Because of their highly charged nature, ADP-ribose analogues are not cell-permeable and are therefore unsuitable for use as inhibitors in a biological context. Additionally, such substrate analogues closely resemble biologically relevant molecules, which may ultimately be a source of unwanted side effects. Despite the lack of optimal inhibitors, it is clear that PARG is an attractive target for drug discovery and development, and it seems as if the scope of potential inhibitors is only beginning to be understood. In consideration of the aforementioned problems, other classes of novel and potent PARG inhibitors are currently being sought. Armed with the new methods to rapidly screen for PARG inhibition, it is possible that the next potent inhibitor might be a member of a combinatorial library or small molecule collection. As new and selective inhibitors of PARG are developed, it is certain that these tools will be used to elucidate and define new and exciting roles for cellular poly(ADP-ribosyl)ation.

REFERENCES


8.1 INTRODUCTION

Ca$^{2+}$ signaling is one of the most important intracellular signal transduction pathways in diverse cells. The Ca$^{2+}$ mobilization is crucial for a wide variety of cell functions—for example, muscle contraction; secretion of neurotransmitters, hormones, and enzymes; fertilization of oocytes; and lymphocyte activation and proliferation. Local Ca$^{2+}$ signals are mainly due to release from internal stores, and these calcium-release events are controlled by various second messengers and regulated by Ca$^{2+}$ channel-associated proteins. Some small molecules—d-myo-inositol 1,4,5-triphosphate (IP$_3$) (Streb et al., 1983), cyclic adenosine diphosphate ribose (cADPR) (Lee et al., 1989), and nicotinic acid adenine dinucleotide phosphate (NAADP) (Lee and Aarhus, 1995)—are Ca$^{2+}$ mobilizing messengers that act on specific intracellular receptors.

*Corresponding author: zdszlh@mail.bjmu.edu.cn

Drug Discovery Research: New Frontiers in the Post-Genomic Era, Edited by Ziwei Huang
Copyright © 2007 John Wiley & Sons, Inc.
The different isoforms of ryanodine receptors (RyRs) (Galeone et al., 1991) and inositol triphosphate receptors (IP₃R) (Joseph et al., 1995) are involved in both local and global Ca²⁺ responses.

cADPR (Fig. 8.1), a metabolite of NAD⁺, was discovered in 1987 and is a signaling molecule to regulate calcium mobilization via RyR from intracellular stores in a wide variety of biological systems such as sea urchin eggs, pancreatic beta cells, smooth and cardiac muscles, T-lymphocytes, and cerebellum neurons (Lee, 2002). The structure of cADPR was determined by X-ray crystallographic analysis in 1994 (Lee et al., 1994). Three genes encoding mammalian RyR1, RyR2, and RyR3 proteins show a high degree of homology except in three divergent regions (Rossi and Sorrentino, 2000; Conklin et al., 1999; Takeshima et al., 1998; Bhat and Ma, 1999). Each protein consists of about 5000 amino acids. A large number of key proteins are involved in specifically shaping an RyR-dependent Ca²⁺ signal. FK 506-binding protein (FKBP), protein kinase A, calmodulin-dependent protein kinase, protein phosphatases 1 and 2A, calsequestrin, sorcin, triadin, and junctin are RyR-associated proteins that contribute to the endoplasmic/sarcoplasmic reticulum Ca²⁺ signaling system (Bers, 2004). Calmodulin also modulates RyR channels from both skeletal and cardiac muscles (Mackrill, 1999). For example, it was found that the Ca²⁺ release channels in skeletal muscle comprise four 565-kD type 1 RyR subunits and four molecules of the 12-kD protein FKBP12. FKBP12 stabilizes the RyR1 complex and enables the four subunits to open and close coordinately (Marx et al., 1998). In pancreatic β cells and in smooth muscle binding of ligands to FKBP, either the naturally occurring ligand cADPR or the drug FK 506 induces a conformational

![Figure 8.1](image-url)  
**Figure 8.1** The structures of cADPR, cIDPR, 3-deaza-cADPR, 7-deaza-cADPR, N¹-pentyl-cIDPcR, and cADPR triphosphate mimic.
change of FKBP resulting in its release from RyR, which in turn resulted in the opening of the RyR Ca\(^{2+}\) channel (Noguchi et al., 1997). However, it is unclear whether this is the general mechanism of action of cADPR or whether cADPR elicits calcium release by direct binding to RyR. A significant approach to exploring the molecular mechanism of calcium release is to investigate the structure–activity relationship of cADPR and cADPR analogues with their calcium-mobilizing activities. Agonists or antagonists of cADPR that resulted from such studies are valuable tools to study cellular signal transduction and are also useful as lead structures for drug discovery.

Since the discovery of cyclic ADP-ribose as a messenger for agonist-induced Ca\(^{2+}\) release, more than 100 papers have been published describing a role for this novel metabolite in regulating calcium mobilization in different biological systems. Some reviews have described the syntheses and properties of cADPR analogues, including the modifications on the adenine moiety, the southern and northern ribose, and the pyrophosphate groups (Guse, 2004; Potter and Walseth, 2004; Shuto and Matsuda, 2004). However, due to the difference in pharmacological properties between the various modified cADPR analogues and the different cell models employed, our view of the structure-activity relationship is still ambiguous. According to the biosynthesis of cADPR from \(\beta\)-NAD\(^+\) (Scheme 8.1), a variety of cADPR analogues were obtained by the enzymatic and chemo-enzymatic approach using nicotinamide dinucleotide-type precursors and ADP-ribose cyclase for the synthesis of the cyclic pyrophosphate. cADPR analogues with a substitution at the 8-position of the adenine ring (including 8-Br, NH\(_2\), CH\(_3\) and MeO) (Walseth and Lee, 1993; Reyes et al., 1999; Guse et al., 1997) were found to be antagonists of cADPR (Fig. 8.1). 8-Br-cADPR, although not as potent as 8-NH\(_2\)-cADPR, has been shown to be cell-permeant and can be used as a valuable tool. 7-Deaza-8-Br-cADPR is a cell-permeant antagonist and is about twice as effective as 8-Br-cADPR (Guse et al., 1999). 7-Deaza-cADPR was a partial agonist in sea urchin egg homogenates and

---

**Scheme 8.1** Biosynthesis of cADPR.
3-deaza-cADPR was about 70-fold more potent in sea urchin egg homogenates, as compared to cADPR (Bailey et al., 1997; Sethi et al., 1997; Wong et al., 1999) (Fig. 8.1). It seems that either the 3-deaza- or the 7-deaza-purine ring in cADPR could retain the conjugated plane shape and be recognized by the receptor for cADPR. In the case of the \( N^1 \)-coupled cyclic inosine diphosphate ribose mimics (cIDPR 2) (Fig. 8.1), different functional groups at the 8-position resulted in completely different biological activities; for example, 8-Br-cIDPR acts as a membrane-permeant agonist of \( \text{Ca}^{2+} \) signaling (Guse et al., 1999). However, due to the limitation of substrate-specificity of the ADP-ribosyl cyclase, the enzymatic and chemo-enzymatic method cannot be used for the construction of a larger variety of mimics modified in the northern or southern ribose moiety of cADPR. To build up a more detailed structure-activity profile for the calcium-releasing activity of cADPR, a method for the total syntheses of cADPR and its mimics is urgently needed.

### 8.2 INOSINE-BASED ANALOGUES OF cADPR WITH MODIFICATIONS IN THE NORTHERN RIBOSE MOIETY

Matsuda’s group developed a total synthetic approach to construct cADPR analogues (Shuto et al., 1998). They found an efficient intramolecular cyclization for the construction of cyclopyrophosphate (Scheme 8.2). Their strategy was to introduce a bulky group into the 8-position of purine nucleoside; this maneuver can restrict the conformation of diphosphate intermediate in a syn-form in which the two phosphate moieties are close to each other and facilitate intramolecular condensation in the presence of EDC. Matsuda and co-workers also reported that the molecular cyclization was completed by I2/3A MS as a promoter in pyridine in quantitative yield. In this way, they synthesized cIDP-carbocyclic-ribose (Fukuoka et al., 2000) (cIDPcR) and cyclic ADP-carbocyclic-ribose (Shuto et al., 2001) (cADPcR). cADPcR, in which the 4'-oxo in the northern ribose moiety of cADPR was substituted by a methylene group caused a significant release of \( \text{Ca}^{2+} \) in sea urchin eggs; in the same system, cADPcR was 3–4 times more potent than cADPR while the 8-substituted analogues were less efficacious.

![Scheme 8.2](image-url)
In contrast, in mammalian cells, cADPcR acted much more weakly as compared to the sea urchin system (Guse et al., 2002). However, it was found that cIDPcR showed only an insignificant Ca\(^{2+}\)-mobilizing effect (EC\(_{50} > 10^6\) M), and 8-Br-cIDPcR was almost inactive in sea urchin egg homogenates (Shuto and Matsuda, 2004). Using a similar synthetic approach, an \(N^1\)-penty1 substituted cIDPcR analogue was reported, but the biological data have not appeared yet (Galeone et al., 2002a). A novel class of stable cIDPR mimics (3–5) (Fig. 8.2) was synthesized, in which the \(N^1\)-ribosyl moiety is replaced by different configurational glycosyl moieties and the \(N^1\)-glycosyl linkage is shifted from the \(N^1-C^1\) to the \(N^1-C^2\) position of furanose to improve the stabilization (Huang et al., 2002a).

The key steps for the synthesis of \(N^1\)-glycosyl cIDPR analogues involve substitution at the \(N^1\) position of inosine and intramolecular cyclophosphorylation of the corresponding bisphosphate (Scheme 8.3). 1,4-anhydro-2-\(O\)-triflyl-3,5-O-benzylidene-\(t\)-xylitol,5-MMTr-3-\(O\)-allyl-2-\(O\)-triflyl-1,4-anhydro-\(d\)-xylitol and 5-\(O\)-benzoyl-3-\(O\)-allyl-2-\(O\)-triflyl-1,4-anhydro-\(d\)-ribitol were synthesized from the corresponding sugar (Scheme 8.3); the active triflates reacted with protected 8-bromo-inosine in the presence of K\(_2\)CO\(_3\) and 18-crown-6 by an S\(_{N2}\) substitution to give the key intermediate, \(N^1\)-glycosyl substituted inosine derivative, accompanied by a litter of \(O^6\)-substituted derivative. The intramolecular cyclization was performed by adding a solution of the cyclic precursor slowly over 20 hr, using a syringe pump, to a large excess of promoter I2/3A MS in pyridine at room temperature. The cyclic products (3, 4a, 4b, 5a, 5b) of triethylammonium salts were obtained, and some debrominated cyclic products (4a, 5a) were also found in addition to the normal products during the cyclization (Scheme 8.3).

Rat brain microsomes were one of the first mammalian cell preparations in which cADPR and IP\(_3\) were shown to trigger Ca\(^{2+}\) release independently of each other. The five novel cyclic nucleotide analogues (3, 4a, 4b, 5a, 5b) were tested for their abilities to release calcium from rat brain microsomes. In addition, Ca\(^{2+}\) mobilizing effects of extracellularly added compounds were analyzed in intact HeLa cells.

cADPR or the synthetic samples (3, 4a, 4b, 5a, 5b) induced Ca\(^{2+}\) release from brain microsomes with a threshold concentration of less than 15 \(\mu\)M under these conditions.
Scheme 8.3  Retrosynthesis of \( N^1 \)-glycosyl c-IDPR.
experimental conditions. Ca\(^{2+}\) was soon resequestered into the microsomes (< 400 sec). It is interesting to find that all samples except compound 4b induced the calcium release from brain microsomes under these conditions. The order of potency is 5a > 4a > cADPR > 5b > 3. It is evident that the 8-nonsubstituted analogues (4a, 5a) have more potent calcium-modulating activity than the corresponding 8-bromo cIDPR mimics (4b, 5b). Compound 4b was not able to induce the calcium release from brain microsomes even at the concentration of 60 \(\mu\)M. In addition to their activity in broken cell preparations, these cIDPR mimics modified in the \(N^1\)-ribosyl moiety were also active in intact HeLa cells. The natural compound cADPR (200 \(\mu\)M) itself exhibited no Ca\(^{2+}\) releasing activity, presumably because it cannot cross the cell membrane. However, samples 3, 4b, and 5b showed significant calcium release under the same conditions. The 8-unsubstituted analogue 5a only elicited calcium release at concentrations as high as 400 \(\mu\)M, and 4a exhibited no calcium release activity. Taken together, these data suggest that the modification in the northern ribose moiety of cADPR could retain the calcium release activity and also increase the membrane permeability of cADPR. It seems that the \(N^1\)-ribosyl moiety and the 6-amino group in cADPR are not critical structural factors for

![Scheme 8.4 Synthesis of cIDPRE.](image-url)
An $N^1$ substitution was carried out regioselectively on the protected inosine with 2-chloromethoxyethyl acetate in the presence of excess DBU to afford $N^1$-substituted inosine in 69% yield (Huang et al., 2002b). After deprotection and phosphorylation, the cyclic precursor was obtained. The intramolecular cyclization of bisphosphate intermediate was completed by using Matsuda’s strategy. Compared to the yields of the cyclization in the syntheses of cIDPR and its analogues, a more flexible $N^1$-ether strand of cIDPRE results in a lower yield for the synthesis. A series of 8-substituted cIDPRE were also reported (Gu et al., 2004). The pharmacological activities of cIDPRE and 8-Br-, 8-N$_3$-, 8-NH$_2$-, and 8-Cl-cIDPRE were analyzed in intact and permeabilized human Jurkat T-lymphocytes (Figs. 8.3 and 8.4). The results indicate that cIDPRE permeates the plasma membrane, releases Ca$^{2+}$ from an intracellular, cADPR-sensitive Ca$^{2+}$ store, and subsequently initiates Ca$^{2+}$ release-activated Ca$^{2+}$ entry. The Ca$^{2+}$-releasing activity of cIDPRE resulted in a lower efficacy of cIDPRE as compared to cADPR (Fig. 8.4d). Analogues of cIDPRE modified at C8 of inosine had strikingly differential effects: while 8-N$_3$-cIDPRE and 8-NH$_2$-cIDPRE were similarly effective in their agonistic activity as compared to cIDPRE, the

![Figure 8.3](image-url) Effect of the cIDPRE and 8-substituted cIDPRE analogues on Ca$^{2+}$ signaling in T cells [Ca$^{2+}_i$], was analyzed in Fura2-loaded Jurkat T cells. (A–C) Characteristic tracings in response to cIDPRE and two of its analogues. (D) Concentration–response relationship of cIDPRE and its analogues.
halogenated derivatives 8-Br- and 8-Cl-cIDPRE did not significantly elevate [Ca\(^{2+}\)]. Comparison of the calcium-release activities of different cIDPRE derivatives revealed a remarkable structure–activity relationship: While any functional group larger than the natural H-atom at C8 in cADPR turned the

**Figure 8.4** Effect of the cADPR mimic cIDPRE on Ca\(^{2+}\) signaling in T cells. (A) cIDPRE was added as indicated (final concentration 0.5 mM). (B) Concentration–response data are given as mean ± SEM (n = 3–5). (C) cIDPRE was added as indicated. Then, CaCl\(_2\) was readded. The dashed line shows a control experiment in which buffer was added instead of cIDPRE. (D) Cells were permeabilized. Lower panel: Vehicle (H\(_2\)O) and cIDPRE (100 μM) were added as indicated. Upper panel: Data are mean ± SEM (n ≥ 3; error bars are partially not visible because the values are smaller than the symbol size). (E) Characteristic confocal ratiometric pseudocolor images of a single Jurkat cell and magnifications of a defined subcellular region are shown. See color plates.
molecule into a cADPR antagonist—for example, 8-NH₂-cADPR, 8-Br-cADPR, 8-CH₃O-cADPR, 8-CH₃-cADPR—in the case of the cyclic IDP-ribose mimics, different functional groups at the 8-position resulted in completely different biological activities.

Thus, cIDPRE, 8-NH₂-, and N₃-cIDPRE are novel and important agonist tools similarly suitable for studying the structure–activity relationship of cADPR and for use as membrane-permeant cADPR mimics in intact cells. As an example subcellular Ca²⁺ signals can be induced by cIDPRE, as demonstrated by confocal Ca²⁺ imaging at the single cell level (Fig. 8.4e). A role for cADPR-mediated amplification of subcellular calcium signals and spatiotemporal signal propagation has been analyzed using the membrane-permeant cIDPRE (Kunerth et al., 2004).

8.3 cIDP-DE: A MIMIC OF cADPR WITH MODIFICATIONS IN THE NUCLEOBASE AND NORTHERN AND SOUTHERN RIBOSE MOIETIES

Concerning modifications on the southern ribose moiety of cADPR, it was found that the 3'-hydroxyl group was essential for calcium release in the sea urchin egg system, whereas the 2'-hydroxyl was not. O-Methylation of the 3'-hydroxyl group generated an antagonist (Ashamu et al., 1997) (Fig. 8.5). The stable cyclic aristeromycin diphosphoribose (c-Aris DPR) contains a replacement of the southern ribose with a carbocyclic five-membered ring and is an agonist slightly more potent as compared to cADPR in the T-lymphocyte system (Bailey et al., 1996) (Fig. 8.5). The Mayol group reported the synthesis of a cIDPcR analogue in which a pentyl group was included instead of the southern ribose, but no biological data were described (Galeone et al., 2002b) (Fig. 8.5). Comparison of the calcium-release activities of cIDPRE and cADPR revealed that the greater flexibility of the whole molecule introduced by replacement of the northern ribose by the ether strand either still allows fitting into the binding site or allows the functional groups to support a conformation much more like cADPR. Because of the same structure of the northern and southern ribose occurring

![Figure 8.5](image_url)
in cADPR, it appeared interesting to synthesize the cIDPR analogue with replacement of both northern and southern riboses using two ether strands. Synthesis and biological evaluation of the respective compound $N^1-[(\text{phosphoryl-O-ethoxy})$-methyl]-$N^9-[(\text{phosphoryl-O-ethoxy})$-methyl]-hypoxantine-cyclicpyrophosphate (cIDP-DE) was reported recently (Guse et al., 2005). The key steps for the synthesis of cIDP-DE include the synthesis of the $N^9$-substituted hypoxanthine (Scheme 8.5) and the intramolecular cyclization of the bisphosphate intermediate (Scheme 8.6). It was very interesting to find that the intramolecular cyclization of bisphosphate intermediate was completed in 61% yield, although more flexible $N^1$ and $N^9$ ether strands probably resulted in a larger spatial distance of the two phosphate groups (Scheme 8.6). Interestingly, cIDP-DE displayed moderate fluorescence ($\lambda_{ex}$ 278 nm, $\lambda_{em}$ 355 nm) in water at room temperature. However, no fluorescence was observed in cIDPR or cIDPRE. The reason for these spectroscopic properties might be the more flexible $N^1$ and $N^9$ ether strands in cIDP-DE, which allow an easier $\pi$-electronic transition in the purine ring. The fluorescent character of cIDP-DE might be used for the mechanistic studies of cADPR-mediated calcium release.

Interestingly, the evaluation of the biological activity of cIDP-DE indicates that it is a membrane-permeant analogue of cADPR with agonistic properties. In saponin-permeabilized Jurkat T cells, cIDP-DE released $\text{Ca}^{2+}$ from intracellular stores but was much weaker as compared to cADPR. Regardless of this weak direct effect, a pronounced activity of cIDP-DE on $\text{Ca}^{2+}$ signaling was observed in intact Jurkat T cells. The $\text{Ca}^{2+}$-mobilizing activity of cIDP-DE was also observed in $\text{Ca}^{2+}$ imaging experiments. The biological activity of cIDP-DE in Jurkat T cells was quite similar to cIDPRE in terms of (i) $\text{Ca}^{2+}$ release activity in permeabilized cells and (ii) induction of both $\text{Ca}^{2+}$ release and $\text{Ca}^{2+}$ entry in a $\text{Ca}^{2+}$-free/$\text{Ca}^{2+}$-reintroduction protocol.

The $\text{Ca}^{2+}$-mobilizing activity of cIDP-DE was also confirmed in cardiac myocytes freshly prepared from mouse hearts. In the myocytes, addition of cIDP-DE induced global $\text{Ca}^{2+}$ signaling with a very slow onset, obviously due to a slower uptake into the cell as compared to the Jurkat T cells. However, spatiotemporal

**Scheme 8.5** Synthesis of $N^9$-substituted hypoxanthine.
Scheme 8.6 Synthesis of cIDP-DE
Ca\(^{2+}\) signaling pattern including local spots with oscillating \([\text{Ca}\(^{2+}\)]_i\) and global Ca\(^{2+}\) waves were observed upon cIDP-DE addition (Fig. 8.6).

Taken together, this similarity of cIDP-DE and cIDPRE, especially in induction of Ca\(^{2+}\) signaling in intact T cells, indicates that the further replacement of the southern ribose by the second ether bridge in cIDP-DE did not significantly influence its biological activity. Compounds of the type cIDP-DE or cIDPRE may thus be used as membrane-permeant agonists of the cADPR-/Ca\(^{2+}\)-signaling system in intact cells.

8.4 MODEL FOR THE INTERACTION OF cADPR WITH ITS RECEPTOR

The structure of cADPR in solution was investigated using NMR by Sekine and co-workers (Wada et al., 1995). They reported a predominant C2'-endo conformation.
for the southern ribose unit and a flat northern ribose conformation in the cADPR structure in solution. These structural characteristics in solution are very similar to those of the X-ray crystallographic analysis (Lee et al., 1994). However, in the structure of cIDP-DE, two more flexible ether strands are included instead of both the northern and southern riboses of cADPR. Nevertheless, cIDP-DE remains biologically active for calcium release in T-lymphocytes and cardiac myocytes. Although the structure of the receptor for cADPR is not clear so far, the data reported above may indicate some clues of the structural requirements for the design of future cADPR agonists.

To model the structure of cIDP-DE, its conformation was minimized by quantum mechanics calculation (Fig. 8.7). The different ways of superimpositions

Figure 8.7 Superimposition of minimized conformations of cADPR and cIDP-DE. The conformations of cADPR (yellow) and cIDP-DE (pink) were minimized using the density functional theory (DFT) method. (A) Superimposition based on nucleobases; (B) superimposition based on diphosphates; (C) superimposition based on whole molecules. See color plates.
of the energetically minimized conformations of cIDP-DE (pink color) and cADPR (yellow color) clearly show that both compounds are characterized by an almost identical three-dimensional structure (Fig. 8.7). Despite the replacement of the northern and southern ribose of cADPR by two flexible ether strands, the distance and orientation between nucleobase and diphospho-bridge were maintained and the whole conformation was still almost unchanged, as visualized by superimposition of either the nucleobase (Fig. 8.7A) or the diphospho-bridge (Fig. 8.7B). The optimal superimposition of the two whole molecules in addition suggests that the two ether groups in cIDP-DE nicely replace the corresponding C1–O–C4 units of the two riboses (Fig. 8.7C). Because of the flexible ether strands used as linkers in cIDP-DE, it may well be that the binding pocket of the cADPR receptor helps induce the correct “binding conformation” during the process of ligand binding, a process known in enzymology as “induced fit.”

In summary, recent data from our own and other laboratories indicate the following crucial elements for the structure–activity relationship of cADPR: (1) 7-Deaza-cADPR and 3-deaza-cADPR are agonists of cADPR in the sea urchin egg system. It seems that the 3-deaza- or 7-deaza-purine ring in cADPR would retain the conjugated plane shape and be recognized by the sea urchin egg cADPR receptor. (2) cADPR analogues with substitutions at the 8-position of the adenine ring (including 8-substituted 7-deaza-cADPR) were found to be antagonists of cADPR in both sea urchin eggs and many high eukaryotic cell types. However, in the case of the \( N^1 \)-coupled cyclic inosine pyrophosphate-ribose mimics (\( N^1 \)-cIDPR), 8-Br-cIDPR acts as a membrane permeant agonist of cADPR-induced \( Ca^{2+} \) signaling. (3) Despite the more radical structure alterations in cIDPRE and cIDP-DE, these analogues are agonists of cADPR. The \( N^1 \)- and \( N^9 \)-ribosyl moieties and the 6-amino group in cADPR obviously are not critical structural factors for biological activity in mammalian cells. (4) Modification of cIDPRE at C8 of inosine had strikingly different effects, with 8-N\( _3 \)-cIDPRE and 8-NH\( _2 \)-cIDPRE being effective agonists and 8-Br- and 8-Cl-cIDPRE without biological activity. (5) With modification in the pyrophosphate moiety of cADPR using three phosphate groups in line, the triphosphate mimic (Fig. 8.1) was a highly effective agonist in bovine brain microsomes (Zhang et al., 1996).

Therefore, the proposed model for the interaction of cADPR with its receptor protein would be as follows: (1) The binding pocket may specifically recognize the purine base or other aromatic rings with similar shape as purine on the one side and the di- or triphosphate bridge or multinegative groups on the other side of the ligand. (2) The two parts of the binding pocket located close to the northern and southern ribose moieties may be sufficiently variable, for example, one electronegative atom, such as oxygen, and two spacers of the right length appear to be sufficient for binding of the ligand and for weak biological activity. (3) The binding pocket interacting with the 8-position in purine moiety may be an active trigger whose function—agonist or antagonist—depends on the structure of adenine or hypoxanthine in the binding molecule. However, as long as the binding protein for cADPR has not been identified, this remains speculation.
ACKNOWLEDGMENTS

We wish to thank all former and current members of our groups who have contributed to this project. The project is supported by National Natural Science Foundation of China (to LHZ), The Ministry of Science and Technology of China (to LHZ), the Deutsche Forschungsgemeinschaft (to AHG), the Hertie-Foundation (to AHG), the Wellcome Trust (to AHG), and the Deutsche Akademische Austauschdienst/Chinese Scholarship Council (to LHZ and AHG).

REFERENCES

9.1 INTRODUCTION

Neurodegenerative disease—including diseases such as Huntington’s disease (HD), Parkinson’s disease (PD), and Alzheimer’s disease (AD)—is caused by chronic degeneration of the nervous system. HD is a terminal progressive neurodegenerative disorder. HD patients mainly develop chorea (uncontrollable jerking movements) and psychiatric dysfunction that generally includes intense irritability and the loss of emotional control. PD is also a terminal progressive neurodegenerative disorder whereby patients experience a debilitating loss of movement functionality that includes muscle rigidity, bradykinesia (slowness of movement), and resting tremor. AD is a form of dementia that invariably leads to complete loss of all cognitive abilities and premature death. The mechanisms of these diseases are still unclear, so there are no specific and effective strategies to treat them. Currently, drugs used to treat HD, PD, and AD mainly include neuroprotectants, nerve growth factor, and other small organic molecules such as neuroimmunophilin ligands.
9.1.1 Neuroprotectants

There are three main mechanisms of neuronal cell death that may act separately or cooperatively to cause neurodegeneration: metabolic compromise, excitotoxicity, and oxidative stress (Alexi et al., 2000): “Metabolic compromise of neurons is caused by stroke, asphyxiation, hypoglycemia, and certain respiratory poisons.” These neurotoxic poisons are mainly mitochondrial poisons and cause a depletion of ATP, which leads to preferential neurodegeneration in the basal ganglia. Excitotoxicity occurs due to a dysfunction of excitatory amino acid neurotransmission—usually a stimulation of glutamate receptors that becomes pathological. The toxic effects of excitatory amino acids are mainly due to the activation of the N-methyl-D-aspartate (NMDA) glutamate receptor, which leads to an influx of Ca²⁺ and then toxic overloading (Greene and Greenamyre, 1996). Per Alexi et al., “Oxidative stress is due to the actions of highly reactive free radicals such as the ROSs superoxide anion and hydroxyl radical, and the RNS peroxynitrite.” These free radicals can lead to the depletion of ATP (Szabo et al., 1996) and induce neuronal cell death (Morel et al., 1999). Alexi et al. add: “The lethal triplet of metabolic compromise, excitotoxicity, and oxidative stress may also act cooperatively in causing neuronal cell death. For example, metabolic impairment may elicit secondary excitotoxicity.” Metabolic compromise also may cause oxidative stress by inducing the production of free radicals both from the electron transport chain and due to the burden of increased intracellular Ca²⁺ on mitochondria function. The oxidative stress may also cause metabolic impairment and initiate excitotoxic pathways. The byproduct HNE that yields from oxidative stress-induced lipid peroxidation can impair glucose transport, which can lead to energetic failure and render neurons more sensitive to excitotoxicity (Pedersen et al., 1999). Thus, all the materials inhibiting or eliminating those mechanisms may have the potential to treat neurodegenerative diseases. These materials that can protect neurons from damage but not promote the regeneration of nerves are termed neuroprotectants.

A characteristic of neurodegenerative disease is formation of β-amyloid. On one side, β-amyloid can activate the Ca²⁺ channel of cell which results in the accumulation of intracellular Ca²⁺ (Hölscher, 1998). This intracellular Ca²⁺ then induces further activation of Ca²⁺-dependent proteases and free-radical generation to lesion neurons. The glutamate mediate system also participates in this procedure. Thus, glutamate receptor antagonists, Ca²⁺ channel inhibitors, and inhibitors of related enzymes all have the potential to treat neurodegenerative disease. On the other side, the accumulation of β-amyloid will trigger abnormal protein phosphorylation reactions responsible for dysfunction and eventual death of neurons in the brain (Jin and Saitoh, 1995). The depositing of β-amyloid is modified by advanced glycation end products (AGEs) (Durany et al., 1999), which implies that the inhibitors of AGEs may also have the potential to treat neurodegenerative disease. The ability of glutamic acid to kill neurons (excitotoxicity) seems to be mediated, in most cases, by an interaction with NMDA receptors, leading to an uncontrollable rise in intracellular calcium concentrations and thence cell lysis and death (Doble, 1995). Since glutamaterigic neurodegeneration appears to better explain the physiological and anatomical hallmarks of neurodegeneration, the focus of research has shifted toward
NMDA receptor inhibitors, such as riluzole (Arendt et al., 1998) and memantine (Parsons et al., 1999). Besides the NMDA receptor family, the AMPA/kainate receptor family is also involved in neurodegeneration and may also have the potential for neurodegenerative disease therapy. Antioxidants are another type of neuroprotectant. Antioxidants can clean the highly reactive free radicals such as the ROSs superoxide anion, hydroxyl radical, and the RNS peroxynitrite, which cause oxidative stress. There are other materials involved in oxidative stress, such as the products of inflammatory reactions—that is, prostaglandins (PGs: PGE1 and PGA1), cytokines, and complement proteins. Thus, the combination of multiple antioxidants and NSAIDs may be more beneficial in the prevention of neurodegenerative disease. Drugs that act mainly on the acetylcholine neurotransmitter system—such as the acetylcholine esterase inhibitors tacrine, velnacrine, or Aricept—are also used to treat neurodegenerative disease, especially AD, because it was believed that the main cause for AD is a degeneration of basal brain nuclei that use acetylcholine as a neurotransmitter.

Considerable efforts have been made to find more effective neuroprotectants, and there must be more progress in this active research field. All the known neuroprotectants can only relieve symptoms or slow the neurodegeneration of neurons, but they cannot stop the progress of neurodegeneration or repair the damaged neurons. Thus, using neuroprotectants cannot completely cure neurodegenerative disease and so is not an ideal therapeutic method.

9.1.2 Nerve Growth Factor (NGF)

Nerve growth factor is an endogenous factor that can stimulate the growth of neurons and the extension of nerve fibers. NGF is a macromolecule including Zn$^{2+}$ and two proteins that consist of three subunits: α unit, β unit, and γ unit. NGF is unstable in acidic and basic environments and dissociates when pH $< 5$ or pH $> 8$. Only the β unit has a biofunction that consists of 118 residues. The net charge on its surface is electropositive since positive groups are greater than negative groups. Whether in vivo or in vitro, the neurons cannot survive and grow without the existence of NGF. There are several cell types that can secrete NGF, including fibroblasts, smooth muscle cells, skeletal muscle cells, glial cells, neurilemma cell, and so on. These cells mainly come from mesenchyme and are related to the normal function of neuraxons. There are specific receptors on the membranes of neuraxons that can recognize and complex with NGF and then transport into cells through a positive transport procedure. The endogenous nerve growth factor acts on sensory and sympathetic neurons of the peripheral nervous system and on basal forebrain and striatal cholinergic neurons of the central nervous system. Evidence showed that NGF effectively attenuates lesion-induced cholinergic deficits and cognitive impairments in animal models. In addition, a clinical study with chronic NGF treatment in an AD patient showed promise. Olson et al. showed that NGF treatment increased blood flow, $^{11}$C nicotine uptake, and $^{11}$C binding in the cerebral cortex. In addition, NGF infusions normalized EEG patterns and improved performance in word-recognition tests (Lapchak, 1993). Besides NGF, other neurotrophic factors such as activity-dependent neurotrophic factor (ADNF) and glial cell line-derived neurotrophic factor (GDNF) all show
notable activity in neuroprotection and neuroregeneration (Bohn, 1999). A short peptide sequence derived from ADNF (a peptidergic active site) protected neurons from death associated with a broad range of toxins, including those related to Alzheimer’s disease, the human immunodeficiency virus, excitotoxicity, and electrical blockade. This peptide is proposed as a lead compound for drug development against neurodegeneration (Gozes and Brenneman, 1996). Although using endogenous factor to stimulate the regeneration of nerves should be the best therapeutic method for neurodegeneration disease, the common shortcomings of biopolymers such as having low bioavailability, degrading easily, passing the brain–blood barrier (BBB) with difficulty, and being difficult to take orally reduce their clinical application in neurodegenerative disease therapy. Studies addressing this problem are applying viral vector-mediated transfer of the NGF, GDNF, and ADNF genes to the CNS in order to deliver biosynthesized neurotrophic factors to a specific location in a chronic manner. Recent studies suggest that these GDNF gene-therapy approaches are effective in rat models of Parkinson’s disease (Fink et al., 1997). However, the effectiveness of gene therapy in humans still needs to be proven.

9.1.3 Neuroimmunophilin Ligands

The ideal way to treat neurodegenerative disease would be to use NGF-like organic compounds that can be taken orally, are stable in vivo, and pass through the BBB easily. The ligands of neuroimmunophilin are the best choice. Neuroimmunophilin is primarily a type of immunophilin that complexes with immunosuppressants such as FK506, CsA, and rapamycin (RAPA) and produces immunosuppressive responses. These proteins are also known as peptidylprolyl cis–trans isomerases (PPIases) since they can catalyze the interconversion of cis and trans rotamers of amide bonds adjacent to proline residues in peptidic substrates (Fischer et al., 1984). The enzymes complex with CsA termed cyclophilin (CyP), and the enzymes complex with FK506 or RAPA termed FKBPs (FK506 bind proteins).

FKBPs are a family of proteins that exist extensively in nature. The three-dimensional structure of this family is highly conserved. The sequence homology is more than 50% among different species. A large number of immunophilins belonging to the FKBP family have been discovered in the past several years. Currently, there are about 20 known FKBPs in nature. FKBPs immunophilins known to be present in humans include FKBP12, FKBP12.6, FKBP13, FKBP25, FKBP37, FKBP51, and FKBP52 (by convention, members of the FKBP family are named by appending to the prefix “FKBP” the apparent molecular weight in kilodaltons). In addition to the activity of PPIases, FKBPS are also involved in some other biological processes, such as translation (FKBP25, FKBP37), protein secretion of endoplasm (FKBP13), single mediation of EGFR on the cell membrane (FKBP12), the influx of Ca\textsuperscript{2+} intracellular (FKBP12), and involvement in the formation of steroid receptor complexes (Smith et al., 1993). Functions of the other members of the FKBP family are still unknown. FKBP12 is the most well-known member of the FKBPs. It complexes FK506 and further inhibits the activity of Calcineurin phosphatase, which results in the interruption of proliferation of T-lymphocyte (Liu et al., 1991). FKBP51 acts the same way as FKBP12 (Baughman et al., 1995). The first three-dimensional structure of the
FKBPs is the complex of FKBP12 and FK506 (Van Duyne et al., 1991; Lepne et al., 1992). The binding site consists of residues Tyr26, Phe36, Asp37, Arg42, Phe46, Glu54, Val55, Ile56, Trp59, Tyr82, His87, Ile91, and Phe99. All these residues make a hydrophobic cave and complex with FK506 through hydrophobic and electronic interaction.

Steiner et al. (1992) showed the high density of FKBP mRNA in brain tissue and a close homology between the structure of NINA A, a protein in the neural retina of *Drosophila*, and cyclophilin. This implies a physiological role for the immunophilins in the nervous system. The results of Dawson et al. (1993) showed that FK506 enhanced phosphorylation of nitric oxide synthase and protected against glutamate neurotoxicity. The neuroprotective effect of FK506 presumably involves the inhibition of calcineurin, preventing the dephosphorylation of nitric oxide synthase and its subsequent activation. GAP-43, a growth-associated protein, was also found to be a physiologically important substrate for calcineurin. Also, calcineurin and protein kinase C may regulate the levels of free calmodulin available in neurons (Liu and Storm, 1989). In addition, regeneration of sciatic and facial nerves is always accompanied by the increase of mRNA of GAP-43. All these results imply a physiological role for the FKBPs in the regeneration of the nervous system. Under the direction of this hypothesis, a small organic compound with neuroregenerative activity was found in FKBP ligands. The findings of Lyons et al. (1994) that FK506 enhances neurite outgrowth in PC12 rat pheochromocytoma cells and in rat sensory ganglia by increasing sensitivity to nerve growth factor initiate a new era of finding small organic neuroregeneration compounds.

### 9.2 NATURAL NEUROIMMUNOPHILIN LIGANDS

Being well-known targets of immunosuppressant for many years, many neuroimmunophilin ligands have been found before FKBPs became the targets of neuroregeneration. Among them are the well-known natural neuroimmunophilin ligands FK506, rapamycin, and cyclosporine A.

#### 9.2.1 FK506

FK506 (Fig. 9.1A) is a 23-member ring macrolide lactone that was discovered in 1984 by scientists at Fujisawa Pharmaceutical Company in the fermentation broth of the filamentous bacterium, *Streptomyces tsukubaensis* (Kino et al., 1987), launched in 1993 for the treatment of immunosuppression in organic transplantation (Starzl et al., 1990). It acts by binding to its receptors, FK506-binding proteins (FKBPs), which in turn can bind to and regulate a Ca$^{2+}$-dependent phosphatase, calcineurin; it also inhibits the activity of calcineurin and inhibits interleukin-2 production in activated T cells. Lyons et al. (1994) found that FK506 was significantly effective in neuroregeneration *in vitro*. It promotes neurite outgrowth both in PC12 cells and in sensory neuronal cultures of dorsal root ganglia. It can also stimulate functional regrowth of damaged sciatic, cortical cholinergic, dopamine, and 5-HT neurons (Snyder et al., 1998). The search results of other researchers showed FK506 can
improve H₂O₂-induced cell damage by increasing glutathione levels in NG108-15 cells (Tanaka et al., 2001), increase the rate of axonal regeneration in rat sciatic nerve (Gold et al., 1995), increase neuronal expression of GAP-43, improve functional recovery after spinal cord injury in rats (Madsen et al., 1998), and enhance the survival of cultured and grafted rat embryonic dopamine neurons (Gastilho et al., 2000). In an in vivo model of focal cerebral ischemia, Sharkey and Butcher (1994) found that FK506 exhibited a powerful neuroprotective activity. Administered up to 60 min post-occlusion, the minimum effective neuroprotective dose is comparable with the immunosuppressant dose in humans, suggesting that FK506 may have clinical potential for the treatment of stroke. Kitamura et al. (1994) found that in the MPTP-induced dopamine depletion in the striatum of young C57BL/6 mice model, pretreatment with FK-506 significantly protected MPTP-induced DA depletion in the striatum.

9.2.2 Rapamycin

Rapamycin (Fig. 9.1B) is another kind of immunosuppressant that can bind to FKBP12. The complex of rapamycin and FKBP12 is more stable than that of FK506, and rapamycin can inhibit the neuroprotective activity of FK506. Recent research results showed that rapamycin may be the one of the only immunosuppressants that can stimulate the regeneration of nerves (Parker et al., 2000). In the study, rapamycin markedly increased neurite outgrowth in PC12 cells in the presence of a low concentration of nerve growth factor, while FK506 (0.1–1000 nM) had little effect on neurite outgrowth in PC12 cells in either the presence or absence of nerve growth factor. The authors implied that this result is due to the ability of rapamycin to inhibit cell cycle progression. The results of Costantini and Isaacson (2000) showed that rapamycin can antagonize the elongation effects of DA neurites of FK506 and cyclosporine A.

9.2.3 Cyclosporine A

Cyclosporine A (CsA, Fig. 9.1C) is a potent immunosuppressive drug that has widespread clinical uses in organ transplantation and the treatment of autoimmune disorders. Cyclosporine A is a cyclic peptide that consists of 11 amino acids. It complexes with cyclophilin and leads to immunosuppression. Research results show that the cyclophilin level in neuronal areas is higher than the level in T-cells (Lad et al., 1991), and this initiates the study of its potential function on the nervous system. Uchion et al. (1995) found that cyclosporine A can dramatically ameliorate CA1 hippocampal damage following transient forebrain ischemia in the rat. Other cyclosporine A-mediated neuroprotection includes attenuation of mitochondrial swelling, stabilization of the permeability transition pore, and neurofilament compaction (Herdegan et al., 2000). In animal models of nerve grafts, cyclosporine A has been proved to increase the survival of grafted rat embryonic dopamine neurons (Gastilho et al., 2000). In addition, cyclosporine A (15 mg/kg) can also inhibit xenograft rejection in rats and enhanced tissue survival in vitro (Wennberg et al., 2001). In a study that investigates whether cyclosporine A is able to facilitate axonal
Figure 9.1 Structure of FK506 (a), rapamycin (b), and cyclosporine A (c).
regeneration in rats, strong evidence of morphological axonal regeneration was observed 15 days after surgery in all cyclosporine A-treated animals (Palladini et al., 1996). The function of enhancing neurite branching or elongation is also proved (Costantini and Isaacson, 2000).

In general, FK506 has significant potential for neuroprotection, neuroregeneration, and passing the brain–blood barrier easily. Cyclosporine A also has neuroprotective and neuroregenerative activity but has to pass the brain–blood barrier. Rapamycin has no neuroprotective activity but has significant neuroregenerative potential. Thus, among the natural immunophilin ligands, FK506 should be the best candidate for treatment of neurodegenerative disease. However, immunosuppressive responses limit its clinical application.

The above-mentioned neuroregenerative and neuroprotective properties of FK506 and other immunophilin ligands suggests an exciting new therapeutic approach for the treatment of neurological disorders. Promising drug candidates should be small-molecule FKBP ligands with oral bioavailability, which are devoid of immunosuppressive activity and are able to cross the blood–brain barrier. Many efforts have been made in this area, and hundreds of novel nonimmunosuppressant neuroimmunophilin ligands have been found.

9.3 NEUROREGENERATIVE MECHANISM OF NEUROIMMUNOPHILIN LIGANDS

The neuroregenerative mechanism of neuroimmunophilin ligands is poorly known. The original mechanism is based on the calcineurin and Ca\(^{2+}\) channel. Dawson et al. (1993) assumed that the complex of FKBP12 and FK506 inhibits the function of calcineurin and in turn inhibits the activity of nitric oxide synthase (NOS), which makes FK506 produce the neuroprotective activity in focal ischemia. However, the evidence that nonimmunosuppressive analogues of FK506 that bind FKBP12 but can’t complex with calcineurin also exhibit neuroprotective activity denied the calcineurin-involved neuroregenerative mechanism.

Since FKBP12 is the primary target of many nonimmunosuppressive ligands, the relevant mechanism may involve this protein. This mechanism has been well accepted for several years. To examine if FKBP12 is required for neurotrophic effects of immunophilin ligands, Guo et al. (2001) cultured dopaminergic neurons from FKBP12 knockout mice and found that FK506 still protects dopaminergic neurons against MPP\(^+\) toxicity. Gold et al. also found that FK506 maintains its neurotrophic activity in primary hippocampal cell cultures from FKBP-12 knockout mice (Gold et al., 1999). These two independent studies proved that without the participation of FKBP12, immunophilin ligands can also exhibit neuroregenerative activity. Recently, a novel nonimmunosuppressive analogue of FK506 named V-13661 exhibited the same neuroregenerative activity. However, this compound can’t bind to FKBP12. This result uncovered that the neuroregenerative activity of immunophilin ligands is not related to FKBP12 (Costantini et al., 2001). All the above results show there may be another mechanism for neuroregenerative effects of immunophilin ligands.
The results of Gold showed that FKBP52 may be a key protein in the mechanism of neuroregenerative effects of FKBP ligands (Gold et al. 1999). FKBP52 is a member of the FKBP family, which consists of 459 amino acid residues. It contains three function domains, named D1, D2, and D3, respectively. Each domain is connected with a short hydrophilic amino acid sequence. The sequence homology of D1 (A1-A149, FKBP52-I) and FKBP12 is about 55%. This domain can bind to FK506 and be of the same PPIase activity as FKBP12. The D1 domain can also bind to ATP and GTP, and it phosphorylates under the catalytic of tyrosine protein kinase II. Also, FKBP52 is a component of a subclass of steroid receptor complexes (Tai et al., 1992). The binding of FKBP52 and FK506 is weak, and the complex cannot inhibit the effect of calcineurin. This is the main difference between FKBP12 and FKBP52. Since it participates in the steroid receptor complex, people presumed that the neuroregenerative effects of immunophilin ligands may relate to the steroid receptor/hsp-90/FKBP52 complex.

Two related research results support this presumption: The first is molybdate, which acts in contrast to steroids (Raaka et al., 1985) and can inhibit the neuroregenerative effects of FK506; the other is that the geldanamycin, which inhibits the complex of FKBP52/p23/Hsp-90 and has neuroregenerative activity (Paratt and Toft, 1997). These results imply that FK506 may complex with FKBP52, which in turn disassociated the steroid receptor complex, and the disassociated components may further act as neuroregeneratants. The findings of Ishimoto et al. (1998) that Hsp-90 can promote the regeneration of telencephalic neurons and spinal neurons proved the rationality of this presumption. But how it stimulates the growth of nerves is still unclear. Gold et al. (1997) found that injecting FK506 10 mg/kg daily could significantly increase the mRNA level of GAP-43, a protein related to the growth of neurons.

Based on this evidence, Gold presumed a more rational mechanism as shown in Fig. 9.2 (Gold et al., 1999): “Binding of steroid hormones (steroid) to the steroid receptor complex leads to dissociation of the complex; this enables the ligand binding component to translocate to the nucleus, where it binds to the steroid response element (SRE) on steroid hormone-response genes. In addition, steroid hormones, by disrupting the complex, act like geldanamycin and FK506 to activate hsp-90. This activation leads to a conformational change that dissociates p23, one proposed mediator of nerve regeneration. Hsp-90 may also stimulate mitogen-activated protein kinase/extracellular signal-regulated kinase-2 (MAP kinase/ERK2) pathways, providing a potential cross-talk with signal transduction pathways for neurotrophic factors (e.g., NGF). Downstream effectors ultimately mediating nerve regeneration include c-jun and GAP-43 (which contains an AP-1 site in its promoter region for activation by c-Jun homodimers), both of which show a markedly increased expression during nerve regeneration that is further augmented by FK506. In contrast, molybdate prevents dissociation of the complex and inhibits the ability of all ligands to increase neurite outgrowth (nerve regeneration).”

Based on the FKBP52-mediated hypothesis, Baulieu and co-workers presumed that the neurotrophic activity of immunophilin ligands may relate to the mediation of phytanoyl-COA-α-hydroxylase (PAHX), a particular and specific target of FKBP52 (Chambraud et al., 1999). PAHX is a peroxisomal enzyme. Per the study: “Inactivation of this enzyme is responsible for Refsum disease in humans. The protein also
corresponds to the mouse protein LN1, which could be involved in the progress of lupus nephritis." The results showed that "PAHX has the physical capacity to interact with the FKBP12-like domain of FKBP52, but not with FKBP12..." Also, "...the specific association of PAHX and FKBP52 is maintained in the presence of FK506. Although the function of PAHX/FKBP52 complex is still unclear, this observation suggests that PAHX is a serous candidate for studying the cellular signaling pathway(s) involving FKBP52 in the presence of immunosuppressants."

9.4 NONIMMUNOSUPPRESSANT NEUROIMMUNOPHILIN LIGANDS

As mentioned above, the best way to treat neurodegenerative disease is by using small organic molecular NGF-like compounds. Neuroimmunophilin ligands seem to be the best choice at present. However, the unwanted immunosuppressive response limits their clinical application. It invokes the research of neuroimmunophilin ligands, which satisfies the demand of neurodisorder disease treatment without immunosuppressive activity.

Design and synthesis of low-molecular-mass, monofunctional FKBP ligands first provided compounds with structural modification of the FKBP binding domain of FK506 and the immunosuppressive mechanisms of FK506. FK506 is a bifunctional molecule whose structure is divided into two distinct parts: the part that binds to FKBP12, termed binding domain, and the part that binds to calcineurin, termed effective domain (Fig. 9.3a). FK506 complexes with FKBP12, and the complex further binds to calcineurin. The immunosuppressive response is produced by this trinity (Dumont et al., 1990). Based on the immunosuppressive mechanisms of
FK506, scientists assumed that blocking the binding of the FK506-FKBP12 complex and calcineurin may remove the immunosuppressive activity of neuroimmunophilin ligands. According to this suggestion, Steiner et al. (1997a) designed and synthesized a nonimmunosuppressive analogue of FK506, that is, L-685818. The structural difference between L-685818 and FK506 is evidenced by the ethyl substitution of FKBP12 and calcineurin are displayed in cartoon model. FK506 is displayed in black stick model. (b) In general the minimal binding domain can be divided into four regions: cyclohexylethyl, pipecolyl, dicarbonyl, and pyranosyl.

FK506, scientists assumed that blocking the binding of the FK506-FKBP12 complex and calcineurin may remove the immunosuppressive activity of neuroimmunophilin ligands. According to this suggestion, Steiner et al. (1997a) designed and synthesized a nonimmunosuppressive analogue of FK506, that is, L-685818. The structural difference between L-685818 and FK506 is evidenced by the ethyl substitution of
C-21 and the hydroxyl substitution of C-18. This variation blocks the binding of calcineurin at the effective domain of FK506 and leads to the loss of immunosuppressive response and neurotoxicity of FK506 (Dumont et al., 1992). It also exhibits the same neuroregenerative activity as FK506. This find is of great significance to research in low-molecular-weight neuroregenerative agents because it gets rid of the primary therapeutic obstacle of neuroimmunophilin ligands.

9.4.1 Classical Compounds of Nonimmunosuppressant Neuroimmunophilin Ligands

Following the discovery of L-685818, many analogues of FK506 were designed and synthesized. Among those exciting compounds, the classical candidates are L-685818, V-10367, GPI-1046, and VX-853.

L-685818 (Fig. 9.4a) is the first immunophilin ligand without immnosuppressive response as mentioned above. Compared to FK506, the only structural difference

Figure 9.4  (a) L-685818, (b) VX-853, (c) V-10367, and (d) GPI-1046.
of L-685818 is the additional hydroxy group in C-18 and the ethyl substitution of C-21 propyl. It cannot bind to calcineurin and has no immunosuppressive response. Research results showed that it exhibited the same effect in sciatic nerve regeneration and function recovery as FK506 does. VX-853 (Fig. 9.4b) is a novel compound developed by Vertex with neuroregenerative and neuroprotective responses. It has finished its phase II trial on the treatment of nerve disorder caused by diabetes (Herdegan et al., 2000).

V-10,367 (Fig. 9.4c) was developed by Armistead et al. (1995), who adopted structure-based design strategy to discover FK506 analogues. It highly inhibits the PPIase activity of FKBP12 with \( K_i = 0.5 \) nM. The complex of V-10,367 and FKBP12 cannot bind to calcineurin, so there is no immunosuppressive activity. In vitro, V-10,367 has been shown to potentiate neurite outgrowth induced by submaximal concentrations of NGF from both PC12 cells and the human neuroblastoma cell line SH-SY5Y. It can also accelerate the function recovering of rat sciatic nerves in vivo. Per Costantini et al. (1998), V-10,367... “increased the number of neurites extended by tyrosine hydroxylase positive (TH+) DA neurons in embryonic day 14 primary DA neuronal cultures. In contrast, the immunosuppressive immunophilin ligand FK506 increased the length of TH+ neurites. After oral administration in MPTP-treated mice, V-10,367 completely protected against MPTP-induced loss of striatal TH+ axonal density, while FK506 did not.”

GPI-1046 (Fig. 9.4d) is the most famous nonimmunosuppressant analogue of FK506. It exhibits highly neuroregenerative activity. Per Steiner et al. (1997b): “In vitro, ... GPI-1046 elicited neurite outgrowth from sensory neuronal cultures with picomolar potency with maximal effects comparable to nerve growth factor. In vivo, GPI-1046 stimulated the regeneration of lesioned sciatic nerve axons and myelin levels. In the central nervous system, GPI-1046 promoted protection and/or sprouting of serotonin-containing nerve fibers in somatosensory cortex following parachloroamphetamine treatment.” Steiner et al. (1997b) add: “GPI-1046 also induced regenerative sprouting from spared nigrostriatal dopaminergic neurons following 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine toxicity in mice or 6-hydroxydopamine (6-OHDA) toxicity in rats. The rotational abnormality in 6-OHDA treated rats was alleviated by GPI-1046.” However, some reports showed that in the MPTP lesioning model of monkeys, GPI-1046 did not exhibit neuroregenerative activity (Eberling et al., 2002).

### 9.4.2 Structure-Based Design of Nonimmunosuppressant Neuroimmunophilin Ligands

Discovery of nonimmunosuppressant neuroimmunophilin ligands is mainly based on the three-dimensional structure of FKBP12 and the interaction of FK506 and FKBP12. FKBP12 is a member of the FKBP family. It consists of 107 amino residues and complexes tightly with FK506. The overall structure of FKBP12 is a five-stranded antiparallel \( \beta \)-sheet wrapping with a right-handed twist around a short \( \alpha \)-helix. FK506 and rapamycin bind in a hydrophobic cavity between the \( \alpha \)-helix and the interior wall of the \( \beta \)-sheet that is lined with a number of highly conserved, aromatic residues (Fig. 9.5a). The entrance of the cavity consists of aromatic residues 39–46, 50–55, and
Figure 9.5  (a) Complex structure of FKBP12-FK506; the protein is displayed in cartoon model, and FK506 is shown in stick model. (b) Key interaction between FK506 and FKBP12. I56 and Y82 make two essential hydrogen bonds with FK506. (c) The ε-H of Tyr26, Phe36, and Phe99 build the specific carbonyl oxygen binding site of FK506. See color plates.
The wall of the cavity is formed by side chains of residues Tyr20, Phe46, Phe99, and Val55–Ile56. The indole ring of Trp59 makes the bottom of the cavity (Van Duyne et al., 1993).

The major interactions of FKBP12 and FK506 are hydrogen bond and hydrophobic interactions. In the complex of FK506 and FKBP12 the pipercolinyl ring binds deeply in the cavity defined by Trp59 and the side chains of Tyr26, Phe46, Val55, Ile-56, and Phe99, whereas the \( \alpha \)-dicarbonyl amide is hydrogen bonded to the Ile-56 NH group and the Tyr-82 OH group. The pyranose ring is buried in the hydrophobic pocket formed by Phe36, Asp37, Tyr82, His87, Ile90, and Ile91, and the cyclohexyl ester chain is engaged in hydrophobic interactions within a shallow groove on the surface of the FKBPs. Besides the hydrogen bond of the \( \alpha \)-dicarbonyl amide of FK506, there are other three hydrogen bonds between FK506 and FKBP12. The NH group of Ile56 is hydrogen bonded to oxygen of the carbonyl of C-1. The carbonyl group of Glu54 makes a hydrogen bond with the OH group of C-24. Also, the carbonyl group of Gln53 makes a hydrogen bond with the OH group of C-24 through a water molecule. There is another special interaction between oxygen of carbonyl of C-9 and residues of FKBP12. In Tyr26, Phe36, and Phe99 the aromatic \( \sigma \)-H makes a cavity that interacts with oxygen of the carbonyl of C-9, which makes the complex more stable (Fig. 9.5b).

According to the key interactions between FK506 and FKBP12, the structure of FK506 can be divided into four regions: the pipercolinyl region, the dicarbonyl region, the pyranosyl region, and the cyclohexylethyl region, as shown in Fig. 9.3b (Holt et al., 1994). According to the variation of each group, many novel nonimmunosuppressive neuroimmunophilin ligands were designed and synthesized.

The first simple acyclic analogues of FK506 with high affinity to FKBP12 were reported by SmithKline Beecham. Based on the four-regions model, they designed and synthesized many analogues of FK506 with a generic structure as shown in Fig. 9.6 (Holt et al., 1994). The compound, incorporating an R1 phenyl group, an R2 phenethyl group, and an R3 (dimethylethyl)methyl group, is a very high-affinity FKBP12 ligand. The crystal structure for this compound bound to FKBP12 shows that the pipercolinyl ring and dicarbonyl groups bind identically as they do in FK506. The R3 (dimethylethyl)methyl group fills the same site that the pyran ring of FK506 occupies. The phenethyl R2 group makes hydrophobic surface contacts in the site that the cyclohexyl group of FK506 occupies. The residue Phe46 makes VDW contacts with R1 phenyl group as well as VDW contacts with the ethyl group of the R3 group (Fig. 9.7a). While solved in different space groups, the structure of the protein is

![Figure 9.6](image)

**Figure 9.6** Generic structure of Holt’s FK506 analogues.
nearly identical between the complexes of FK506 and SmithKline Beecham’s compound (Fig. 9.7b).

A slight conformation change in the side chain of Phe46 of FKBP12 takes place to maximize VDW contacts with the ligand. The side chains of Asp37 are in slightly
different conformations in the two complexes. In the FK506 complex the side chains of Asp37 approach the ligand more closely to make a hydrogen bond with the hemiketal hydroxyl group. The flexible region of the protein between Ala84 and His94 is very similar in both complexes. The crystal structure means that the simplified structure of FK506 binding domain can interact with FKBP12, as FK506 does (Holt et al., 1993).

Computational methods are commonly applied for lead identification and optimization. Although there are only a few examples of completely new chemical entities discovered by such approaches, many scientists still believe in its power and adopt it to find novel neuroimmunophilin ligands. Ludi is a suggestive program for the \textit{de novo} design of ligands for proteins. Babine et al. (1995) applied this approach and found a new type of ligand of FKBP12. Using the crystal structure of the FKBP12-FK506 complex as a starting point, Ludi suggested complementary fragments for the hydrophobic pipecolinic acid binding site of FKBP12. Per Babine et al. (1995): “Among the many fragments suggested, adamantane appeared to fill this site well and seemed well suited for further elaboration. Examination of the docked ligand for the formation of potential hydrogen bonds suggested the removal of a methylene bridge with the introduction of a carbonyl group. This should allow for a hydrogen bond to the backbone amide NH of Ile-56.” Among the suggestions of potential groups to fill the FK506 pyran binding site of FKBP12, Babine et al. found that aromatic groups appended through a linker to the bridgehead position of the new ring system seemed be the best choice. Babine et al. add: “Finally, Ludi suggested a hydroxyl group in the meta position of an aromatic ring tethered by a methylene linker [Fig. 9.8]. This hydroxyl group has the potential to make favorable electrostatic contact with Asp-37. Thus, compounds of the generic structure 1 were \textit{de novo} designed . . .” Among the designed compounds, the most effective one was compound A. A 2.4-Å crystal structure was obtained for the complex of compound A and FKBP12. The complex showed that the bicyclic ring system makes VDW contact with the hydrophobic residues Val55, Ile56, Trp59, Phe46, Tyr82, and Phe99. The carbonyl group makes an H-bond to the backbone NH of Ile56. The sulfur atom fills a small hydrophobic, electropositive cavity formed by Tyr26, Phe36, and Phe99. The pyridine ring makes VDW contact with His87, Ile91, Tyr82, and Phe99.

In another paper, Burkhard et al. (1999) utilized the molecular docking computer program SANDOCK to screen small molecule three-dimensional databases for

![Diagram](https://example.com/diagram.png)

**Figure 9.8** Ludi-based ligand design pipeline. See color plates.
FKBP ligands. Based on the site point of FKBP (Fig. 9.9), many potential compounds were selected. As confirmed by fluorescence quenching tests, several compounds with micromolar $K_d$ values were identified. Interestingly, several steroids were characterized to be novel leads for FKBP12 ligands—for example, 5β Pregnan-3.20 dion ($K_d = 7 \mu\text{M}$). The biological significance of these FKBP-steroid interactions remains to be elucidated. However, so far there is no evidence that the binding of steroids to FKBPs accounts for neuronal actions of steroids.

Our laboratory has researched neuroimmunophilin ligands since 1997. With the integration of virtual library, virtual screening, and combinatorial chemistry, we designed, built, and screened several different generic structure libraries, each consisting of thousands of compounds. Compound 308 exhibits remarkable activity in many in vitro and in vivo neuroregeneration models and was chosen as the prime candidate for further evaluation. The co-crystal structure of compound 308 and FKBP12 was obtained and solved the key interaction between them (Sun et al., 2003). The complex shows that the 1,4-thiazine and sulfanilamide groups bind identically as the pipercolinyl ring and dicarbonyl groups do in FK506. The $p$-methyl-phenyl group fills the same site that the pyran ring of FK506 occupies. The 3-carbonyl group of the 1,4-thiazine group makes a hydrogen bond to the backbone amide NH of Ile56. The carbonyl of the carboxyl group makes another
hydrogen bond to the hydroxyl group of Tyr82. The 1,4-thiazine group and p-methyl-phenyl group make VDW contact with the hydrophobic residues Tyr26, Phe36, Phe46, Val55, Trp59, His87, and Ile90 (Fig. 9.10). At present, this compound is going to phase I clinical trial for therapy of stroke.

9.4.3 SAR of Nonimmunosuppressant Neuroimmunophilin Ligands

Structure–activity relationship (SAR) studies were performed for a large number of simple, analogous molecules (Fig. 9.11) in terms of their ability to inhibit FKBP12 (Armistead et al., 1995) and were reviewed recently (Hamilton and Steiner, 1997, 1998). As mentioned above, the binding domain of FK506 is consisted of four regions: the pipecolyl region, the dicarbonyl region, the pyranosyl region, and the cyclohexylethyl region. It is suggested that the minimal binding requires the central pipecolic acid ring with the ester group (Holt et al., 1994). SAR study results show that there are three key factors of neuroregenerative activity of this minimal binding structure. They are the presence of ester bond, the chirality of ester bond, and the presence of pipecolic ring. The disappearance of the ester group will lead to the decrease of activity. Changing the chirality of the ester bond from S-type to R-type will cause the activity to decrease in two degrees, while open-chain compounds have
approximately 10-fold weaker affinity for the enzyme as compared to FK506 and rapamycin (Teague and Stocks, 1993; Teague et al., 1994). These findings imply that the binding of the ester bond and the surface of FKBP12 needs a stable conformation to support it, which makes the binding more complete. The pipecolic acid, pyrrolidine, pyridine, and other 5 or 6 heterocyclic ring just satisfied with this requested conformation.

Figure 9.11  Generic structures of nonimmunosuppressive neuroimmunophilin ligands ($n = 1, 2; X = O, S; Y = O, S, NH; Z = C, SO$).

approximately 10-fold weaker affinity for the enzyme as compared to FK506 and rapamycin (Teague and Stocks, 1993; Teague et al., 1994). These findings imply that the binding of the ester bond and the surface of FKBP12 needs a stable conformation to support it, which makes the binding more complete. The pipecolic acid, pyrrolidine, pyridine, and other 5 or 6 heterocyclic ring just satisfied with this requested conformation.
The \( \alpha \)-dicarbonyl amide group was identified as being essential for enzyme inhibition because derivatives obtained by replacement of either or both of the carbonyl groups corresponding to positions C-8 and C-9 of FK506 such as peptides \( (K_i = \geq 1 \mu M) \) (Hauske et al., 1992), sulfonamides \( (K_i = \geq 160 nM) \) (Holt et al., 1994; Duffi, 1992) and ureas \( (K_i = \geq 120 nM) \) (Dragovich et al., 1996) were substantially less potent. The Abbot group examined the effect of modifications to the ketone carbonyl groups in simple acyclic compounds (Wang et al., 1994). Their results were consistent with the conclusion that ketone is the most potent group.

Although the binding properties for some of these \( \alpha \)-dicarbonyl amides have been critically discussed, the variation of pyranose ring and cyclohexylethyl moiety of the FK506 binding domain is also studied extensively. The SmithKline Beecham results showed that substitution of the pyranosyl group with a (dimethylethyl)methyl group resulted in a threefold increase in affinity for FKBP12 (Luengo et al., 1994). The Vertex group has reported that a 3,4,5-trimethoxy phenyl group is a useful pyran replacement in acyclic FKBP12 ligands (Armistead et al., 1995). Substitution of cyclohexylethyl moiety with either a phenyl group or a phenethyl group resulted in an 11-fold and 18-fold increase in FKBP12 affinity, respectively. Incorporating both phenyl groups and phenethyl groups is a very-high-affinity FKBP12 ligand (7 nm). Among these compounds, high-affinity FKBP12 ligands were bulky hydrophobic alkyl groups such as 1,1-dimethyl-propyl or (3,4,5-trimethoxyl)phenyl as substituents for the pyranose ring region of FK506, as well as simple alkyl or alkyl aryl esters, instead of the lead cyclohexylethyl moiety. This confirmed the conclusion that these moieties mainly make hydrophobic VDW contact with FKBP12.

### 9.5 CONCLUSIONS AND FUTURE PROSPECTS

Because of the complicated mechanisms of neurodegenerative disease, there are still few available clinical therapeutants for these diseases’ treatment. All available drugs for the treatment of frequently occurring neurodegenerative disorders such as Alzheimer’s and Parkinson’s diseases only alleviate symptoms and delay the neuronal atrophy by compensating or increasing impairments of the neurotransmitter metabolism (Brodaty, 1999; Grunblatt et al., 2000). Growth factors such as BDNF, GDNF, NGF, and NT3 have been shown to promote nerve regrowth \textit{in vivo} (Apfel et al., 1994; McMahon and Priestley, 1995; Gash et al., 1996) and seem be the perfect therapeutants for nervous disorders. Most of them are currently being evaluated in animal models of neurodegeneration. But the therapeutic utility of these proteins is severely limited due to their lack of oral bioavailability and inability to cross the blood–brain barrier, necessitating their delivery directly to the site of injury. Abnormal neuronal sprouting has also been observed following administration of protein growth factors (Voolf et al., 1996). Clinical potential of other available or presently examined neuroregenerative therapeutics, such as gangliosides (Ekstrom and Tomlinson, 1990) or the dihydropyridine Ca\(^{2+}\) antagonist nimodipine, is hampered by the observed adverse side effects (Candelise and Ciccone, 2000; Ahmed et al., 2000). As an alternative approach, neural transplantation of competent neuronal cells or tissues is considered to hold promise as a future
therapeutic tool to treat progressive and irreversible neural disorders (Dunnett, 1999). These neural transplantation techniques, however, are still in an early developmental stage, and many of the adherent problems such as viability and function of the graft remain to be solved. New therapeutic approaches are required to treat neural diseases.

The neuroimmunophilin ligands have been demonstrated to promote regrowth of damaged facial and sciatic nerves in the peripheral nervous system and the regeneration of damaged dopaminergic, serotonergic, and cholinergic neurons in the central nervous system. This broad scope of action suggests therapeutic utility in a variety of neurodegenerative disorders such as Parkinson’s disease, Alzheimer’s disease, diabetic and peripheral neuropathies, and spinal cord injuries. The studies surveyed above document the remarkable ability of FKBP ligands to cause functional regeneration of a variety of damaged neuronal pathways. In terms of specificity of action, bioavailability, and stability, these compounds might be essentially advantageous compared to peptidic growth factors, being presently evaluated, such as NGF, BDNF, NT3, and GDNF. Unlike the growth factors, which exert neurotrophic activities on overlapping but limited populations of CNS neurons, FKBP ligands do not cause aberrant sprouting of healthy neuronal processes in vivo (Steiner et al., 1997b).

Many unanswered questions remain regarding the actions of these compounds. The mechanism whereby the nerve regenerative effects of FK506 and small FKBP ligands are exerted is still unknown. FKBP12 is present in high concentrations in neurons, and many compounds produce neurotrophic effects in vitro in concentrations at which less than 0.1% of the enzyme would be inhibited. This implies that there may be other FKBPs, present in lower concentrations in nerve cells, that mediate the actions of these compounds. The ability of neuroimmunophilin ligands to produce potent neurotrophic effects in vitro and in vivo at low concentrations, along with the lack of correlation of these effects with inhibition of rotamase activity, suggests another possibility that interaction of the compounds with FKBP or an FKBP-like protein results in formation of an activated complex, leading to a gain of function for the FKBP. All these hypotheses need more evidence to prove.

Although there are many things such as targets and mechanism that need to be elucidated, the intensive interest and effort devoted to exploring the therapeutic utility of neuroimmunophilin ligands is continuing. It is likely that this will be a major new area of medicinal chemistry and neuroscience research for years to come.

**ACKNOWLEDGMENTS**

We would like to acknowledge Dr. Wang Lili for their contribution during the preparation of the manuscript, especially for discussion of ligand efficiency and potential neuroregenerative mechanism. We are also grateful to A. B. for administrative assistance and for overall scientific input and abbreviation preparation.
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATP</td>
<td>Adenosine triphosphate</td>
</tr>
<tr>
<td>5-HT</td>
<td>5-Hydroxytryptamine</td>
</tr>
<tr>
<td>6-OHDA</td>
<td>6-hydroxydopamine</td>
</tr>
<tr>
<td>AD</td>
<td>Alzheimer’s disease</td>
</tr>
<tr>
<td>ADNF</td>
<td>activity-dependent neurotrophic factor</td>
</tr>
<tr>
<td>AGEs</td>
<td>advanced glycation end products</td>
</tr>
<tr>
<td>AMPA</td>
<td>alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid</td>
</tr>
<tr>
<td>AP-1</td>
<td>activator protein 1</td>
</tr>
<tr>
<td>BBB</td>
<td>brain–blood barrier</td>
</tr>
<tr>
<td>BDNF</td>
<td>brain-derived neurotrophic factor</td>
</tr>
<tr>
<td>CA1</td>
<td>hippocampal domain CA1</td>
</tr>
<tr>
<td>CNS</td>
<td>central nervous system</td>
</tr>
<tr>
<td>CsA</td>
<td>Cyclosporine A</td>
</tr>
<tr>
<td>DA</td>
<td>dopamine</td>
</tr>
<tr>
<td>EGFR</td>
<td>epidermal growth factor receptor</td>
</tr>
<tr>
<td>FK506</td>
<td>an immunosuppressant named Tacrolimus</td>
</tr>
<tr>
<td>FKBP12</td>
<td>FK506 binding protein with molecular weight 12 kDa</td>
</tr>
<tr>
<td>FKBP12.6</td>
<td>FK506 binding protein with molecular weight 12.6 kDa</td>
</tr>
<tr>
<td>FKBP13</td>
<td>FK506 binding protein with molecular weight 13 kDa</td>
</tr>
<tr>
<td>FKBP25</td>
<td>FK506 binding protein with molecular weight 25 kDa</td>
</tr>
<tr>
<td>FKBP37</td>
<td>FK506 binding protein with molecular weight 37 kDa</td>
</tr>
<tr>
<td>FKBP51</td>
<td>FK506 binding protein with molecular weight 51 kDa</td>
</tr>
<tr>
<td>FKBP52</td>
<td>FK506 binding protein with molecular weight 52 kDa</td>
</tr>
<tr>
<td>FKBPs</td>
<td>FK506 binding protein family</td>
</tr>
<tr>
<td>GAP-43</td>
<td>growth-associated protein 43</td>
</tr>
<tr>
<td>GDNF</td>
<td>glial cell-line-derived neurotrophic factor</td>
</tr>
<tr>
<td>GTP</td>
<td>guanosine triphosphate</td>
</tr>
<tr>
<td>HD</td>
<td>Huntington’s disease</td>
</tr>
<tr>
<td>HNE</td>
<td>human neutrophil elastase</td>
</tr>
<tr>
<td>Hsp-90</td>
<td>heat shock protein 90</td>
</tr>
<tr>
<td>LN1</td>
<td>lupus nephritis protein</td>
</tr>
<tr>
<td>MAP</td>
<td>a 42-kDa serine/threonine protein kinase</td>
</tr>
<tr>
<td>MAP kinase/ERK2</td>
<td>N-methyl-4-phenylpyridinium</td>
</tr>
<tr>
<td>MPP</td>
<td>1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine</td>
</tr>
<tr>
<td>MPTP</td>
<td>nerve growth factor</td>
</tr>
<tr>
<td>NGF</td>
<td>a protein in the neural retina of <em>Drosophila</em></td>
</tr>
<tr>
<td>NMDA</td>
<td>N-methyl-d-aspartic acid</td>
</tr>
<tr>
<td>NOS</td>
<td>nitric oxidase synthase</td>
</tr>
<tr>
<td>NSAIDs</td>
<td>nonsteroidal anti-inflammatory drugs</td>
</tr>
<tr>
<td>PAHX</td>
<td>phytanoyl-CoA-α-hydroxylase</td>
</tr>
<tr>
<td>PC12</td>
<td>rat adrenal pheochromocytoma cell line</td>
</tr>
<tr>
<td>PD</td>
<td>Parkinson’s disease</td>
</tr>
<tr>
<td>PGA1</td>
<td>prostaglandin A1</td>
</tr>
<tr>
<td>PGE1</td>
<td>prostaglandin E1</td>
</tr>
<tr>
<td>PGs</td>
<td>prostaglandins</td>
</tr>
<tr>
<td>PPIases</td>
<td>peptidylprolyl cis–trans isomerases</td>
</tr>
</tbody>
</table>
REFERENCES


REFERENCES


Luengo et al. (1994).


10.1 INTRODUCTION

Since Arthur Cushny (Challener, 2001a) established 80 years ago the relevance of chirality to the pharmaceutical industry by showing that one enantiomer of hyoscyamine possessed greater pharmacological activity than the other, drug chirality has become more and more important in both pharmaceutical industries and the academic world over the past three decades. Drug chirality is now a major theme in the design, discovery, development, launching, and marketing of new drugs with the development of stereochemistry, pharmacology, and other subjects. The importance of drug chirality is not only a major challenge to the minds of practicing scientists, but also a highly fertile field for the development of technologies for the production of high-value pharmaceuticals. New developments in technologies for isolating, preparing, and purifying chiral materials have greatly increased the opportunities for utilizing enantiopure compounds in commercial applications. Novel techniques for classical resolution, new methodologies for developing selective enzymes for biocatalysis, advances in the application of microorganisms for chemical production, and continued progress in the area of asymmetric synthesis have all contributed to the growth of this field.

* Corresponding author: lingq@mail.sioc.ac.cn
10.2 SIGNIFICANCE OF CHIRALITY

10.2.1 Terminology in Stereochemistry

10.2.1.1 Superimposability

The word “chiral” is derived from the Greek word “cheir,” meaning hand, or “hand-edness” in a general sense. The left and right hands are mirror images of each other, and no matter how the two are arranged, one cannot be placed directly over the top of the other in the exact same orientation. Chiral objects are commonly known as a person’s right and left hands, snail shells, and clockwise- or counterclockwise-threaded screws. Simply speaking, this is a unique character of nature. Accordingly, a chiral molecule is a molecule that is not superimposable on its mirror image, whereas an achiral compound has a superimposable mirror image (Müller and Kohler, 2004). Nowadays, chiral molecules have ever-increasing importance in the pharmaceutical area in both academy and industry settings. Chiral drugs refer to those molecules that belong to chiral molecules where the chirality exists. It is not surprising that the synthesis of chiral compounds has become an important subject for research. At the moment, most new drugs and those under development consist of a single optically active isomer.

10.2.1.2 Stereoisomerism and Enantiomers

Chirality is a fundamental property of many three-dimensional objects—in other words, a phenomenon caused by a process called stereoisomerism. Stereoisomers are compounds that have the same atoms connected in the same order but differ from each other in the way that the atoms are oriented in space. Chiral molecules that behave like the image and mirror image of each other and are not superimposable are called enantiomers. To take a simple example, lactic acid can be obtained in two forms, 1 and 2, which are clearly related as mirror images to each other (Fig. 10.1).

10.2.1.3 Absolute Configuration

It is very important to define the absolute configuration of a chiral molecule in order to understand its function in a biosystem. Many biological activities are exclusive to one specific absolute configuration. Without a good understanding of absolute configuration

![Figure 10.1 Mirror images of lactic acid.](image-url)
of a molecule, we often cannot understand its chemical and biological behavior. Under
normal conditions, the two enantiomers of a chiral compound will have identical
chemical and physical properties such as the same boiling/melting point and the
same solubility in normal achiral solvent (Eliel and Wilen, 1994). Also, their chemical
reaction will be identical under achiral conditions. Enantiomers do differ from each
other in the direction in which they rotate the plane-polarized light, and this is called
optical activity or optical rotation. Thus, 2, which rotates the plane-polarized light in a
clockwise direction, is denoted as (+)-lactic acid by a plus sign (+) prior to the name of
the compound, while the enantiomer (1), which has an equal but opposite rotation under
the same conditions, is denoted as (−)-lactic acid by a minus sign (−) prior to the name
of the compound. However, under chiral conditions, these enantiomers may behave in a
very different way. For example, the physical property or chemical reactivity may
change significantly under chiral conditions. Determination of absolute configuration of
a chiral center is accomplished by assigning the spatial orientation of a molecule and
correlating this orientation to the negative or positive sign of the rotation of polarized
light by the substance under a given condition (Lin et al., 2001).

10.2.1.4 Diastereoisomers

Many molecules contain more than one chiral center. If a compound has two chiral
carbon atoms, there are a total of four possible stereoisomers. Two of the isomers will
be mirror images of each other, and therefore enantiomers. The other two isomers are
called diastereoisomers—that is, stereoisomers that are not enantiomers (Challener,
2001a). The relationship for 2-chloro-3-hydroxybutane is exemplified by Fig. 10.2. In
such conditions, there exist possibly two pairs of enantiomers and two pairs of
diastereoisomers, for four individual compounds.

10.2.1.5 Determining Enantiomer Composition (ee)

Measuring the enantiomer composition is very important in the chemistry of chiral
drugs, because people working in this area need this information to evaluate the
biological data to make sure they are viable. The enantiomer composition of the
sample is described by the enantiomer excess (ee), which describes the excess of one
enantiomer over the other, where (R) and (S) are the composition of R and S
enantiomers, respectively. In a similar manner, the diastereoisomer composition of
a sample can be described by the diastereoisomer excess (de), which refers to the

![Figure 10.2](image-url) Enantiomers and diastereomers.
excess of one diastereoisomer over the other. A variety of methods for determination of ee are available (see Eliel and Wilen, 1994).

10.2.2 Chirality in Drug Development

10.2.2.1 Chirality in Living Systems

Louis Pasteur, over 100 years ago, might have been the first scientist to propose the idea of chemical “handedness.” His initial theories arose from his study of racemic tartaric acid, which is composed of visibly discernible crystalline forms of enantiomeric D- and L-tartaric acid. From this, the concepts of asymmetry and chirality have been arisen (Hillier and Reider, 2002).

One of the main features of the living world is its chirality, because most of the biological macromolecules of living systems occur in nature in one enantiomer form only (Mislow, 1999). An intrinsic property of the “building blocks of life,” such as the amino acid and carbohydrate building blocks of proteins, carbohydrates, and nucleic acids, is that they are all composed of chiral molecules (Maier et al., 2001). Moreover, they are homo-chiral: The 21 essential amino acids are all L-enantiomers, whereas most carbohydrates have the D-configuration (Bonner, 2000). The structures of the L- and D-configurations of alanines 3 and 4 are shown in Fig. 10.3.

10.2.2.2 Stereoselective Actions of Chiral Drugs

Almost 80 years ago Arthur Cushny (Challener, 2001a) first found a difference in pharmacological action between optical isomers, or enantiomers, of a chiral molecule. He showed that the natural, levorotatory alkaloid hyoscyamine was twice as potent as atropine, which is racemic, in antagonizing cholinergic stimuli. A few years later, Easson and Stedman postulated a three-point interaction between a drug and its receptor to explain stereoselectivity in drug action (Waldeck, 2003).

Metabolizing enzymes, protein-binding sites, and/or drug receptor sites have three-dimensional structures; therefore, if the drug molecule is chiral, thus forming optically active isomers, this interaction is usually stereoselective. This means that with the chiral human enzymes and cell surface receptors, the two enantiomers of a racemic drug may be absorbed, activated, or degraded in very different ways, both

\[
ee = \left( \frac{(S) - (R)}{(S) + (R)} \right) \times 100\
\]

Figure 10.3  D- and L-configurations of alanine.
The two enantiomers may have unequal degrees or different kinds of activity. One may be therapeutically effective, while the other may be ineffective or even toxic. This is illustrated by the following examples.

Propafenone (PPF) (5) is an anti-arrhythmic drug that is clinically used as a racemic mixture. The enantiomers have almost the same activity in their sodium channel-blocking activity, but (S)-PPF is 100 times more potent at the β-adrenergic receptors (Pires de Abreu et al., 1999). Amphetamine (6) and its analogues are substances that have a stimulating effect on the central nervous system (CNS). d-Amphetamine is 3–4 times more effective in central nervous system stimulation than its isomer, whereas the latter is slightly more potent in its cardiovascular action (Varesio and Veuthey, 1995). β-Adrenergic block agents exhibit stereoselective actions at β-adrenoceptors, whereby (−)-(S)-propranolol (7) is about 40 times more potent than the enantiomer. Nevertheless, for local anesthetic and antiarrhythmic activities, the enantiomers are essentially equipotent (Triggle, 1997). Ephedrine and its alkaloids are used for the treatment of asthma, nasal congestion, and obesity. There are two chiral centers in the molecule, and therefore there exist four possible isomers: (+)-ephedrine (8a), (−)-ephedrine (8b), (+)-pseudoephedrine (9a), and (−)-pseudoephedrine (9b). (−)-Ephedrine and (−)-pseudoephedrine are isomers occurring in nature. The ephedrine isomers exhibit direct and indirect effects on both α- and β-adrenergic receptors (ARs). (−)-Ephedrine (8b) has been shown to be the most potent of the four isomers with various behavioral measures (Young and Glennon, 1998). In addition to sympathomimetic effects, it has been reported that ephedrines can exert CNS dopaminergic effects similar to those of the amphetamines. Systemic administration of (−)-ephedrine in rats increased extracellular levels of dopamine (DA) in the nucleus accumbens, decreased feeding, and increased locomotor activity (Wellman et al., 1998).

In order to further investigate the mechanism of action for the ephedrine isomers in vivo, the direct effects of the four-ephedrine isomers on human β1-, β2-, and β3-AR subtypes expressed in CHO cells were examined by Vansal and Feller. The
results are listed in Table 10.1. Their studies demonstrated that the potencies of the four ephedrine isomers on the human β-AR subtypes are different. (−)-Ephedrine was the most potent of the four ephedrine isomers on all three human β-AR subtypes and was the only ephedrine isomer to possess some agonist activity on the human β3-AR (Vansal and Feller, 1999).

Today, enantioselectivity is well accepted in molecular and atomic detail. Through site-directed mutagenesis, the amino acids responsible for the stereoselective recognition of ligand molecules have been identified in a number of cases, in particular for the β-adrenergic receptor, a prototypic G-protein-coupled receptor. Novel ligands with one or more chiral centers and complicated stereochemistry continue to be developed as potential medicines, despite reservations put forward by regulatory agencies (Soudijn et al., 2004).

The HIV-1 proteinase has been recognized as an attractive target for antiviral therapy. The structure of the naturally occurring L-enzyme, the HIV-1 proteinase with 99 amino acids, had been determined earlier (Miller et al., 1989). Later, the D-enantiomer was synthesized by chemical methods (Milton et al., 1992). Each enantiomer is equally active on hexapeptide substrate analogues; however, only the D-enzyme only being active on D-peptides and the L-enzyme on L-peptides (Lamzin, 1995). Therefore, HIV-1 protease in its D- and L-forms exhibits opposing chiral substrate selectivity (Milton, 1992). Crixivan (10) is a protease inhibitor for AIDS, a single enantiomer in pure form. The X-ray structure at 1.9-Å resolution showed a stick model of Crixivan (10) bound to the active site of HIV-1 protease (Zurer, 2005), as shown in Fig. 10.4.

### TABLE 10.1 Cyclic AMP Effects of Ephedrine Isomers on Human β-Adrenoceptor Subtypes Expressed in CHO Cells

<table>
<thead>
<tr>
<th>Compound</th>
<th>(β₁-Adrenoceptors&lt;sup&gt;a&lt;/sup&gt;)</th>
<th>(β₂-Adrenoceptors&lt;sup&gt;a&lt;/sup&gt;)</th>
<th>(β₃-Adrenoceptors&lt;sup&gt;a&lt;/sup&gt;)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>pK&lt;sub&gt;act&lt;/sub&gt; (a)</td>
<td>I.A. (%)</td>
<td>R.P.</td>
</tr>
<tr>
<td>(−)-Isoproterenol</td>
<td>8.13 ± 0.10&lt;sup&gt;b&lt;/sup&gt;</td>
<td>100</td>
<td>1</td>
</tr>
<tr>
<td>(−)-Ephedrine</td>
<td>4.14 ± 0.07&lt;sup&gt;c&lt;/sup&gt;</td>
<td>66</td>
<td>9,773</td>
</tr>
<tr>
<td>(+)-Ephedrine</td>
<td>6.26 ± 0.12&lt;sup&gt;c&lt;/sup&gt;</td>
<td>68</td>
<td>74</td>
</tr>
<tr>
<td>(+)-Pseudoephedrine</td>
<td>3.51 ± 0.05&lt;sup&gt;c&lt;/sup&gt;</td>
<td>53</td>
<td>41,688</td>
</tr>
<tr>
<td>(−)-Pseudoephedrine</td>
<td>2.95 ± 0.19&lt;sup&gt;c&lt;/sup&gt;</td>
<td>53</td>
<td>151,358</td>
</tr>
</tbody>
</table>

<sup>a</sup>Data for I.A. (intrinsic activity) are expressed as percentage responses to that of (−)-isoproterenol (3 × 10<sup>−5</sup> M = 100%). pK<sub>act</sub> values are expressed as means ± SEM.R.P. = relative potency, calculated as EC<sub>50</sub> of drug/EC<sub>50</sub> of (−)-isoproterenol.  
<sup>b</sup>ND, not determined due to low intrinsic activity.  
<sup>c</sup>Source: Data were taken from Vansal and Feller (1999).
10.2.3 Pharmacological Implications of Chiral Drugs

Molecular chirality is a fundamental consideration in drug discovery (Zhang et al., 2005). In the asymmetric environment of receptors and enzymes, two enantiomers of chiral drugs may form different spatial relationships; therefore, there may be significant differences in their pharmacodynamic activity and their pharmacokinetic properties.

10.2.3.1 Pharmacodynamic Activity

Differences in the potency of stereoisomers may be related to differences in their affinity or intrinsic activity at receptor sites (Rentsch, 2002). The effects of differently acting stereoisomeric drugs can be categorized basically as follows: (a) Two enantiomers may have nearly identical qualitative and quantitative pharmacological activity. For example, carvedilol (11) is a newer agent that is marketed as the racemate for the treatment of hypertension and congestive heart failure (Tenero et al., 2000). Carvedilol is a nonselective β- and α-adrenergic receptor blocking agent, and the nonselective β-blocking activity resides mainly in the (S)-carvedilol, while the α-blocking effect is shared by (R)- and (S)-enantiomers (Bartsch et al., 1990). Sotalol (12) is a chiral β-adrenergic blocking drug used in therapy as a racemic mixture. The (R)-enantiomer possesses the majority of the β-blocking activity; however, the (R)- and (S)-enantiomers of sotalol share an equivalent degree of III antiarrhythmic potency (Mehvar and Brocks, 2001). (b) The two enantiomers of a chiral drug may have similar modes of action but may differ in their affinity to a receptor or an enzyme, resulting in different reaction rates. Under these conditions, enantiomers may have opposite or different effects—for example, isoprenaline (13), where the (−)-enantiomer is a β1-adrenoceptor agonist and the (+)-enantiomer has antagonistic effects (Lopez-Vidriero et al., 1985). Picenadol (14) is a unique opioid mixed agonist–antagonist analgesic currently under clinical evaluation. The (+)-isomer of picenadol is a potent opiate agonist, whereas the (−)-isomer is an opioid antagonist (Zimmerman et al., 1985). (c) All the pharmacological activity may reside in one enantiomer; in this case the other enantiomer may be regarded as an impurity.
The impurity may be inactive or have desirable or undesirable activity. The maleate of (S)-(−)-Timolol \(^{15}\) is a potent and nonselective β-adrenergic blocking agent. In comparison with timolol, the maleate of (R)-(+)--enantiomer \(^{16}\) is 4 times less potent in reducing intraocular pressure in human subjects, 49 times less potent as a β\(_1\)- and β\(_2\)-adrenoceptor or antagonist in animal subjects, 13 times less potent as a bronchoconstrictor in normal subjects, and considerably less effective in reducing the heart rate of exercising human subjects \(\text{Hanna and Lau-Cam, 1995}\). Ofloxacin \(^{17}\) is a fluoroquinolone compound being able to be concentrated in phagocytic cells \(\text{Garcı́a et al., 2000}\). (S)-(−)-Ofloxacin and (R)-(−)-ofloxacin are optically active isomers of racemic ofloxacin. However, (S)-(−)-ofloxacin is considered to be the more active isomer and is 8–128 times more potent than (R)-(−)-ofloxacin with the different bacterial strains \textit{in vitro} \(\text{Zeng et al., 1999}\). As a single enantiomer, ofloxacin has been used in a clinical medicine called levofloxacin. (d) The two enantiomers of a chiral drug may act on different receptors or enzymes, resulting in different pharmacological activity. The (R)- and (S)-isomers of verapamil \(^{18}\) have different magnitudes of effect on hemodynamics and the atrioventricular conduction system. (S)-Isomer is 6–10 times more potent than (R)-isomer in regard to the Ca\(^{2+}\) channel blocking effect \(\text{Sakuta and Okamoto, 1994}\). However, both enantiomers are equally effective in increasing cellular accumulation of anticancer drugs. Verapamil may be a useful candidate drug for the treatment of multidrug resistance in cancer patients \(\text{Höll et al., 1992}\). (e) There are many chiral drugs for which one or both enantiomers have the desired effect and only one enantiomer causes unwanted side effects. A well-known disaster is the contergan tragedy. Contergan was a sedative that contained racemic thalidomide \(^{19}\). Both enantiomers had the desired therapeutic effects, whereas only the (S)-enantiomer had teratogenic effects and caused severe malformations of human babies \(\text{Man et al., 2003}\). Racemic dropropizine \(^{20}\) has long been used in human therapy as an antitussive agent. Recent studies have revealed that (S)-dropropizine possesses the same antitussive activity as the racemic mixture but has much lower selective activity on the CNS \(\text{Salunkhe and Nair, 2001}\). Therefore, particular clinical significance is attached to drugs of which one enantiomer may contribute to the side effects or toxic effects. (f) In contrast to the above-mentioned enantiomers, the inactive enantiomer might antagonize the side effects of the active isomer. In such cases, an enantiomerically pure compound is not preferred. Tramadol \(^{21}\) is a centrally acting analgesic that is not associated with the classical side effects of opiate drugs, such as respiratory depression, constipation, or sedation \(\text{Sacerdote et al., 1999}\). This compound is manufactured as an equal ratio of two enantiomers, each one displaying different affinities for various types of receptors. The (+)-enantiomer is a selective agonist for μ receptors with preferential inhibition to serotonin reuptake and enhances serotonin efflux in the brain, whereas the (−)-enantiomer mainly inhibits noradrenaline reuptake. The incidence of side effects, particularly opioid-mediated ones, was higher with (+)-enantiomer than with (±)-tramadol or (−)-enantiomer. In regard to efficacy and side effects, the racemate of tramadol is superior to the administration of either enantiomer alone for the treatment of severe postoperative pain \(\text{Rojas-Corrales et al., 1998}\). The (−)-enantiomer of dobutamine \(^{22}\), which possesses mainly α\(_1\)-adrenoceptor agonist
activity, is in contrast to (+)-dobutamine, which possesses predominantly β₁- and β₂-adrenoceptor agonist activity. The effects of (±)-dobutamine on cardiac performance are mediated, which is due mainly to an increase in stroke volume, with heart rate being only minimally affected (Ruffolo and Messick, 1985).
10.2.3.2 Pharmacokinetic Properties

The administration of racemic mixtures often results in far more complex pharmacokinetics than is the case when single enantiomers are used. Different enantiomers may have different plasma concentration time profiles as a result of chiral discrimination in pharmacokinetic processes. These differences may be further influenced by other factors such as the route of administration, the age and sex of the subject, disease states, and genetic polymorphism in cytochrome P450 (CYP) isoenzymes involved in drug metabolism (Lane and Baker, 1999). Thus, it is necessary to measure the stereoselectivity of pharmacokinetic processes (absorption, protein binding, metabolism, transport, and excretion) and to determine their contribution to the observed overall stereoselectivity of drug action.

Methadone (23) was introduced to manage opioid dependence in 1965. Methadone’s therapeutic benefits reside in the (R)-enantiomer. One reason is that methadone’s (R)-enantiomer shows 10-fold higher affinity for μ and κ opioid receptors and up to 50 times the anti-nociceptive activity in clinical studies and animal models compared with the (S)-enantiomer. Another important reason is that methadone’s enantiomers show markedly different pharmacokinetics. (R)-Enantiomer shows a significantly greater unbound fraction and total renal clearance than (S)-enantiomer. This reflects (S)-enantiomer’s higher affinity for the main plasma-binding protein. These stereoisomeric differences contribute to the wide variations in some of methadone’s pharmacokinetic parameters between patients (Foster et al., 2000; Baumann et al., 2002).

The potent histamine H1-receptor antagonist cetirizine (Zyrtec) is a racemic mixture of (R)-levocetirizine (24) (now available under the trademark Xyzal) and (S)-dextrocetirizine (25). In binding assays, levocetirizine has demonstrated a two-fold higher affinity for the human H1-receptor compared to cetirizine and an approximately 30-fold higher affinity than dextrocetirizine (Gillard et al., 2002). However, the pharmacokinetic parameters indicate that levocetirizine is rapidly and extensively absorbed and poorly metabolized, compared after administration alone or in the racemate. Its apparent volume of distribution is smaller than that of dextrocetirizine (0.41 L kg\(^{-1}\) vs. 0.60 L kg\(^{-1}\)). Moreover, the nonrenal (mostly hepatic) clearance of levocetirizine is also significantly lower than that of dextrocetirizine (11.8 mL min\(^{-1}\) vs. 29.2 mL min\(^{-1}\)). All evidence available indicates that levocetirizine is intrinsically more active and more efficacious than dextrocetirizine and for a longer duration (Tillementa et al., 2003).

The main metabolites of trimipramine (26) are desmethyl-trimipramine (27), 2-hydroxy-trimipramine (28), and 2-hydroxy-desmethyl-trimipramine (29) (as shown in Fig. 10.5). However, trimipramine appears to show stereoselective metabolism with
preferential \(N\)-demethylation of \(\alpha\)-trimipramine and preferential hydroxylation of \(\beta\)-trimipramine, mediated by CYP2-D6. CYP2C19 appears to be involved in demethylation and favors the \(\alpha\)-enantiomer, while CYP3A4 and CYP3A5 seem to metabolize \(\beta\)-trimipramine to a currently undetermined metabolite. Direct investigations of the implications of this stereoselective differences are needed (Eap et al., 2000).

10.2.3.3 Drivers for Chiral Drugs Market

For about the past two decades, the pharmacopoeia was dominated by racemates, but since the emergence of new technologies in the 1980s that allowed the preparation of pure enantiomers in significant quantities, the awareness and interest in the stereochemistry of drug action has increased (Eichelbaum and Gross, 1996; Caldwell, 1999).

Drug chirality became a major theme in the design, discovery, development, launching, and marketing of new drugs after the importance of stereoselective pharmacodynamics and pharmacokinetics was widely recognized (Agranat et al., 2002). It further became evident that the use of stereochemically pure drugs could be advantageous, because they would be expected to reduce the total dose given, simplify the dose–response relationship, remove a source of intersubject variability, and minimize toxicity due to the inactive stereoisomer and enhanced therapeutic window (Caldwell, 1999). These factors have led to an increasing preference for single enantiomers in both industry and regulatory authorities.

10.2.3.4 Market Information

Economic interests are obvious and essential driving forces in the spectacular development of new chiral substances and technological improvements (Maier et al., 2001). Today, approximately 80% of chiral intermediates and related products go into the pharmaceuticals market, and it is expected that, in the future, pharmaceuticals will remain the key driver for chiral chemistry development. In 1999, worldwide single-isomer drug sales reached $115 billion, up 16% from $99 billion in 1998, according to Richard L. DiCicco, president of the consulting firm Technology Catalysts International Corp (TCI). Single-isomer drugs accounted for 30% of the $335 billion total drug sales worldwide in 1998, edging up to 32% of the $360 billion market in 1999 (Stinson, 2000). These data are shown in Table 10.2. Worldwide sales of chiral drugs in single-enantiomer dosage forms continued
growing at a more than 13% annual rate to $133 billion in 2000, according to TCI in Falls Church, Virginia. At a future growth rate estimated by TCI, the figure could hit $200 billion in 2008 (Stinson, 2001). According to the market research firm Freedonia Group (Rouhi, 2002), demand for chiral raw materials, intermediates, and active ingredients would grow by 9.4% annually between 2000 and 2005.

An overall look at the 20-year period 1983–2002 (Bristol, 1984–2003) indicates that single enantiomers surpassed achirals, whereas racemates represented the minority category at 23% of worldwide approved drugs. However, even in the 1980s, single enantiomers were not a minor component of approved drugs and of approved chiral drugs. In subsequent years the percentage of achiral drugs decreased gradually from 43% to 34% of all approved drugs. The achirals:chirals ratios were 43:57 (1983–1986), 36:64 (1987–1990), 34:66 (1991–1994), 41:59 (1995–1998), and 34:66 (1999–2002).

The distribution of FDA-approved NMEs (new molecular entities) according to chirality character covers only 12 years (1991–2002). The distribution of FDA-approved drugs for the whole 12-year period bears some similarity to the worldwide distribution, in that 44% were single enantiomers, 42% were achiral drugs, and a minority (14%) were racemic drugs. The distribution of the 15 FDA-approved drugs in the period of January–August 2003 was 64% from single enantiomers, 14% from racemates, and 22% from achirals (Caner, 2004). All the data are shown in Table 10.3.

The chiral technology market is also evolving due to regulatory changes and the demand for single isomers. The U.S. market for chiral technology reached $1.2 billion in 2003, comprising over 40% of the total worldwide market. Rising at an average annual growth rate of 8.8%, the U.S. market is expected to exceed $1.8 billion in 2008.

The chiral industry may be divided into two main categories: the manufacturing of chiral compounds and the analysis of chiral compounds. Chiral

### TABLE 10.2 Chiral Drug Sales Hurtle Past $100 Billion and Show No Sign of Slowing

<table>
<thead>
<tr>
<th></th>
<th>1998</th>
<th>1999</th>
<th>2003</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cardiovascular</td>
<td>21,906</td>
<td>24,805</td>
<td>26,012</td>
</tr>
<tr>
<td>Antibiotics/antifungals</td>
<td>19,756</td>
<td>20,907</td>
<td>23,265</td>
</tr>
<tr>
<td>Hormones/endocrinology</td>
<td>12,297</td>
<td>13,760</td>
<td>17,345</td>
</tr>
<tr>
<td>Cancer</td>
<td>8,006</td>
<td>9,420</td>
<td>13,360</td>
</tr>
<tr>
<td>Central nervous system</td>
<td>7,027</td>
<td>8,592</td>
<td>13,720</td>
</tr>
<tr>
<td>Hematology</td>
<td>6,730</td>
<td>8,580</td>
<td>11,445</td>
</tr>
<tr>
<td>Antiviral</td>
<td>6,131</td>
<td>7,540</td>
<td>13,446</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1998</th>
<th>1999</th>
<th>2003</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respiratory</td>
<td>4,305</td>
<td>5,087</td>
<td>8,795</td>
</tr>
<tr>
<td>Gastrointestinal</td>
<td>1,718</td>
<td>2,998</td>
<td>5,355</td>
</tr>
<tr>
<td>Ophthalmic</td>
<td>1,482</td>
<td>1,794</td>
<td>2,070</td>
</tr>
<tr>
<td>Dermatological</td>
<td>1,124</td>
<td>1,270</td>
<td>1,540</td>
</tr>
<tr>
<td>Analgesics</td>
<td>842</td>
<td>1,045</td>
<td>1,135</td>
</tr>
<tr>
<td>Vaccines</td>
<td>568</td>
<td>676</td>
<td>1,100</td>
</tr>
<tr>
<td>Other</td>
<td>7,947</td>
<td>8,527</td>
<td>7,425</td>
</tr>
</tbody>
</table>

**TOTAL**

<table>
<thead>
<tr>
<th></th>
<th>1998</th>
<th>1999</th>
<th>2003</th>
</tr>
</thead>
<tbody>
<tr>
<td>$ (Millions)</td>
<td><strong>99,389</strong></td>
<td><strong>115,001</strong></td>
<td><strong>146,013</strong></td>
</tr>
</tbody>
</table>

*Source: Data were taken from Stinson (2000).*

The distribution of FDA-approved NMEs (new molecular entities) according to chirality character covers only 12 years (1991–2002). The distribution of FDA-approved drugs for the whole 12-year period bears some similarity to the worldwide distribution, in that 44% were single enantiomers, 42% were achiral drugs, and a minority (14%) were racemic drugs. The distribution of the 15 FDA-approved drugs in the period of January–August 2003 was 64% from single enantiomers, 14% from racemates, and 22% from achirals (Caner, 2004). All the data are shown in Table 10.3.

The chiral technology market is also evolving due to regulatory changes and the demand for single isomers. The U.S. market for chiral technology reached $1.2 billion in 2003, comprising over 40% of the total worldwide market. Rising at an average annual growth rate of 8.8%, the U.S. market is expected to exceed $1.8 billion in 2008.

The chiral industry may be divided into two main categories: the manufacturing of chiral compounds and the analysis of chiral compounds. Chiral
manufacturing will continue to dominate the market as it rises from revenues of over $1 billion in 2003 to the estimated over $1.6 billion in 2008. The manufacturing category is further subdivided into synthesis and separation. The chiral analysis figures include the market for chiral chromatography, chiral polarimetry, chiral NMR (nuclear magnetic resonance) shift, and others. Total sales of reagents in the analysis of chiral products reached $164.6 million in 2003 and are expected to grow at an AAGR of 6.6% to $226.4 million by 2008 (see Fig. 10.6).

TABLE 10.3 Annual Distribution of Worldwide and FDA-Approved Drugs (NMEs) According to Chirality Character in the Period 1983–2002

<table>
<thead>
<tr>
<th>Year</th>
<th>World-wide</th>
<th>FDA (%)</th>
<th>World-wide</th>
<th>FDA (%)</th>
<th>World-wide</th>
<th>FDA (%)</th>
<th>World-wide</th>
<th>FDA (%)</th>
<th>World-wide</th>
<th>FDA (%)</th>
<th>World-wide</th>
<th>FDA (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1983</td>
<td>37</td>
<td>NA</td>
<td>26</td>
<td>NA</td>
<td>37</td>
<td>NA</td>
<td>16</td>
<td>17</td>
<td>45</td>
<td>29</td>
<td>39</td>
<td>54</td>
</tr>
<tr>
<td>1984</td>
<td>28</td>
<td>NA</td>
<td>26</td>
<td>NA</td>
<td>46</td>
<td>NA</td>
<td>38</td>
<td>5</td>
<td>38</td>
<td>57</td>
<td>24</td>
<td>38</td>
</tr>
<tr>
<td>1985</td>
<td>38</td>
<td>NA</td>
<td>22</td>
<td>NA</td>
<td>40</td>
<td>NA</td>
<td>21</td>
<td>37</td>
<td>46</td>
<td>30</td>
<td>33</td>
<td>33</td>
</tr>
<tr>
<td>1986</td>
<td>26</td>
<td>NA</td>
<td>26</td>
<td>NA</td>
<td>48</td>
<td>NA</td>
<td>9</td>
<td>14</td>
<td>41</td>
<td>43</td>
<td>50</td>
<td>43</td>
</tr>
<tr>
<td>1987</td>
<td>18</td>
<td>NA</td>
<td>49</td>
<td>NA</td>
<td>33</td>
<td>NA</td>
<td>24</td>
<td>8</td>
<td>30</td>
<td>38</td>
<td>46</td>
<td>54</td>
</tr>
<tr>
<td>1988</td>
<td>26</td>
<td>NA</td>
<td>39</td>
<td>NA</td>
<td>35</td>
<td>NA</td>
<td>15</td>
<td>9</td>
<td>50</td>
<td>41</td>
<td>35</td>
<td>50</td>
</tr>
<tr>
<td>1989</td>
<td>29</td>
<td>NA</td>
<td>26</td>
<td>NA</td>
<td>45</td>
<td>NA</td>
<td>13</td>
<td>4</td>
<td>52</td>
<td>46</td>
<td>35</td>
<td>50</td>
</tr>
<tr>
<td>1990</td>
<td>33</td>
<td>NA</td>
<td>35</td>
<td>NA</td>
<td>32</td>
<td>NA</td>
<td>9</td>
<td>19</td>
<td>62</td>
<td>37</td>
<td>29</td>
<td>44</td>
</tr>
<tr>
<td>1991</td>
<td>20</td>
<td>9</td>
<td>40</td>
<td>65</td>
<td>40</td>
<td>26</td>
<td>0</td>
<td>0</td>
<td>68</td>
<td>60</td>
<td>32</td>
<td>40</td>
</tr>
<tr>
<td>1992</td>
<td>21</td>
<td>33</td>
<td>44</td>
<td>42</td>
<td>35</td>
<td>25</td>
<td>6</td>
<td>0</td>
<td>55</td>
<td>53</td>
<td>39</td>
<td>47</td>
</tr>
</tbody>
</table>

*Including diastereomeric mixtures.

NA, not applicable.

Source: Data were taken from Bristol (1984–2003) and Caner et al. (2004).

Figure 10.6 Sales of chiral technology products in the United States, 2001–2003 and 2008 (estimated) (taken from www.bccresearch.com).
10.2.3.5 Regulatory Issues

Regulatory changes have a significant effect on the production of chiral chemicals, particularly those prepared for use in formulated pharmaceutical and agrochemical products. Regulatory control of chiral drugs began in the United States with the publication in 1992 of formal guidelines on the development of chiral drugs in a document entitled *Policy Statement for the Development of New Stereoisomeric Drugs* (FDA, 1992) and was followed in the European Union (EU) in 1994 by *Investigation of Chiral Active Substances*. The governing bodies have issued specific rules pertaining to the development of stereoisomeric drugs. These guidelines require that applicants must recognize the occurrence of chirality in new drugs, attempt to separate the stereoisomers, assess the contribution of the various stereoisomers to the activity of interest, and then make a rational selection of the stereoisomeric form that is proposed for marketing. These common features are the basis for the consideration of this topic under the auspices of the International Conference on Harmonisation (Caldwell, 1996).

10.3 SOURCES OF CHIRAL DRUGS

There are three main methodologies for preparing chiral molecules in homo- or enriched enantiomeric forms: The first technique involves isolation of natural abundant products or semisynthetic drugs derived from natural products, the second involves biotransformation of inexpensive, available feedstocks, and the third involves a chemical control method using optically active compounds (obtained from one of the first two methods) or prochiral starting materials. The chemical control methods of the preparation of enantiomers can be characterized in detail as shown in Fig. 10.7.

A survey by Frost and Sullivan estimated that in 2002, 55% of chiral products worldwide were generated by traditional chiral pool and separation, 35% by chemocatalysis, and 10% by biocatalysis.

10.3.1 Natural Products

Natural products or semisynthetic drugs derived from natural resources present fertile preparation of enantiopure compounds by various chemical methods and diversified field for discovering therapeutic of single enantiomers (Rouhi, 2003).

Figure 10.7 Chemical control methods of the preparation of enantiomers.
For example, (−)-quinine is an antimalarial (30), whereas its quasienantiomer, (+)-quinidine (31), is an antiarrhythmic compound. Similarly, the (−)-enantiomer of the predominant natural (+)-gossypol (32) has male-contraceptive and anticancer activities.

Isolation of natural products can be achieved in several ways. For example, many terpenes are extracted from the various plants in which they are produced. Steam distillation at atmospheric pressure is often used to extract the terpene molecules from plants. Essential oils are analyzed by gas chromatography in order to estimate the concentration of active molecule that is removed from the mixture by preparative low-pressure liquid chromatography and a polar solvent as an elutant. The chemical and optical purities of the monoterpenes are dependent upon the specific plant source, the location where it was grown, and the processing conditions. Table 10.4 lists different terpenes and plant sources they are derived from (Martin et al., 2004).

<table>
<thead>
<tr>
<th>Terpene</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>(R)-(+) Citronellol</td>
<td>33 Citronella, Eucalyptus citriodora</td>
</tr>
<tr>
<td>(S)-(−) Citronellol</td>
<td>34 Geranium, rose</td>
</tr>
<tr>
<td>(R)-(+) Citronellal</td>
<td>35 Citronella</td>
</tr>
<tr>
<td>(S)-(−) Citronellal</td>
<td>36 Eucalyptus citriodora</td>
</tr>
<tr>
<td>(R)-(+) Limonene</td>
<td>37 Orange, lemon, grapefruit, tangerine, bergamot, caraway, dill</td>
</tr>
<tr>
<td>(S)-(−) Limonene</td>
<td>38 Commercial</td>
</tr>
<tr>
<td>(S)-(+) Carvone</td>
<td>39 Caraway, dill</td>
</tr>
<tr>
<td>(R)-(−) Carvone</td>
<td>40 Mentha dulcis, Mentha spicata (spearmint)</td>
</tr>
</tbody>
</table>
10.3.2 Biotransformation (Biocatalysis)

The pharmaceutical industries are doing everything they can to develop a mass-synthesis technique that yields single enantiomer compounds, the drug intermediates, or active pharmaceutical ingredients. While enantiomerically selective organic synthesis is the traditional approach, using enzymes and enzyme-containing microorganisms to biocatalyze a reaction is becoming increasingly important. There has been an increasing awareness of the enormous potential of microorganisms and enzymes for the transformation of synthetic chemicals with high chemoselectivity, regioselectivity, stereoselectivity, and specificity (Patel, 2002). Large numbers of microorganisms or enzymes have been employed for the production of optically active compounds by means of kinetic resolution or stereospecific chemical transformations (Rasool et al., 2005).

10.3.2.1 Kinetic Resolution—Enzymatic

Resolutions generally employ isolated enzymes. They are pretty basic technologies but can be very effective (McCoy, 2001). Naproxen (41) is one of the early chiral molecules developed and used as a single enantiomer. *Trichosporon* sp. (TSL), a newly found strain isolated from a locally fermented cottage cheese, has been found to be highly stereoselective in the resolution of (S)-(+)naproxen (ee >99%, E ~ 500) from the corresponding racemic methyl ester. The viability of the resolution process has been further improved by the development of a simple racemization process for the enriched (R)-(−)-ester (Koul et al., 2003) (Scheme 10.1).

Secnidazole (42c) showed an anti-amoebic potency 10 times higher than that of metronidazole (42a) and ornidazole (42b). The latest developments of those drugs reveal that ornidazole (42b) is still a subject of research for use as antifertility agents in male animals. In order to further reveal the relationship between configuration and bioactivities of these two chiral drugs, the necessity of gaining the optically pure enantiomers of ornidazole (42b) and secnidazole (42c) is quite obvious. Our laboratory attempted to acquire enantiomerically pure ornidazole (42b) and secnidazole (42c) via a lipase-catalyzed resolution; accordingly, ornidazole (42b) and secnidazole (42c) were esterified with vinylacetate employing different lipases (Scheme 10.2), and the results are listed in Table 10.5. Among the available lipases tested, AKL exhibited the best result for the kinetic resolution of both (42b) and (42c), given the high values and isolated yields (Tian et al., 2003).
10.3.2.2 Chemo-enzymatic Synthesis

Fermentation and enzyme-based catalysis are starting to challenge traditional synthetic methods of producing optically active pharmaceutical intermediates. As these biocatalytic methods gain ground, a few organic-chemistry-based intermediate manufacturers have climbed aboard, but not many open seats remain for the rest (McCoy, 1999). Historically, the synthesis of single isomers of complicated molecules has been expensive and time-consuming, and biocatalysis was not employed because of many factors restricting its use. Improvements in genetic engineering, immobilization, and stabilization technologies have enabled the introduction of stable enzymes that meet the highly specific needs required for chemical manufacturing, such as solvent stability and the ability to function at different temperature and pH ranges (Rogers, 1999).

<table>
<thead>
<tr>
<th>Entry</th>
<th>Compound</th>
<th>Lipase</th>
<th>Time (d)</th>
<th>Conversion (%)</th>
<th>Product</th>
<th>Yield (%)(^a)</th>
<th>ee (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>rac-(42b)</td>
<td>PSL</td>
<td>20</td>
<td>14</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>2</td>
<td>rac-(42b)</td>
<td>AKL</td>
<td>18</td>
<td>44</td>
<td>((-)-(S)-(42b)/ ((+)-(R)-(42b)</td>
<td>43.0/ 95(&gt;99(^b))/94(&gt;99(^b)</td>
<td>52.2</td>
</tr>
<tr>
<td>3</td>
<td>rac-(42b)</td>
<td>CAL</td>
<td>10</td>
<td>5</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>4</td>
<td>rac-(42c)</td>
<td>AKL</td>
<td>10</td>
<td>44</td>
<td>((-)-(R)-(42c)/ ((+)-(S)-(42c)</td>
<td>37.5/ 94(&gt;99(^b))/90(&gt;99(^b)</td>
<td>51.8</td>
</tr>
</tbody>
</table>

\(^a\)Isolated yield.
\(^b\)Enantiomeric excess (ee) after crystallization.
\(^c\)Determined by chiral HPLC analysis of secnidazole acetate.
As part of our continuing work toward utilizing enzymes and enzyme-containing microorganisms to biocatalyze organic synthesis, biotransformation has become a powerful tool to provide chiral synthons of drugs or pharmaceutical intermediates, which are otherwise difficult to obtain by conventional chemical methods.

The Asymmetric Carbonyl Reduction Mediated by Geotrichum sp.

Baker’s yeast (*Saccharomyces cerevisiae*) is well known for its ability to convert various classes of compounds, especially carbonyl-containing substrates, into the corresponding optically active alcohols, usually with (S)-configuration. For the synthesis of various kinds of natural products and physiologically active compounds, it is desirable to find a microorganism that is capable of performing reduction to provide products with the opposite enantioselectivity, a complement to baker’s yeast. For this purpose, we carried out a number of screenings on soil samples and found a strain of fungus, *Geotrichum* sp., called G.38, which shows the desired properties (Chen, 2003) when the carbonyl group is reduced to alcohol with anti-Prelog stereogenic outcome (Scheme 10.3).

Synthesis of (−)-Pyrenophorin α, β-Unsaturated ketone (44) could be converted to the major product (45) with >98% ee and 86% chemical yield by G.38. 45 is a precursor of 7-(S)-hydroxy-4,4-ethylenedioxy-2-octenoic acid (46), which in turn is a building block for the synthesis of (−)-pyrenophorin (47) and its analogues, a kind of naturally occurring antifungal macrodiolide (Scheme 10.4).
Synthesis of (R)- and (S)-Denopamine

(R)-Denopamine is the active form of denopamine, a new selective β₁-agonist that is important for the treatment of congestive heart failure without promotion of increased myocardial oxygen consumption or heart rate (Inamasu et al., 1987). There are few examples where enantioselective synthesis of denopamine has been achieved by using the CBS (boron-oxazolidin reaction) enantioselective catalytic reduction process (Corey and Link, 1991). The demand for finding more-convenient and higher-efficiency methods of synthesis remains. The microbial transformation of ketones with G.38 and its application to the syntheses of the key intermediates of (R)-denopamine were performed (Gu et al., 1993, 1995). Ketones 48 and 49 were incubated with G.38 to produce the corresponding alcohols 50 and 51 with (S)- and (R)-configurations, respectively (Scheme 10.5). It is necessary for protecting the phenolic hydroxyl group of ketones to obtain a high chemical and optical yield. In the case of unprotected ketones, poor ees (30% for R = CH₂Cl, 39% for R = CO₂Et) were observed. Coupling of 50 with 3,4-dimethoxyphenyl-ethylamine afforded (S)-denopamine (52). On the other hand, 51 reacted with 3,4-dimethoxyphenyl-ethylamine and was then reduced by diborane, giving the (R)-enantiomer 53 (Scheme 10.6).

Synthesis of (R)- and (S)-Fluoxetine

Fluoxetine is an important antidepressant drug for the treatment of unipolar mental depression. The enantiomers of fluoxetine exhibit their selective 5-hydroxy-tryptamine (5-HT) uptake inhibition with about equal potencies (Piperaki et al., 1995). However, many differences of pharmacological properties, as well as that of NE, have been reported in the literature.
regarding their necessitating the enantioselective synthesis of these compounds. Chemical enantiomeric routes to fluoxetine have been described by Sharpless (Gao and Sharpless, 1988), using selective reduction of 2,3-epoxycinnamoyl alcohol with Red-Al, and by Corey and Reichard, (1989), comprising the CBS enantiomeric catalytic reduction process, respectively. With the aid of microorganisms, we have accomplished the synthesis of both enantiomers of fluoxetine in good yields with high optical purities (Gu et al., 1992). Ethyl benzoyl acetate was fermented with baker’s yeast or Geotrichum sp. to give (S)- and (R)-alcohol, respectively, which provided the target compounds (S)-fluoxetine and (R)-fluoxetine through four steps of transformation as shown in Scheme 10.7. In the sense of overall yield, these processes provide the highest yield, to the best of our knowledge.

**Asymmetric Hydrocyanation Mediated by (R)-Oxynitrilase or (S)-Oxynitrilase** The (R)-oxynitrilase (HNL, hydroxynitrile lyase, EC.4.1.2.10), which catalyzes the reversible condensation of hydrogen cyanide with aldehydes, is a useful and promising enzyme for biotransformation. The resulting optically active cyanohydrins are expedient starting materials for the preparation of several important classes of compounds such as α-hydroxyketones, α-hydroxyc acids, and β-aminoalcohols as well as amino-nitriles and aziridines. It has so far been reported that the (R)-oxynitrilase catalyzed reactions were carried out with purified enzymes or crude enzyme preparations from almonds, flax, plums, cherries, apricots, and so on (Lin et al., 1999). In general, micro-aqueous buffer media were chosen as favorable reaction conditions (Scheme 10.8).

![Scheme 10.7](image)

![Scheme 10.8](image)
We (Han et al., 1998) have disclosed that it was not a problem over a wide range of temperatures when the reactions were carried out under the micro-aqueous conditions. In that case the organic phase served as a big reservoir of substrates and products, where the catalytic enzyme power that retained essential water spread homogeneously and acted as a highly effective mode of transfer. Higher yields and enantiomeric selectivities were achieved as the nonenzymatic addition and decomposition were almost suppressed. To avoid using highly toxic potassium or sodium cyanide, acetone cyanohydrin was used as a cyano donor. In the last few years, a representative example of the utility of this reaction in the synthesis of pharmaceuticals was the preparation of adrenergic bronchodilators (\(R\))-terbutaline (Effenberger and Jäger, 1997). With regard to the synthesis of (\(R\))-terbutaline (63), Effenberger and his co-workers found that the protected benzaldehydes (57) are the most suitable substrates. The corresponding (\(R\))-cyanohydrins (58) can be obtained with high conversion and excellent optical yields (\(>98\% ee\)) (Scheme 10.9).

Stereoselective syntheses of (\(R\))-terbutaline (63) started from the \(O\)-bisallyl protected cyanohydrin (\(R\))-58 via a Ritter \(N\)-tertiary butylation to provide the amide (\(R\))-60. Hydrogenation of (\(R\))-60 gave the amino alcohol (\(R\))-61. After deprotection of the hydroxyl functions, (\(R\))-terbutaline hydrochloride (\(R\))-63·HCl was obtained in an overall yield of 44% with \(>98\% ee\) (Scheme 10.10).

Asymmetric Oxidations  The direct conversion of a carbon–hydrogen bond to a carbon–hydroxyl bond with defined regio- and stereospecificity has given the organic chemist the ability to produce a wide range of compounds with the use of a
microorganism. The first application of this technique was for the synthesis of the anti-inflammatory corticosteroids and followed the discovery in 1950 by Murray and Peterson that the fungus *Rhizopus arrhizus* (isolated from the air in Kalamazoo, Michigan) was able to convert progesterone (64) into its 11α-hydroxy derivative (65) in high isolable yield. Since that time, microbial hydroxylation has been extensively applied in the production of steroids on a commercial scale (Holland, 1982).

Microbial hydroxylation of progesterone (64) in the culture of *Acremonium strictum* PTCC 5282 produced two hydroxylated pregnene-like steroids (Faramarzi et al., 2003) (66 and 67). Steroid 15α-hydroxylation of 13β-ethyl-4-gonene-3,17-dione (68) is an industrially important intermediate of the synthesis of Gestoden, a widely used contraceptive drug. Jekkel et al. (1998) reported that 68 could be hydroxylated to produce the 15α-hydroxylated product 69 by the enzyme of *Fusarium nivale* (VJ-63 strain) in a high yield.

Omapatrilat (70) is an antihypertensive drug, which acts by inhibiting angiotensin-converting enzyme (ACE) and neutral endopeptidase (Patel, 2001b). A novel L-lysine-ε-aminotransferase has been isolated and then cloned, overexpressed in *Escherichia coli*, and produced on large-scale fermentation by using the selective enrichment technique. The isolated enzyme is capable of oxidizing the ε-amino group of L-lysine, which can be used in the oxidation of the dipeptide 71 to the aldehyde intermediate 72. Through the conversion of 72 to 73, a key chiral intermediate for the synthesis of omapatrilat (70) was obtained. The aminotransferase reaction required α-ketoglutarate as the amine acceptor. Glutamate formed during this reaction was recycled back to α-ketoglutarate by glutamate oxidase (GOX) from *Streptomyces noursei* SC 6007. The conversion of 71 to 73 proceeded with an overall yield of 65–70% (Zaks, 2001) (Scheme 10.11).

**Enantioselective Hydrolysis and Esterification** Esterases (EC 3.1.1.x) represent a diverse group of hydrolases catalyzing the cleavage and formation of ester bonds and are widely distributed in animals, plants, and microorganisms. Besides lipases, a considerable number of microbial carboxyl esterases have also been discovered and overexpressed (Bornscheuer, 2002). For example, the S-monoester 74 is a key intermediate for the synthesis of β₃-receptor agonists. The enantioselective enzymatic hydrolysis of diester 75 to the desired acid ester 74 by pig liver esterase has been demonstrated. The reaction yields and ee of 74 were dependent upon the solvent used. High ee values (>91%) were obtained with methanol, ethanol, and toluene as a cosolvent; ethanol gave the highest reaction yield (96.7%) and ee (96%) (Patel et al., 1998).

Chiral tertiary benzylic centers are present in the family of aromatase inhibitory drugs such as (R)-(++)-amino-glutethimide (76), which was originally developed as an anticonvulsant. The first enzymatic attempt to construct an asymmetric quaternary carbon of (R)-(++)-76 has been undertaken by Fadel and Garcia-Argote. (R)-(++)- and (S)-(---)-aminogluthemide was synthesized in 97% ee via enzymatic hydrolysis (PLE acetonic powder) of malonate derivative (R)-(++)-77 followed by chemical modification (Fadel and Garcia-Argote, 1996) (Scheme 10.12).

Another enantioselective synthesis of (R)-(++)-76 via lipase-catalyzed kinetic resolution of (±)-4-cyano-4-phenyl-1-hexanol (78) as appropriate precursors possessing
an asymmetric quaternary carbon has been reported. Enzymatic transesterification of primary alcohol \((\pm)-78\) using *Pseudomonas cepacia* (Amano PS, PCL) provided the enantiopure alcohol \((R)-(\pm)-78\) with 99% ee at conversion of 86%, while that of \((\pm)-78\) using *Pseudomonas fluorescens* (Amano AK, LAK) provided the \((S)-(\pm)-78\) with 96% ee at conversion of 86%. Chemical transformation of substrate \((R)-(\pm)-78\) gave \((R)-(\pm)-\text{aminoglutethimide (76)}\) enantioselectively high yield (Im, 2003) (Scheme 10.13).

Two *Nocardioides* strains containing the novel hydrolases C-13 taxolase and C-10 deacetylase, capable of cleaving the corresponding sidechain esters from taxol, have been isolated. Interestingly, in addition to the hydrolysis reaction, the intracellular C-10 deacetylase was found to catalyze the acylation of 10-deacetyl baccatine
(80) to baccatine III (81), an intermediate in the semisynthetic synthesis of taxol (82). The process proceeded in 51% yield and required no protection of the 7-hydroxyl group because of the high regioselectivity of the hydrolase (Patel et al., 2000) (Scheme 10.14).
Other Chemo-enzymatic Syntheses The use of different classes of enzymes for the catalysis of many different types of chemical reactions has enabled a wide variety of chiral drugs to be generated. These include numerous hydrolytic enzymes and oxidoreductases. With the development of biocatalysts and fermentation technologies, a variety of enzymes and microbial cultures are regarded as tools for organic synthesis, such as aminotransferases, aldolases, decarboxylases, mono-oxygenases, and dioxygenases (Patel, 2001a).

L-Ephedrine (8a) hydrochloride, used as a decongestant and anti-asthmatic, is obtained from the extracts of a medicinal plant belonging to the genus Ephedra. In the pharmaceutical industry, L-phenylacetylcarbinol (L-PAC) is the commercial precursor for the production of L-ephedrine (8a) and D-pseudoephedrine (9b). L-PAC is produced from benzaldehyde and pyruvate catalyzed by the enzyme pyruvate decarboxylase (PDC), which is known to exist in certain yeast strains, especially the genera Succharomyces and Candida (Leksawasdi et al., 2005; Tripathi et al., 1997) (Scheme 10.15).

For enzymatic synthesis of eugenyl α-glucoside (83) as a promising pro-drug for a hair restorer and a derivative of spices in one step, selective α-glucosylation of eugenol (84) was carried out using the α-glucosyl transfer enzyme of Xanthomonas campestris WU-9701. The maximum molar conversion yield based on the amount of eugenol supplied reached 52% (Sato, 2003) (Scheme 10.16).
Crixivan (10) is an orally active HIV protease inhibitor for the treatment of AIDS. It contains five asymmetric centers allowing 32 possible stereoisomers, only one of which, in enantiomeric purity in the structure of 10, is the active ingredient. Production of the key intermediate cis-amino indanol (85) is technically demanding for synthesis of Crixivan (10). Its chemical synthesis proceeds through epoxidation of indene to indan oxide at ee 87%, followed by two steps to afford 85 in greater than 99% ee. Two alternative processes currently exist: biotransformations of indene (86) either through the transformation of indandiol (88) at >95% ee followed by its conversion to 85, or through from indan oxide (87) at high enantio-purity followed by its conversion to 85 (Buckland, 1999). In both, cis-(1S, 2R)-indandiol (88) or trans-(1R, 2R)-indandiol (89) are potential precursors to 85. Enrichment and isolation of microbial cultures yielded two Rhodococcus species—strains B 264-1 (MB 5655) and I-24 (MA 7205)—enabling the biotransformation of indene (86) to cis-(1S, 2R) indandiol (88) and trans-(1R, 2R) indandiol (89), respectively. Isolated MB 5655 was found to have a toluene dioxygenase, whereas MA 7205 was found to harbor toluene and naphthalene dioxygenases as well as a naphthalene monooxygenase, which catalyzes the above biotransformation (Patel et al., 2000) (Scheme 10.17).

10.3.3 Chemical Control

10.3.3.1 Separation of Enantiomers of Racemate of Chiral Drugs

If the chiral pool is not, or cannot, be employed, resolution is thought to be the most frequently used alternative industrial method for obtaining single enantiomers, although it is so far difficult to estimate exactly what the balance is between the use of chiral pool, asymmetric synthesis, and resolution (Crosby, 1997).

Despite the disadvantage that over half the amount of the racemate has to be discarded or transformed to a racemate again for reuse via chemical sequences, our rough impression based on current business trends is that some 30–50% of the single enantiomers required are obtained by classical resolution procedures (Rouhi, 2004). Although a number of stereoselective syntheses have been described and applied to the production of single enantiomeric substances, relatively few are selected for large-scale preparations, particularly at the early stages of development of new drugs. Obviously, it is urgent for researchers to provide some amounts of pure enantiomers for primary pharmacological tests before a manufacturing route (protocol) can be
selected. At these early stages, the development of an asymmetric synthesis would be both expensive and time-consuming, and thus techniques as simple as possible for the separation of enantiomers have an interesting potential (Maier et al., 2001).

**Preferential Crystallization** Spontaneous resolution is a phenomenon in which a particular racemate (racemic mixture or conglomerate) undergoes separate crystallization of both of (+) and (−)-enantiomers (antipodes) from solution. This was discovered by Pasteur in 1848 in the aqueous crystallization of ammonium sodium (+)-tartarate (92). As shown in Scheme 10.18, 50% each of (+)-92 and (−)-92 enantiomers are theoretically formed (Matsuura and Koshima, 2005).

Although the conglomerate formation is valuable for optical resolution, the frequency of the conglomerate occurrence has been estimated to be in the range of only 5–10% of all racemates. In order to find the controlling factors of the conglomerate crystal formation, extensive research has been carried out. The molten state of 1,1-binaphthyl and the mechanically stirred crystallization of NaClO$_4$ from its aqueous solution are two typical chiral breaking examples of crystallization by “external influences or measures” to affect the resolution of racemates (Kitoha et al., 2005). In another example, the spontaneous resolution of racemic 93 with Cd(ClO$_4$)$_2$·6H$_2$O in the presence of 2-butanol under solvothermal conditions favors the formation of crystal 94, while a similar reaction in the presence of ethanol only favors the formation of crystal 95 (Scheme 10.19). The crystal structural determination shows that both 94 and 95 crystallize in chiral enantiomorphous space groups (Ye et al., 2005).
Because of practical difficulties in discriminating one enantiomeric crystal from another by the visual observation of their crystal morphology, the so-called preferential crystallization is carried out. The preferential crystallization is one of the practical techniques for obtaining enantiomerically pure compounds, where the seeding of a small amount of the enantiomer in its supersaturated solution can preferentially crystallize one of the enantiomers. Resolution via preferential crystallization is known to be practical and cost-effective in that it requires no chiral reagent or enzyme and can be applied to industrial-scale preparations (Jacques et al., 1981). The Merck Company has a long history in the preparation of L-$\alpha$-methyldopa (96) via a continuous fluidized-bed crystallization resolution-racemization process. The intermediate (±)-97 for the synthesis of L-$\alpha$-methyldopa possesses the properties of a conglomerate, which can be separated by preferential crystallization. The resulting L-isomer (98) was hydrolyzed to L-$\alpha$-methyldopa with concentrated HCl. The undesired D-isomer (99) was racemized with NaCN in DMSO and returned for the resolution (Scheme 10.20). Using this process, Merck produced millions of kilograms of L-$\alpha$-methyldopa over the lifetime of the product (Grabowski, 2005).

The antibiotic chloramphenicol has been used widely for the treatment of numerous microbial infections for nearly 50 years. Only the D-(−)-threo-isomer (100) of four chloramphenicol stereoisomers has antibacterial activity. The stereochemistry of chloramphenicol is another example of industrial-scale preparation via preferential crystallization. The three-racemate of a synthetic intermediate is separated from the corresponding erythro-racemate at some stage in the synthetic sequence, and the desired D-(−)-threo isomer chloramphenicol is then obtained via resolution of the fully elaborated bases 101 and 102 into its enantiomers using the D-threo base, 101 as crystal seed, followed by introduction of the dichloroacetyl group (Liu and Fang, 1997).

**Diastereomer Crystallization** Preferential crystallization is a cheap and efficient resolution process, but the major limitation is that the racemic mixture must crystallize as a conglomerate, and the resolution results are composed of simultaneous nucleation and growth steps (Beilles et al., 2001). However, approximately 10% of all compounds that exist as racemates can be classified as...
conglomerate while the rest of them, ~90%, cannot be separated by preferential crystallization. However, these racemic compounds can be separated by diastereomer crystallization. In this method, a target racemate is mixed with a chiral reagent (known as the resolving agent) to give a mixture of diastereomers, which is repetitively recrystallized from an appropriate solvent to give the enantiomerically pure diastereomer. After removal of the resolving agent, the target enantiopure isomer can be finally isolated. Most commonly, the racemic mixture is a carboxylic acid, and the chiral reagent is an amine (or vice versa, a racemic mixture of an amine and a chiral carboxylic acid). An example of this process is illustrated in the resolution of (+)-lactic acid using ethylamine (Scheme 10.21).

The first amino acids were resolved by Emil Fischer in 1899 using the naturally occurring alkaloids strychnine (103) and brucine (104) to separate the racemic N-benzoyl derivatives of alanine, glutamic acid, and aspartic acid (Brittain, 1990). Although quite a lot of racemic compounds have been separated by diastereomer crystallization, it is still difficult to predict appropriate conditions for the successful resolution of a given target racemate. The practical and economical factors associated with the resolution of compounds by diastereomer crystallization must be considered, such as the choice of resolving agent, solvent composition, racemization of the unwanted isomer, and recovery of the chiral auxiliary (Challener, 2001b).

**Resolving Agents** Although the process of crystallization is affected by numerous factors, such as cooling rate, stirring, trace contaminants, and solvent, the crystal structures of substances obtained during optical-resolution advancements in the use of computer-aided molecular modeling have provided some assistance to professionals in the rational design and selection of resolving agents (Kinbara, 2005). The criteria for good resolving agents are as follows: (a) Diastereomeric compounds crystallize well; (b) formation of diastereomeric salts occurs readily; (c) diastereomeric salts can be separated easily; (d) inexpensive; (e) available in optically active pure form; (f) chiral center is as close as possible to the functional group; (g) responsible for salt formation; (h) available as both enantiomers; (i) have tight rigid structures; (j) strong acids/base give best results; (k) molecular weight as low as possible; and (l) easily
recovered (Sheldon, 1993). Many types of pharmaceuticals have been prepared utilizing resolving agents by formation of diastereomeric salts. A list of a few selected drugs and the associated resolving agents is shown in Table 10.6.

The chiral pool is one key source for resolving agents. There are also several commercially available synthetic resolving agents. Figures 10.8 and 10.9 demonstrate examples of common chiral acids (Sakai et al., 2003) and bases used as agents.

### Designed New Resolving Agents
Most of the popular basic resolving agents such as brucine, ephedrine, and morphine are naturally originating, highly toxic amines. Furthermore, only one enantiomer of each is commonly available.

#### TABLE 10.6 Drugs Prepared via Classical Resolution

<table>
<thead>
<tr>
<th>Chiral Drugs</th>
<th>Resolving Agent</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ampicillin</td>
<td>d-Camphosulfonic acid</td>
<td>Antibiotic</td>
</tr>
<tr>
<td>Ethambutol</td>
<td>l-(+)-Tartaric acid</td>
<td>Tuberculostatic</td>
</tr>
<tr>
<td>Chloramphenicol</td>
<td>d-Camphosulfonic acid</td>
<td>Antibiotic</td>
</tr>
<tr>
<td>Propoxyphene</td>
<td>d-Camphosulfonic acid</td>
<td>Analgesics</td>
</tr>
<tr>
<td>Dextromethorphan</td>
<td>d-Phenylsuccinic acid</td>
<td>Antihistamine</td>
</tr>
<tr>
<td>Fosfomycin</td>
<td>r-(+)-Phenethylamine</td>
<td>Antibiotic</td>
</tr>
<tr>
<td>Thiamphenicol</td>
<td>d-(−)-Tartaric acid</td>
<td>Anti-infective</td>
</tr>
<tr>
<td>Naproxen</td>
<td>Cinchonidine</td>
<td>Anti-inflammatory</td>
</tr>
<tr>
<td>Diltiazem</td>
<td>r-(+)-Phenethylamine</td>
<td>Calcium antagonist</td>
</tr>
</tbody>
</table>

*Source: Bayley and Vaidya (1992).*

![Figure 10.8](image1)  
*Figure 10.8* Examples of common chiral acids used as agents.
Therefore, the development of novel basic resolving agents, both of whose enantiomeric forms are readily accessible, is a subject of practical importance. A new chiral acid, the hydrogen phthalate of isopropylidene glycerol (105), has been shown to resolve a wide range of 1-arylalkylamines with high efficiency and with the same stereochemical outcome (Pallavicini et al., 2000).

Systematic studies on the resolution of 2-arylalkanoic acids (106) by enantiopure amino alcohols, expecting that the hydroxy group of the amino alcohol would play the same role as that of 1-arylglycolic acids in the formation of a hydrogen-bond sheet in the less-soluble diastereomeric salts, have been done by the group of Kinbara et al. They found that (1S, 3S)-109 is directly hydrogen-bonded to the

![Diagram of common chiral bases used as agents.](https://example.com/diagram.png)
adjacent hydrogen-bond column to form the supramolecular hydrogen-bond sheet. In addition, water molecules assisted stabilization of the supramolecular sheet externally. Hence, the skeleton of (1S, 3S)-109 is the most preferable for formation of a supramolecular sheet, and further design of basic resolving agents should be possible based on this skeleton (Brittain, 1990).

**Diastereomer Interconversion** A perceptible fact is that crystallization of diastereomeric salts can occur in conjunction with racemization of the unwanted isomer. Asymmetric transformation of diastereomeric salts, also known as diastereomer interconversion, can yield a theoretical 100% of one isomer. 2-Acyl-3, 6-dihydroxycyclo-hex-2-enones are isolated from the insects that exhibit kairomonal activity. The optically pure isomers (−)-(S)- and (+)-(R)-2-acetyl-3,6-dihydroxy cyclohex-2-enone (110) can be obtained through an enantioconvergent synthesis, starting from the racemic α-ketol (111) via diastereomeric N-tosyl-(S)-proline esters (112) by a repeated 4-dimethylaminopyridine (DMAP)-catalyzed equilibration and crystallization of the undesired diastereomer. At last, the enantiomeric 110 can be given by hydrolysis of the diastereomERICALLY pure ester 112 (Zaitsev and Mikhal’chuk, 2001) (as shown in Scheme 10.22).

**Kinetic Resolution—Chemical** Kinetic resolution can be achieved with chemical or enzymatic methods to promote selective reaction of one enantiomer over the other, giving a mixture of enantio-enriched starting material or product. The desired component is then isolated. Kinetic resolutions with enzymatic method have been described in Section 10.3.2.1. An interesting example of a parallel kinetic resolution involving asymmetric ketone reduction was reported by Kishi in the context of batrachotoxin synthesis. Batrachotoxins (113) are extremely potent neurotoxins that act as selective and irreversible Na⁺-channel activators (Kurosu and Kishi, 1998).

As shown in Scheme 10.23, essentially perfect selectivity was achieved in the reduction of model steroidal ketones using the commercial CBS catalyst, and the epimeric alcohol products were separated as the corresponding acetates by

![Scheme 10.22](https://via.placeholder.com/150)
chromatography. Limiting the reaction to 60% conversion led to kinetic resolution with recovered ketone in 40% isolated yield and 93% ee ($k_{rel} = 27$).

Sulfoxides possess the same stereochemical behavior as other chiral compounds that exist between the substituents of the stereogenic sulfur atom: a lone electron pair, an oxygen, and two different carbon ligands, which are able to differentiate the diastereotopic faces of a proximal or even remote reaction center (Carreno, 1995). Many chiral sulfoxides exhibit interesting biological activities, such as omeprazole (114), used as proton pump inhibitors to treat acid-related diseases, and OPC-29030 (115), used as a platelet adhesion inhibitor (Jia et al., 2004). Through kinetic resolution or, perhaps better, enantioselective sulfide oxidation coupled to kinetic resolution, enantiopure sulfoxides can be obtained in high ee.

The kinetic resolution of racemic sulfoxides alone with the Ti (IV)/BINOL/H₂O system has been shown to be an effective method for the preparation of enantiomerically pure sulfoxides, albeit with moderate isolated yields in most cases. The most promising system developed thus far is one in which alkyl aryl sulfides undergo enantioselective oxidation with aqueous TBHP as a stoichiometric oxidant in the presence of the Ti(OiPr)₄/BINOL catalyst. The same catalyst system can be applied to the kinetic resolution of racemic substrates as well as to provide enantiopure sulfoxides in moderate yield (Scheme 10.24) (Komatsu et al. 1992, 1993).

10.3.3.2 Asymmetric Synthesis and Its Application

In addition to chiral intermediates, enantioslective technologies (i.e., asymmetric synthesis) will also be on display at CPhI where the pharmaceutical industries have done a close collaboration with academicians. Innovative drug companies have always faced competition from one another as each strives to bring out the next member of a certain class of compounds for treatment of a disease. Asymmetric reactions have been applied
in several industrial processes, such as the asymmetric synthesis of L-DOPA (116), a drug for the treatment of Parkinson’s disease, via Rh(DIPAMP)-catalyzed hydrogenation of the enamide (Knowles, 1986), as shown in Scheme 10.25.

The industrial synthesis of (−)-menthol (117) and (−)-citronellal (118) through asymmetric isomerization of an allylic amine and asymmetric hydrogenation reaction (Akutagawa, 1997) are shown in Scheme 10.26.

Currently, the side chain (119) of efficient semisynthesis of taxol and decetaxol (Deng and Jacobsen, 1992) can also be synthesized via a highly enantioselective epoxidation reaction catalyzed by the readily available (salen)Mn(III) complex (120), as shown in Scheme 10.27.

**Asymmetric Hydrogenation Reaction** Hydrogen is the simplest molecule and is fully available in abundance at very low cost; therefore, asymmetric hydrogenation is a core technology in research and industry. Since Kang’s discovery of a practical synthesis of an air-stable ferroceny l bis(phosphine) and its application in the rhodium(I)-catalyzed enantioselective hydrogenation of dehydroamino acid derivatives, several ferrocenyl phosphines have found industrial applications in the synthesis of chiral
pharmaceuticals and agrochemicals. A process has been developed by Ciba-Geigy (now Novartis) for the production of an herbicide, (S)-metolachlor (121), which has been synthesized by chloroacetylation of the corresponding amine obtained by the Ir-catalyzed enantioselective hydrogenation of the requisite imine (Blaser and Spindler, 1997), as shown in Scheme 10.28. Similar technologies involving an Rh catalyst-containing ligand (123) were developed by Lonza Fine Chemicals in partnership with Ciba-Geigy for the production of the vitamin (+)-biotin (124) (Imwinkelried, 1997) (Scheme 10.29).

The antibacterial agent levofloxacin (125) can be prepared by utilizing catalytic asymmetric hydrogenation of the cyclic imine (126) with the iridium (I) complex of (2S,4S)-BPPM in the presence of bismuth (III) iodide to get the key intermediate (127), which can be easily converted to levofloxacin (125) in six steps, with high enantiomeric purity (90% ee) and a chemical yield of 96% (Satoh, 1998) (Scheme 10.30).
Asymmetric Hydrogenation of Ketones  Asymmetric hydrogenation of ketones is one of the most efficient methods for making chiral alcohols. This methodology has been used in the industrial production of synthetic intermediates for some important antibiotics: carbapenem (128) and sanfetrinem (129). Ru-BINAP catalysts are highly effective in the asymmetric hydrogenation of functionalized ketones. The Ru-BINAP complexes effect highly enantioselective hydrogenation of various β-keto esters to yield chiral β-hydroxy esters. By combining this asymmetric reaction with dynamic kinetic resolution, an important chiral intermediate for the synthesis of carbapenem antibiotics (Noyori and Ohkuma, 2001) was synthesized, as shown in Scheme 10.31. In the case of the preparation of sanfetrinem (129), the key step is the asymmetric hydrogenation of racemic 2-methoxycyclohexanone (130) with an Ru-(S)-3,5-xylyl-BINAP-(S, S)-DPEN-KOH catalyst leading to (1R, 2S)-2-methoxycyclohexanol (131) as the nearly sole stereoisomer (Matsumoto et al., 1999), as shown in Scheme 10.32.

Asymmetric Cyclopropanation Reaction  The chiral cyclopropyl group is one of the most important structural motifs responsible for biological activity of molecules, such as curacin A (132) and plakoside A (133). Curacin A (132) is a novel antimitotic agent recently isolated from a Caribbean cyanobacterium, Lyngbya majuscula (Onoda et al., 1995). Therefore, asymmetric cyclopropanation has been used widely in synthetic organic chemistry. Plakoside A (133) is a prenylated galactosphingolipid isolated from the marine sponge Plakortis simplex and exhibits strongly immunosuppressive activity with rather low cytotoxicity (Seki et al., 2001). Asymmetric synthesis of the cyclopropane moiety of curacin A (132) is shown in Scheme 10.33. Double-asymmetric Simmons-Smith cyclopropanation of the diene
(134) derived from diethyl L-tartrate in four steps proceeded with excellent diastereofacial selectivity (>99% de) to give the dicyclopropane (135), which was converted to the desired carboxylic acid (–)-136 in three steps (Onoda et al., 1996).

Cilastatin (137), an excellent inhibitor of dehydropeptidase-I, which increases the in vivo stability of imipenem, has desired pharmacological properties (Graham et al., 1987). The (S)-(+)–2,2-dimethylcyclopropane carboxylic acid (++)-138, a key building block for the synthesis of cilastatin (137), was prepared from 2-methylpropene and chiral iron carbene complex via asymmetric cyclopropanation reaction and exhaustive ozonolysis with up to 92% ee (Wang et al. 1998) (Scheme 10.34).

Asymmetric Hydroformylation Reaction 1β-Methylcarbapenem (139) is an unnatural antibiotic that possesses an excellent antibacterial profile as well as enhanced chemical and metabolic stability. Monocyclic β-lactam (140), a valuable intermediate in the preparation of 139, was synthesized by the asymmetric hydroformylation of the 4-vinyl β-lactam (141) catalyzed by rhodium aminophosphonite-phosphinite (143) and rhodium aminophosphine-phosphite (144) complexes. The stereoselectivity to the desired β-isomer was related to the presence of a substituent at the N atom of the β-lactam ring. The regioselectivity (branched/
linear) but not the stereoselectivity ($\beta/\alpha$) was found to be dependent on the ratio of substrate to catalyst (Cesarotti and Rimoldi, 2004) (Scheme 10.35).

**Asymmetric Dihydroxylation Reaction** Since Sharpless’s discovery of the asymmetric dihydroxylation reaction of alkenes mediated by osmium tetroxide-cinchona alkaloid complexes, continuous efforts have been made to utilize this reaction in preparation of chiral drugs or intermediates. For example, the

```
Scheme 10.34
```

```
Scheme 10.35
```

![Diagram](image-url)
non-steroidal anti-inflammatory drug (NSAID) naproxen (145) has been synthesized in high enantiomeric excess by Sharpless asymmetric dihydroxylation of the appropriate methyl styrenes (Griesbach et al., 1997) (Scheme 10.36).

Asymmetric Aminohydroxylation Reaction Loracarbef (146) is a carbacephalosporin antibiotic with extended chemical and serum stability, which is especially used in the treatment of pediatric ear infections. A formal synthesis of loracarbef was accomplished with high stereoselectivity. Stereoselective construction of the key intermediate cis-3,4-disubstituted azetidinone (147) was performed by employing intramolecular cyclization of cis-substituted azetidinone skeleton (148), which was obtained from Sharpless asymmetric aminohydroxylation of $\alpha,\beta$-unsaturated ester 149 in good yield with high regioselectivity (>13:1) and enantioselectivity (89% ee) (Lee et al., 2001) (Scheme 10.37).

Asymmetric Sulfoxidation Reaction Sulfide (150) is a potent systematically available ACAT inhibitor, which is rapidly metabolized into the corresponding sulfoxide. Of the two-sulfoxide enantiomers, only one—RP 73163—exhibits useful biological activity. The asymmetric synthesis of sulf oxide RP 73163 is based on the enantioselective oxidation of a suitably designed prochiral methyl sulfide, followed by $\alpha$ alkylation of the resulting sulfoxide (Scheme 10.38).
Asymmetric Codimerization Reaction  Substituted 8-oxabicyclo[3.2.1]octanes (151) are used as a key intermediate for the synthesis of tromboxane-A_2 (TXA_2) analogues. A novel and unexpected stereoselective synthesis of functionalized 151 has been finished by interaction α,β-unsaturated ketones 152 with γ,δ-unsaturated ketones

![Scheme 10.38](image_url)

**TABLE 10.7 Synthesis of Substituted 8-Oxabicyclo[3.2.1]octanes Catalyzed by Rh(I)-Sn(II) System**

<table>
<thead>
<tr>
<th>R_1</th>
<th>R_2</th>
<th>Conversion of 153 (%)</th>
<th>To converted 153 (%)</th>
<th>To mol/g at Rh</th>
</tr>
</thead>
<tbody>
<tr>
<td>Me</td>
<td>Me</td>
<td>31</td>
<td>39</td>
<td>12</td>
</tr>
<tr>
<td>Ph</td>
<td>Me</td>
<td>54</td>
<td>26</td>
<td>11</td>
</tr>
<tr>
<td>t-Bu</td>
<td>Me</td>
<td>26</td>
<td>14</td>
<td>4</td>
</tr>
<tr>
<td>5-(2-methylfuryl)</td>
<td>Me</td>
<td>47</td>
<td>33</td>
<td>15</td>
</tr>
<tr>
<td>Me</td>
<td>Ph</td>
<td>94</td>
<td>13</td>
<td>12</td>
</tr>
</tbody>
</table>

\(α,β\)-Unsaturated ketones (9.26 mmol), \(γ,δ\)-unsaturated ketones (9.26 mmol), [RhCl(C_2H_4)_2] (0.046 mmol), SnCl_2·2H_2O (0.19 mmol), acetone 3 mL, 80°C, 10 h.

The yield of isolated products.

Source: Data were taken from Pitchen et al. (1994).
catalyzed by the \([\text{RhCl(C_2H_4)2}_2\text{SnCl}_2]\) system (Kovalev et al., 1992). The reaction proceeds stereoselectively and gives only one *endo*-isomer with the carbonyl group in the equatorial position, as shown in Scheme 10.39. The results of the reaction are listed in Table 10.7.

### 10.4 Biotargets-Directed Design and Discovery of Bioactive Compounds

It has long been recognized that knowledge of the 3D structures of proteins has the potential to accelerate drug discovery. Many new protein targets have been identified from genome analyses and studied by X-ray analysis or NMR spectroscopy (Congreve et al., 2005). Described here is an example of using GABA transporters as targets for a drug discovery project.

#### 10.4.1 GABA Transporter

\(\gamma\)-Aminobutyric acid (GABA), the major inhibitory neurotransmitter in the brain, has been the target of extensive pharmacological research. GABA reuptake from the synaptic cleft is an important mechanism in the regulation of GABA activity, which is effected by a selective transporter named a GABA transporter (GAT) (Radian et al., 1986). GAT, together with the betaine, taurine, and creatine transporters, belongs to a subfamily of \(\text{Na}^+\) (and \(\text{Cl}^-\)) ion-dependent transporters. The human GAT proteins consist of 599 amino acids and a putative structure (Nelson et al., 1990; Kanner, 1994; Gao et al., 1999), as shown in Fig. 10.10.

In the central nervous system, GABA transport between neurons is effected by the release of neurotransmitters from the presynaptic neuron into the synaptic cleft to interact with GABA receptors in the membrane of the postsynaptic neuron. The GABA transport is ended by reuptake of the transmitter from the synaptic cleft by transporters located in the presynaptic membrane and is then partly inactivated enzymatically and partly stored in presynaptic vesicles for further use. The uptake in vesicles is caused by transporters located in the vesicular membrane. Glial cells surrounding the synapse also inactivate the transmitter by uptake from the synaptic cleft by transporters followed by enzymatic inactivation. The reuptake is effected by selective GABA transporters. Inhibitors of uptake transporters enhance the concentration of the transmitter in the synaptic cleft, thereby intensifying the action of the transmitter on its postsynaptic receptors, a useful property in case of transmitter malfunction (Soudijn, 2000). The process of normal release and reuptake of GABA is illuminated in Fig. 10.11.
In order to elucidate the functional roles of GABA transport processes, it is necessary to obtain detailed knowledge about the anatomical distribution of GABA transporters as well as their pharmacological properties. The cloning of four GABA transporters responsible for the cellular reuptake of GABA (mouse nomenclature GAT1-4) has allowed such studies to be performed over the past 15 years (Schousboe et al., 2004). Inhibition of the reuptake of GABA by potent and selective inhibitors of the GABA transporter enhances GABA activity. A therapeutic benefit is obtained by inhibiting these proteins in disorders in which GABA reserves are depleted—for instance, epilepsy or psychiatric disorders. So far, only GAT1 has convincingly been targeted in a selective manner, and little is known about the significance of the other subtypes (Dalby, 2003).

Figure 10.11  The process of normal release and reuptake of GABA (taken from www.gabitril.com). See color plates.
10.4.2 Inhibitors of the GABA Transporter

Some neuromedicinal chemistry groups have previously made significant contributions in the development of compounds that selectively interfere with the systems transporting this crucial neurotransmitter without affecting signaling mechanisms directly. GABA and several of its conformationally restricted analogues effectively inhibit the uptake of [3H]-GABA into rat brain synaptosomes. The rank-order of potency is homo-β-proline 154 = (R)-nipecotic acid 155 > GABA 156 > cis-4-OH-nipecotic acid 157 > guvacine 158 > (S)-nipecotic acid 159 (Ali et al., 1985).

Although (R)-nipecotic acid and guvacine are effective inhibitors of the GABA transporter in vitro, the compounds are not centrally active after systemic administration. This is probably due to the hydrophilic character of these cyclic amino acids. Therefore, in order to develop more lipophilic GAT inhibitors, the nitrogen atom of nipecotic acid and guvacine was substituted with diarylalkenyl groups. Some optimal active compounds had been obtained, such as SKF89976-A (160), CI-966 (161), SKF100330- A (162), tiagabine (163), NO-711 (164), and other analogues (Andersen et al., 1999, 2001).

From the data of GABA transporter inhibitory activity, it can be seen that (R)-nipecotic acid derivatives have more potent effects than (S)-nipecotic acid derivatives. For example, the GABA transporter inhibitory activity of (R)-nipecotic acid (155) (IC\textsubscript{50} = 1.69 μM) is more potent than that of (S)-nipecotic acid (159) (IC\textsubscript{50} = 9.88 μM). (R)-SKF89976-A (160) (IC\textsubscript{50} = 0.11 μM) has a more potent inhibitory effect than (S)-SKF89976-A (165) (IC\textsubscript{50} = 1.91 μM).

The (R)-nipecotic acid and guvacine derivatives have been the starting key for drug development programs in different companies and have led to a compound applied clinically as an antiepileptic drug as well as pointing to other potential therapeutic areas—that is, sleep disorders and pain. This drug is tiagabine hydrochloride (163) (TGB, Gabitril\textsuperscript{R}), a selective γ-aminobutyric acid (GABA) reuptake inhibitor (SGRI), which increases synaptic GABA availability via inhibition of the
GAT-1 on presynaptic neurons and glial cells (Borden et al., 1994). TGB has been shown to be effective as add-on therapy in adults and children 12 years and older with refractory partial seizures with or without secondary generalization. To date, exposure to TGB amounts to 147,686 patient-years (Bialer et al., 2004).

### 10.4.3 Design of New Inhibitors of the GABA Transporter

Knud Erik Andersen and co-workers (Knutsen et al., 1999) have designed some novel GABA uptake inhibitors of remarkable potency, using a putative new model of ligand interaction at GAT-1, as shown in Fig. 10.12. This model involves the postulated interaction of an electronegative region in the small-molecule GABA uptake inhibitor with a positive domain in the protein structure of the site controlling GABA uptake. The increasing electronegative character of the linker is coupled to the increase in potency of the compounds as inhibitors of $[^{3}H]$-GABA uptake \textit{in vitro} (Liu et al., 1993; Borden et al., 1994; Bennett and Kanner, 1997).

On the basis of the putative new model of ligand interaction at the GABA transporter type 1 and the efforts of our continued research in this area, we have designed and synthesized a series of new (R)-nipecotic acid derivatives containing thio/sulfinyl groups. The general structure of these compounds is shown in Fig. 10.13. Extensive animal studies with these compounds have revealed that some of them shown good inhibitory activity of $[^{3}H]$-GABA uptake \textit{in vitro} of culture cells. Hopefully, through further research of these compounds, new compounds will be discovered that can help elucidate the mechanisms of ligand interaction at the GABA transporter type 1 as well as provide drug candidates.

![Figure 10.12](image1.png)  
**Figure 10.12** Putative new model of ligand interaction at GAT-1.

![Figure 10.13](image2.png)  
**Figure 10.13** General structure of (R)-nipecotic acid derivatives containing thio/sulfinyl groups.
REFERENCES


REFERENCES


11.1 INTRODUCTION

The α-helix, first characterized by Linus Pauling in 1951 (Pauling et al., 1951), has been extensively studied due to its prevalence in structural biology (Brandon, 1991). α-Helices are the most common secondary conformation in natural proteins (~40% of amino acids adopt helical conformations) (Ruan et al., 1990). A typical α-helix completes one rotation with 3.6 amino acid residues, each with backbone dihedral angles of $\Psi = -50^\circ$ and $\Phi = -60^\circ$ (Fig. 11.1) (Pauling and Corey, 1954). The helix has a rise of 1.5 Å/residue or 5.4 Å/turn (Regan, 1993); as a result, the side chain of a certain residue at position $i$ projects from the same face as those from the $i + 3$, $i + 4$, and $i + 7$ residues in the sequence. There is a large entropy cost in helix formation, which is thermodynamically compensated by the formation of intramolecular hydrogen bonds. The backbone of the α-helix is primarily stabilized by hydrogen bonds between the carbonyl oxygen at the $i$ position and the carboxamide hydrogen at the $i + 4$ position (Pauling and Corey, 1954). Because all amides are oriented in the same direction and also the hydrogen bonding sites on the first and last turns are unfulfilled, a macro-dipole is produced (Galoppini and Fox, 1996; Scholtz et al., 1991). The positive end of the dipole is centered at the N-terminus, and the negative end is centered at the C-terminus. The total dipole is augmented if the peptide exists in conditions where both the termini are ionized. α-Helices composed of L-amino acids

*Corresponding author: andrew.hamilton@yale.edu

Drug Discovery Research: New Frontiers in the Post-Genomic Era, Edited by Ziwei Huang
Copyright © 2007 John Wiley & Sons, Inc.

281
almost always adopt a right-handed twist. The alternative left-handed conformation is destabilized due to steric crowding of the side chains and the backbone carbonyl groups. The helix propensities of the amino acids have been studied extensively in an attempt to predict secondary structure from primary sequences (Lifson, 1961; Zimm and Bragg, 1959). Certain amino acids, such as alanine, are more prevalent in $\alpha$-helices, while other residues, such as proline, are rarely found. The principal factors that dictate helix propensities are conformational restrictions of the side chains and steric considerations. $\alpha$-Helices are the bioactive regions of numerous proteins, playing critical roles in mediating many protein–protein, protein–DNA, and protein–RNA interactions (Fairlie et al., 1998). Therefore, they are attractive targets for the design of molecular mimetics, which could be very useful in the development of potential drugs and de novo proteins. In the past decade, major advances have been made in this field, evolving from strategies based on induced helix stabilization to the recent advent of helix proteomimetics, molecules that mimic the surface functionalities presented by $\alpha$-helical secondary structures (Peczuh and Hamilton, 2000).

### 11.2 CONSTRAINED PEPTIDES THAT ADOPT $\alpha$-HELICAL CONFORMATIONS

Peptides that adopt helical conformations would serve as good inhibitors of macromolecular interactions (Fairlie et al., 1998). However, peptide sequences derived from $\alpha$-helical domains in proteins frequently lack structural integrity due to the loss of the stabilizing environment provided by the protein. The modest thermodynamic gain of internal hydrogen bonding in an isolated helix is simply not enough to compensate for the large entropic penalty of conformational restriction for short peptides. Therefore, the goal of much research has been to develop methods that address this phenomenon by designing molecular systems that either initiate or stabilize $\alpha$-helicity in short peptides (Andrews and Tabor, 1999).

![Figure 11.1](image-url) (A) An $\alpha$-helix that is stabilized by internal hydrogen bonding. (B) Ribbon diagram showing rise per turn. (C) Amide backbone showing dihedral angles. See color plates.
11.2.1 Stabilization of α-Helical Structures Using Covalent Interactions

The formation of covalent linkages between adjacent residues in peptides has been shown to impart stabilization to the helical form of the peptide (Fig. 11.2). Disulfide bonds, lactam linkages, hydrazones, and carbon–carbon bonds have all been used to stabilize helices (Fig. 11.2). Noteworthy examples include a redox-triggered disulfide bond (1) used by Jaskson et al. (1991) and the lactam tethers (2) developed by Yu and Taylor (1999) and by Osapay and Taylor (1992) to promote helicity. Flint et al. (2002) have shown that an azobenzene (AZO) (3) linkage through two cysteine thiols can be used to both stabilize and destabilize the helical form of a peptide under photocontrol of the AZO trans/cis-isomerization. Blackwell and Grubbs (1998) have demonstrated that olefin metathesis (4) of two O-allyl serines at the i and i + 4 positions can facilitate the transition from an α-helix to a 310-helix in nonpolar solvents. In a more recent example, Schafmeister et al. (2000) have elaborated upon this work to show that enhanced and compromised helix stability can be conferred upon the C-terminal peptide of RNase A via incorporation and subsequent metathesis of α,α-disubstituted amino acids.

11.2.2 Stabilization of α-Helical Structures Using Noncovalent Interactions

Stabilization of α-helical structures has also been achieved using noncovalent interactions between appropriately spaced residues in a peptide chain. Metal–ligand interactions, designed host–guest interactions, salt bridges, cation–π interaction, and π–π stacking have resulted in helix stabilization (Fig. 11.3). Albert and Hamilton (1995) have studied hydrophobic interactions in the stabilization of helical structure in short peptides that incorporated two ε-(3,5-dinitrobenzoyl)Lys residues (5) at various positions. Solvent mixture titrations in TFE/water demonstrated that helical stability is highest for the peptide having a pair of modified residues at the i and i + 4 positions. Using an intermolecular hydrogen-bonding strategy, the same group has shown in several studies that guanidinium-based receptors can recognize aspartate residues at multiple i and i + 3 or i + 4 positions and induce α-helix formation (Peczuh and Hamilton, 2000; Albert et al., 1995; Peczuh et al., 1997). Scholtz et al. (1993) showed that the (i, i + 3) and (i, i + 4) Glu–Lys interactions (6) are

![Figure 11.2](image-url) Some examples of helical stabilization by covalent linkages: disulfide (1), lactam (2), azobenzene (3), olefin (4). See color plates.
helix-stabilizing and are similar to each other in strength, regardless of the orientation of the side chains. The same group also reported that an Arg–Asp salt bridge, spaced either \( i, i + 3, \) or \( i, i + 4 \) apart, stabilizes the \( \alpha \)-helical conformation in a similar fashion (Huyghuesdespointes et al., 1993). Olson et al. (2001) demonstrated that the cation–π interaction of a Trp–Arg \( (i, i + 4) \) pair stabilizes the \( \alpha \)-helical conformation with a quantified stabilization energy of \(-0.4 \text{ kcal/mol}\). Tsou et al. (2002) showed that the Phe–Lys cation–π interaction \( (6) \) contributed similar energy \((−0.4 \text{ kcal/mol})\) to the stability of an \( \alpha \)-helical peptide.

### 11.3 LOW-MOLECULAR-WEIGHT \( \alpha \)-HELICAL PROTEOMIMETICS

Although many of these helical stabilization methods have been very successful, the use of such peptide systems for exerting biological effects is often compromised due to poor cellular uptake and proteolytic degradation. The structural characteristics of \( \alpha \)-helices make them intriguing targets for low-molecular-weight proteomimetics. Such approaches offer the advantages of improved biostability, lower molecular weight (<750 Da), and better bioavailability (Lipinski et al., 1997). The challenge is to design synthetically viable scaffolds that project functionality with similar distance and angular constraints to the analogous side chains projecting from an \( \alpha \)-helical surface.

The development of small molecule modulators of protein–protein and protein–peptide interactions has been widely regarded as a formidable goal. Large interfacial areas, typically more than 1600 Å\(^2\) of buried area (around 170 atoms), are involved in forming a protein–protein interface, posing a serious challenge for any small molecule to be competitive (Stites, 1997). The binding regions of protein partners are often discontinuous and thus hard to be mimicked by simple synthetic peptides. Conventional methods for identifying inhibitors of protein–protein interactions require much input in the preparation and screening of a diverse chemical library to discover lead compounds. An alternative approach for identifying inhibitors is to design synthetic scaffolds that reproduce structural features of the protein secondary structure.
11.3.1 Indane-Based Mimetics that Target the Tachykinin Receptors

At this stage, only a few synthetic mimetics of α-helices have been reported. The pioneering work of Horwell et al. (1996) showed that 1,6-disubstituted indanes present functionalities in a similar spatial arrangement to the i and i + 1 residues of an α-helix (Horwell et al., 1996a,b). Trisubstituted indanes 8 (Fig. 11.4) were used to mimic the presentation of the i, i + 1 residues of an α-helix. Molecular modeling experiments involving the (S)-enantiomer of an Ala–Ala mimetic were conducted, overlaying the two Cα and the two Cβ carbon atoms of the lowest energy conformation of the indane system with the corresponding positions in a polyalanine α-helix. A root mean square (RMS) deviation of 0.15 Å was reported, suggesting that the template was orientated within the same space occupied by the helix backbone. Derivatives with large hydrophobic peptide side chains (Phe–Phe and Trp–Phe) were found to bind to the tachykinin receptors, NK1, NK2, and NK3, with micromolar affinity. This work represents a peptidomimetic approach to the design of molecules based on templating the i and i + 1 side chains of an α-helix, elucidating nicely the potential of using nonpeptide scaffolds to present critical recognition functionality in a suitable spatial orientation for binding to a target protein. However, these mimics do not cover a surface large enough to sufficiently represent an α-helical mimic. Although initial results suggest that this strategy is partly successful by producing compounds with biological activity similar to their dipeptide target, further application of this concept to the mimicry of other biologically active helices has not been reported. Also, extension of this strategy to mimic larger α-helical regions remains to be explored.

11.3.2 Pentasaccharide-Based Mimetic of GCN4

The work of Xuereb et al. (2000) represents the first example of a synthetic scaffold to mimic an extended surface area of an α-helix, in which a pentasaccharide scaffold (9, Fig. 11.5) was used to mimic the basic region of the transcription factor GCN4. GCN4 contains a bZIP protein that binds DNA in the major groove with one surface of an α-helical region (Ellenberger et al., 1992). A number of positively charged side chains emanating from an α-helical region of GCN4 are responsible for binding contacts to

![Figure 11.4](image-url)  
**Figure 11.4** A 1,6-disubstituted indane molecule, 8, that mimics of the i and i + 1 side chains (shown in yellow) of an α-helix. See color plates.
both the phosphate backbone and the nucleotide bases. To mimic the basic domain, a conformationally rigid oligosaccharide scaffold was designed in which positively charged guanidinium groups project from the surface in a similar array to that displayed by the α-helix target. The monomer chosen for the design is 2-deoxyfucose because it is more hydrophobic than fully oxygenated sugars, and this was expected to help facilitate groove binding (Kahne, 1995; Ding and Ellestad, 1991). Conformational analyses and molecular mechanics calculations showed that a pentasaccharide of this monomer with guanidinium groups attached at the C₃ positions was a suitable α-helix mimetic. The preferred conformation of this oligosaccharide in D₂O was calculated through a combination of 2D-NMR experiments and was found to match the predicted glycosidic linkage conformation. Using an ethidium bromide competition assay, pentasaccharide 9 was found to bind calf thymus DNA with a dissociation constant (K_d) of 10⁻⁶ M, which is similar to an unstructured pentalysine control peptide. However, the nonpolyelectrolyte contribution to binding was calculated to be 1 kcal/mol higher for the pentasaccharide, suggesting that there is a contribution to the binding that is not associated with electrostatics. Nuclear Overhauser effects (NOE) suggested that 9 binds to DNA in the minor groove with one surface of the oligosaccharide contacting the floor of the DNA. Preliminary experiments to determine selectivity suggested that 9 discriminates duplex DNA from other types of nucleic acid. Even at high concentrations, pentasaccharide 9 was unable to displace

---

Figure 11.5 Structure of pentasaccharide 9.
etidium bromide from double-stranded regions of yeast tRNA. However, the limits of this approach are quite apparent. The mimetic binds relatively weakly in comparison to the natural protein, although it mimics a large percentage of the side chains responsible for the binding energy of GCN4 to DNA. More importantly, pentasaccharide 9 fails to bind in the major groove of DNA, a point that weakens the argument of the helix mimicry design, especially since naturally occurring proteins rarely use the \( \alpha \)-helix motif in minor groove recognition.

### 11.3.3 Hairpin-Based Mimetics of p53 that Disrupt p53–HDM2 Interaction

Fasan et al. (2004) have recently reported structural \( \alpha \)-helical mimetics as antagonists of p53–HDM2 interaction based on a \( \beta \)-hairpin scaffold. p53 is a transcription factor mutated in approximately 50% of human tumors (Hollstein et al., 1994). Overexpression of wild-type p53 induces a large number of downstream genes that lead to cell cycle arrest or apoptosis. In unstressed normal cells, p53 is present at a very low level due to rapid degradation through the ubiquitin-dependent proteasome pathway. Human double minute 2 (HDM2) is an important regulator of p53 and acts by binding p53 to promote its ubiquitination (Haupt et al., 1997; Honda et al., 1997; Kubbutat et al., 1997). The overexpression of HDM2 abrogates the ability of p53 to induce cell cycle arrest and apoptosis (Chen et al., 1997). Therefore, deregulation of the overexpression or hyperactivation of HDM2 plays a key role in tolerance of wild type p53, making it an attractive target for the development of antitumor agents.

Recently, several approaches have been pursued to disrupt the p53–HDM2 interaction. García-Echeverría et al. (2000) have reported that a 13-residue peptide derived from the p53 protein (a.a. 16–28) gave a potent inhibition (IC \(_{50} = 8.7 \mu M\) of HDM2. Blaydes et al. (1997) used microinjection of the 3G5 antibody to inhibit p53 binding to HDM2 and activated p53 function \textit{in vivo}. Chen et al. (1998) showed that antisense oligonucleotides can inhibit MDM2, a murine analogue protein of HDM2, resulting in activation of p53 and cell growth arrest or apoptosis. These studies provided a proof-of-concept for activating p53 by targeting HDM2; yet the design of synthetic molecules that target the HDM2 exterior surface has been barely explored.

Robinson’s helical mimetic design is based on the X-ray crystal structure of the HDM2/p53 complex (Fig. 11.6a), which reveals that the p53 peptide adopts a largely amphipathic \( \alpha \)-helical conformation, with side chains of Phe19, Trp23, and Leu26 inserting into a hydrophobic cleft on the HDM2 surface. \( \beta \)-Hairpins (Fig. 11.6b) naturally occur in many protein–protein interfaces (Fasan et al., 2004). The distance between the C\(_2\) atoms of two residues \( i \) and \( i + 2 \) along one strand of a \( \beta \)-hairpin is similar to the \( \alpha \)-helical side chains at the \( i \) and \( i + 4 \) residues (Fasan et al., 2004). Incorporated with the functional epitope, a \( \beta \)-hairpin can act as the scaffold to present the Phe19 and Trp23 in a similar fashion as the p53 peptide. To test this design, a series of \( \beta \)-hairpins with different peptide sequences grafted onto one of the hairpin strands was prepared by solid-phase synthesis. \( \beta \)-Hairpin 9, which has a sequence of Phe1, Trp3, and Leu4, showed the ability to inhibit p53 binding to HDM2 with an IC\(_{50}\) of 125 ± 8 \( \mu M\) using a surface plasmon resonance (SPR) assay. The control compounds,
with Ala replacing either the Phe or Trp residue, were not able to block the HDM2–p53 complexation (IC₅₀ > 1000 μM), indicating the importance of the side-chain complementarity. The most potent β-hairpin inhibitor, with a sequence of Phe₁, (6Cl)-Trp (6-chlorotryptophan), and Leu₄, showed an increased affinity over the initial lead 9 by a factor of almost 900-fold (IC₅₀ = 0.14 ± 0.06 μM). As a comparison, the linear p53-derived peptide (a.a. 11–29) showed an IC₅₀ of 1.1 μM in the same assay. The binding modes of these β-hairpin mimetics were studied by ¹⁵N-HSQC NMR experiments. The residues that showed chemical shift changes upon addition of the hairpin mimetics are consistent with the p53-binding site revealed by X-ray crystallography, confirming that β-hairpins indeed act as structural mimetics of the p53 peptide.

11.3.4 Terphenyl-Based Mimetics that Target Calmodulin, gp41, and Bcl-x₅ Proteins

Hamilton and co-workers have reported a strategy of α-helix mimicry using oligo-phenylene scaffolds (Orner et al., 2001; Ernst et al., 2001; Kutzki et al., 2002). The goal was to reduce the complexity of the helical backbone while keeping the critical functionality in a similar spatial arrangement (Fig. 11.7). The first generation of the

Figure 11.6  (A) X-ray crystal structure of the HDM2/p53 complex. (B) β-hairpin inhibitor that disrupts the HDM2-p53 interaction. See color plates.

Figure 11.7  An α-helix backbone can be replaced by a synthetic scaffold that projects the functionalities in a similar fashion. See color plates.
design was based on a 3,2',2''-trisubstituted terphenyl scaffold (Orner et al., 2001; Ernst et al., 2001; Kutzki et al., 2002). This molecule was chosen for its low molecular weight, the modularity of its synthesis, and most importantly for its conformational characteristics. The ortho-functionalization serves to project the critical side chains that mimic the \(i, i+3, i+4,\) and \(i+7\) residues on one face of the helix (Fig. 11.8).

Due to the low rotational energy barrier about the phenyl-phenyl bonds in this system (Carreira and Towns, 1977), energetically minimized conformations are accessible that present the substituents in a way that mimics \(\alpha\)-helices (Fig. 11.8). An X-ray crystal structure of 3,2',2''-trimethyl-4-nitro-4''-hydroxy-terphenyl revealed that the molecule takes up the expected staggered conformation with dihedral angles of 59.1° and 120.7° (Fig. 11.8b) (Orner et al., 2001). Due to the low barrier of rotation, interconversion between different conformations in solution should be facile, thus easily populating the desired conformation for binding. The low rotational energy barrier (<15 kcal/mol), as exemplified by similarly substituted biphenyls, permits some desired conformational flexibility for adaptation to the target surfaces (Carreira and Towns, 1977). However, this flexibility will likely come with an entropic cost due to the recognition of the terphenyl on induced-fit-binding.

Computer modeling experiments comparing a polyalanine \(\alpha\)-helix to 3,2',2''-trimethylterphenyl indicated that the phenyl–phenyl torsion angles of 55° gave a conformation with close correspondence (RMS deviation = 0.9 Å) to the positions of the three methyl groups and the \(i, i+4, i+7\) alanine methyl side chains (Fig. 11.8c). Overlay of the isomeric low-energy conformation (in which the middle ring has phenyl–phenyl torsion angles of \(-55^\circ\)) with the \(i, i+3,\) and \(i+7\) alanine methyl side chains gave an RMS deviation of 1.06 Å (Fig. 11.8d).
11.3.4.1 Inhibitors of the Calmodulin/Smooth Muscle Myosin Light Chain Kinase Association

The complex of calmodulin (CaM) and smooth muscle myosin light chain kinase (smMLCK) was chosen as an initial system to study because of its extensive structural characterization (Meador et al., 1992, 1993). Calmodulin (calcium modulated protein) is a ubiquitous and essential Ca$^{2+}$-binding protein found in all eukaryotes (Vogel, 1994). It plays a pivotal role in the regulation of gene expression, protein syntheses, ion channel functions, cell motility, and chemotaxis by sensing the Ca$^{2+}$ level in the cytoplasm. CaM is a 17-kDa acidic protein containing two globular EF-hand motifs, each of which binds two Ca$^{2+}$ ions. Upon binding of Ca$^{2+}$, CaM undergoes a significant conformational change, resulting in exposure of two hydrophobic patches to recruit CaM-binding proteins (CaMBP). Many CaM-binding proteins recognize an $\alpha$-helix of roughly 20 amino acids. A representative example is the binding of smMLCK (RS20) to CaM. The crystal structure of this complex is shown in Fig. 11.9 (Meador et al., 1992). Mutational studies of RS20 have established that an $i$, $i + 3$, and $i + 7$ arrangement of residues (Trp800, Thr803, and Val807) is critical for binding.

Compound 11a was designed using this terphenyl mimetic strategy, and for synthetic simplicity the indole segment of Trp800 residue was changed (Fig. 11.10) to a phenyl group and the hydroxyl group of Thr803 was removed. Compound 11a was determined to bind CaM using an enzymatic assay (Johnson et al., 1987). The enzyme 3',5'-cyclic nucleotide phosphodiesterase (PDE) is a CaM-activated enzyme, which can be inhibited by RS20 via competition binding to CaM. Compound 11a effectively inhibits the ability of PDE to hydrolyze substrate with a comparable affinity (IC$_{50} = 0.8$ μM) to RS20 (IC$_{50} = 0.08$ μM). The control molecule 10, possessing three methyl side chains, showed no inhibitory effect under the same conditions (IC$_{50} > 20$ μM).

Figure 11.9  X-ray crystal structure of the CaM-binding region of smMLCK (residues 796-815, shown in red) in a complex with CaM. Side chains of the critical residues (Trp800, Thr803 and Val807) are shown in stick representation. See color plates.
A competition experiment using fluorescence spectroscopy was used to elucidate the binding site of these inhibitors. An α-helical peptide derived from the plasma membrane calcium pump, C20W, has been shown to bind CaM at the same site as smMLCK RS20 (Elshorst et al., 1999). Addition of dansylated C20W to CaM resulted in an increase in fluorescence, which was indicative of complex formation. Addition of \(11a\) to the complex showed the reverse effect, suggesting that \(11a\) and C20W are competing for the same binding site. The results from these assays suggested that an improved representation of the indole moiety (Trp800) in the design could perhaps optimize the fit into the CaM binding pocket and enhance potency. Molecules \(11b\) and \(11c\) were prepared in which a naphthylethene side chain replaces the benzyl substituent in \(11a\). These molecules showed more potency to inhibit PDE with \(IC_{50}\) values of 9 and 20 nM, respectively.

### 11.3.4.2 Inhibitors of the Assembly of the Hexameric Helical Core of the gp41 Protein

To show the generality of this terphenyl helical mimetic design, Hamilton et al. (2001) extended the use of the terphenyl compounds to the disruption of the assembly of the hexameric helical core of the gp41 protein. The envelope glycoproteins of HIV-1, gp120, and gp41 are generated from the polyprotein precursor gp160 by proteolytic cleavage. While gp120 is responsible for the target cell recognition and viral tropism through the CD4 and CCR5/CXCR4 cytokine coreceptors, the membrane spanning gp41 promotes the fusion of the viral and cell membrane. The helical domain of gp41 has an α6-helical bundle composed of trimers of C34 and N36 peptides (Lu et al., 1995). During the target cell binding, gp41 undergoes a conformational change and exposes its hydrophobic N36, allowing the fusion peptide to insert into the host cell membrane. The surface N36 has a particularly deep cavity into which three hydrophobic residues from C34: Trp628, Trp631, and Ile635 (Chan et al., 1997).
Synthetic molecules mimicking the interaction surface of C34 can be effective inhibitors to abolish the C34 and N36 interaction. Consequently, the C34 peptide of gp41 can be a useful target for inhibiting HIV-1 entry. A series of terphenyl derivatives was designed to mimic the hydrophobic functionality at the $i$, $i+4$, and $i+7$ positions of this $\alpha$-helical region of the C34 peptide, to disrupt binding to the trimeric N-terminal core, and to prevent HIV-1 from entering into host cells. Terphenyl 12 was found to disrupt the interaction of two peptides (N36 and C34) to form a stable six-helix bundle that is analogous to the gp41 core (Fig. 11.11). Circular dichroism (CD) spectroscopy demonstrated that the C34 peptide forms a random coil on its own in solution while the N36 peptide forms concentration-dependent aggregates. Titration of 12 into a PBS-buffered solution of this model system resulted in a decrease of the CD signal at 222 nm and 208 nm corresponding to a reduction in helicity of the hexameric bundle that saturates at three equivalents of terphenyl. The melting point ($T_m$) value of the N36 core in the presence of excess 12 is significantly lower than that of the gp41 core and resembles that of the N36 region alone. These results suggested that the terphenyl 12 competes for the C34 binding site. The electrostatic and hydrophobic features of 12 were found to be critical for this behavior because both the terphenyls lacking alkyl substituents and those with positively charged substituents had little effect on the CD signal of the gp41 core. An ELISA experiment using an antibody that binds to the N36/C34 complex but not the individual peptides showed that 12 disrupts the formation of the hexameric bundle with an IC$_{50}$ of 13.18 ± 2.54 µg · mL$^{-1}$, while a dye transfer fusion assay indicated that 12 inhibits HIV-1 mediated cell–cell fusion with an IC$_{50}$ of 15.70 ± 1.30 µg · mL$^{-1}$.

### 11.3.4.3 Inhibitors of the Bcl-xL/Bak Interaction

B-cell lymphoma-2 (Bcl-2) proteins are critical members in programmed cell death, also known as apoptosis (Adams and Cory, 1998; Reed, 1997). The Bcl-2 proteins...
include two groups: the proapoptotic and the prosurvival subfamilies. The proapoptotic subfamily proteins, such as Bak, Bad, and Bax, possess the death-promoting BH3 domain and act as a checkpoint downstream of p53 (Lane, 1992), and upstream of mitochondrial rupture and caspase cysteine proteases, which transduce the apoptotic signal (Adams and Corey, 1998, 2001, 2002). These molecules (Bak, Bax, Bad, and Bid) are able to induce apoptosis through heterodimerization with the prosurvival Bcl-2 family members (Chao and Korsmeyer, 1998). Previous studies showed that oncogenic mutations induced apoptotic defects through a Bcl-2-dependent pathway (Graeber et al., 1996; Fearon and Vogelstein, 1990). Overexpression of the prosurvival proteins, such as Bcl-2 and Bcl-xL, can inhibit the potency of many currently available anticancer drugs by blocking the apoptotic pathway (Strasser et al., 1997). Several low-molecular-weight inhibitors of Bcl-2 (Bcl-xL) have been identified by screening diverse chemical libraries (Enyedy et al., 2001; Wang et al., 2000; Lugovskoy et al., 2002; Tzung et al., 2001; Degterev et al., 2001). The rational design of agents that directly mimic the death-promoting BH3 domain of the proapoptotic subfamily of Bcl-2 proteins is an efficient alternative because it allows structure-based optimization of the initial hits (Adams and Cory, 2001).

A current strategy for inhibiting Bcl-xL function is to block the Bak-recognition site on Bcl-xL and disrupt the protein–protein contact. A Bcl-xL/Bak complex structure determined by NMR spectroscopy showed that a helical region of Bak (a.a. 72–87, \(K_d = 340\) nM) binds to a hydrophobic cleft on the surface of Bcl-xL, yet in solution this isolated 16-mer peptide exists as a random coil (Sattler et al., 1997). Furthermore, the crucial residues for binding were shown by alanine scanning to be Val74, Leu78, Ile81, and Ile85, which project at the \(i, i + 4, i + 7,\) and \(i + 11\) positions along one face of the helix (Fig. 11.12). A series of terphenyl derivatives was prepared

Figure 11.12  Results of the \(^{15}\text{N}-\text{HSQC}\) experiments of \(13\) binding to Bcl-xL. The residues that showed significant chemical shift changes in the presence of \(13\) are shown in yellow. The highest ranked binding mode of inhibitor \(13\) predicted from a computational docking simulation has been superimposed on the helical Bak BH3 domain for comparison. See color plates.
with different sequences and substitution patterns on the 3,2′,2″-positions. A fluorescence polarization assay was used to monitor the interaction between the inhibitor and the target protein. Terphenyl 13, with two carboxyl groups and a substituent sequence of isobutyl, 1-naphthylmethylene, and isobutyl groups, was identified as a potent inhibitor ($K_d = 114 \text{nM}$). The binding specificity was confirmed by scrambling the sequence of the substitutions, which led to a 30-fold drop in $K_i$. The importance of the two carboxyl groups was confirmed by partial removal or conversion to positively charged groups. $^{15}$N-HSQC NMR experiments with 13 indicated that these terphenyl derivatives target the hydrophobic cleft known to interact with the Bak peptide (shown in blue, Fig. 11.12).

### 11.3.5 Oligoamide Foldamers as Mimetics of the Bak Peptide

The hydrophobicity of these terphenyls compounds and their complex syntheses prompted us to search for simpler scaffolds that could similarly mimic the side chain presentation on an α-helix (Lipinski et al., 1991). The goal of these design modifications was to maintain the similarity between the arrangement of the $i$, $i+4$, and $i+7$ side chains of an α-helix and the substituents on 3,2′,2″-positions on terphenyl 13 (Kutzki et al., 2002), while minimizing the structural complexity and reducing the high hydrophobicity of the inhibitors. This strategy of simplifying a proven proteomimetic was accomplished by using oligoamide foldamer, 14, as the scaffold (Ernst et al., 2003). The anticipated conformation of the molecule 14 is depicted in Fig. 11.13a and should result from a favorable intramolecular hydrogen bond between the amide-NH and the pyridine nitrogen in addition to the disfavored interaction that would occur between the amide carbonyl and the nitrogen of the pyridine. Further conformational biasing is induced by a hydrogen bond between the amide-NH and the oxygen of the alkoxy substituent. The intramolecularly hydrogen-bonded conformation was confirmed

![Figure 11.13](A) X-ray crystal structure of an oligoamide foldamer 14. (B) Overlay of a polyalanine α-helix and 14, showing good agreement of the side chains. See color plates.
by X-ray crystallography, which further revealed that the alkoxy side chains were tilted at 45° to maximize interaction of the lone pair on the oxygen with the amide-NH. The foldamer was tested as a proteomimetic of Bak to disrupt its interaction with Bcl-xL. Three inhibitors, 14a, 14b, and 14c, were identified using fluorescence polarization assays with $K_i$ values of 2.3, 9.8, and 1.6 µM, respectively. The foldamers with tris-benzylxyl and tris-methoxyl groups showed weaker activities, indicating a requirement for size matching and sequence specificity between the protein and the mimetic.

### 11.3.6 Terephthalamide-Based Mimetics of the Bak Peptide

An alternative design for mimicking the binding surface of $\alpha$-helices based on a novel terephthalamide scaffold was reported recently (Yin and Hamilton, 2004). The flanking phenyl rings in 13 were replaced by two functionalized carboxamide groups, which retain the planar geometry that is presented by the phenyl rings, due to the restricted rotation of the amide bonds. These carboxamide groups also increase the polarity of the inhibitors. Computational calculations predicted that terephthalamide 15 ($R_1 = R_2 = R_3 = \text{Me}$) has a log $P$ (partition coefficient for $n$-octanol/water) of 2.73, compared to 10.84 for terphenyl 13, suggesting that the terephthalamide derivatives may have favorable solubility in aqueous solution. Another conformational constraint in the molecule was imposed by an intramolecular hydrogen bond between the amide–NH and the alkoxy oxygen atom, to influence the position of the amino acid side chain $R_1$ (Fig. 11.14a) (Ernst et al., 2003). The increased rigidity of the terephthalamide

![Figure 11.14](https://example.com/figure1114.png)

**Figure 11.14** (A) Terephthalamide helical mimetics. (B) The superimposition of 15 on the $i$, $i + 4$, and $i + 7$ positions of a poly-alanine $\alpha$-helix (RMS deviation = 1.03 Å). (C) The superimposition of terphenyl 13 and terephthalamide 15 (RMS deviation = 0.34 Å). The functional side chains are highlighted. See color plates.
avoids the entropy loss of the terphenyls upon binding to their protein targets. The modularity of inhibitor synthesis is also strengthened. Terphenyl 13 requires difficult C–C bond formations to attach the side chains to the scaffold and to assemble functionalized phenyl building blocks (Orner et al., 2001; Kutzski et al., 2002). Terephthalamide derivatives, on the other hand, were synthesized using O-alkylation and amide bond formation steps, accelerating the process of preparing these compounds. Computational energy minimization using MM2 force field within the Macromodel 7.0 program showed good stereochemical similarity among energy-minimized 15, the \( i, i + 4, \) and \( i + 7 \) side chains of an \( \alpha \)-helix (RMS deviation = 1.03 \( \text{Å} \)), and a comparable terphenyl scaffold (RMS deviation = 0.34 \( \text{Å} \)).

Terephthalamide 16 and 17 (Fig. 11.15) were identified as potent inhibitors of Bcl-xL with \( K_i \) values of 0.78 ± 0.07 \( \mu \text{M} \) and 1.785 ± 0.388 \( \mu \text{M} \), respectively, using the previously reported fluorescence polarization assay (Kutzki et al., 2002). Intact cell assays using human embryonic kidney 293 (HEK293) cells showed that terephthalamide derivatives inhibit the Bcl-xL/BH3 complexation in whole cells. Treatment of HEK293 cells with terephthalamide 17 (100 \( \mu \text{M} \)) resulted in inhibition of the association of Bcl-xL with Bax to an average extent of 81.1 ± 13.5%, compared to the 68.7 ± 20.3% inhibition induced by the Bax BH3 peptide. This result demonstrates that terephthalamide 17 is able to perturb Bcl-xL/Bax interaction in cells, suggesting that it has good uptake properties.

The displacement of the Bak peptide by a terephthalamide in the fluorescence polarization competition assay suggested that they both bind to the same hydrophobic pocket on the Bcl-xL surface. Comparison of the Bak peptide and the top scored computer-docked 17 also suggested that the terephthalamide scaffold mimics the peptide helical backbone (Fig. 11.16). The docking result showed that the side chains of terephthalamide 17 have an analogous spatial arrangement to the three key side chains of Val74, Leu78, and Ile81 of the Bak peptide. The \(^{15}\text{N}-\text{HSQC} \) chemical shift perturbation experiments have confirmed this binding mode, because the addition of 17 led to peak shifts in a number of residues found near this pocket. Furthermore, the residues affected by terephthalamide 17 and terphenyl 13 are similar (Gly94 of the BH3; Arg102 and Leu108 of the BH2; Arg139, Ile140, and Ala142 of the BH1 domains of Bcl-xL), indicating the terephthalamide is a successful alternative scaffold to the terphenyl as \( \alpha \)-helix mimetics.
A foremost goal in chemical biology and medicinal chemistry is to reduce the structural complexity of biomolecules by designing low-molecular-weight synthetic agents that reproduce their essential features (Mutter and Tuchscherer, 2000). Using synthetic agents to mimic α-helices, in particular, has immense interest in drug discovery because of the central role that the α-helical regions play in many biological processes. The advantages of using synthetic agents to target protein–protein interfaces include resistance to proteolytic degradation, lack of need to prepare large chemical libraries, and higher hit rates in preliminary screening. Although still at its early stage, the approach of rationally designing α-helix mimetics has made major strides in the last decade, substantially contributing to our understanding of fundamental principles underlying biological processes and serving as a first step in accessing molecules of therapeutic relevance. Selected examples of recent studies (most of which are based on shape and electrostatic complementarity) in this interdisciplinary research field have been reviewed. In view of the expected exponential growth of knowledge about biomolecular recognition along with improved technology in computer-aided simulations in the coming years (Jorgensen, 2004), the design of therapeutically relevant molecules with higher affinity and selectivity are within reach.

ACKNOWLEDGMENTS

We thank the National Institutes of Health (GM35208 and GM 69850) for financial support of this work.
REFERENCES

12.1 INTRODUCTION

Chemokine (chemoattractant cytokine) receptors are a group of membrane proteins that belong to the superfamily of G-protein-coupled receptors (GPCRs). They possess seven transmembrane (TM) helices and transmit signals from extracellular ligands to

*Corresponding author: ziweihuang@burnham.org

intracellular biological pathways via heterotrimeric G-proteins (Kobilka, 1992; Murphy, 1994; Strader et al., 1994). As the natural ligands of chemokine receptors, chemokines are small soluble proteins of about 70 amino acid residues with a molecular weight of 8–10 kDa. They play prominent roles in leukocyte activation and trafficking to sites of inflammation by interacting with chemokine receptors. Based on the positions of two conserved cysteine residues in their amino (N)-termini, chemokines can be divided into four subfamilies: CC, CXC, CX3C, and C (Berger et al., 1999; Murphy, 1994). The two main subfamilies of chemokines are CXC and CC (Table 12.1). They are important for the selective activation and recruitment of a large variety of cell types in inflammation. CXC chemokines are primarily involved in the activation of neutrophils, whereas CC chemokines do not affect neutrophils and generally stimulate other leukocytes such as monocytes, lymphocytes, and basophils. In addition to important roles in many physiological processes, chemokines are implicated in a wide range of human acute and chronic inflammatory diseases (i.e., acute respiratory distress syndrome, allergic asthma, psoriasis, and arthritis), neurological disorders, cancer, and most notably acquired immune deficiency syndrome (AIDS) (Baggiolini et al., 1997; Berger et al., 1999; Murphy, 1994; Proudfoot, 2002) (Table 12.2).

### 12.2 CHEMOKINE RECEPTORS IN HIV-1 ENTRY

In recent years, chemokine receptors have been found to be involved in the pathogenesis of human immunodeficiency virus type 1 (HIV-1) infection. HIV-1 enters cells through a fusion process in which the HIV-1 envelope glycoprotein gp120 binds to CD4, the main receptor for HIV-1 on the target cell surface. However, it has long been known that CD4 alone is not sufficient for HIV-1 fusion and entry, and that additional receptors may be needed (Clapham et al., 1991; Maddon et al., 1986). In 1996, chemokine receptors CXCR4 and CCR5 were discovered to be the long-sought coreceptors for syncytium-inducing and non-syncytium-inducing HIV-1 strains, respectively (Alkhatib et al., 1996; Deng et al., 1996; Dragic et al., 1996; Feng et al., 1996). While all HIV-1 strains appear to require CXCR4, CCR5, or both
(Simmons et al., 1996; Zhang et al., 1996), some strains can also use other chemokine receptors, such as CCR3 and CCR2b, as coreceptors for fusion and infection (Choe et al., 1996; Doranz et al., 1996). The viral fusion process may involve the initial binding of HIV-1 gp120 to its high-affinity receptor CD4, which results in conformational changes in gp120 and CD4 (Clements et al., 1991; Gershoni et al., 1993; Sattentau et al., 1993). The gp120–CD4 complex then interacts with a chemokine coreceptor such as CXCR4 or CCR5 to form a heterotrimeric complex of gp120–CD4–coreceptor (Lapham et al., 1996; Trkola et al., 1996; Wu et al., 1996) (Fig. 12.1).

During the asymptomatic stage of disease, macrophage (M)-tropic strains of HIV-1 involved in sexual transmission primarily use CCR5 as an entry coreceptor

<table>
<thead>
<tr>
<th>Chemokine Receptors</th>
<th>Associated Diseases</th>
</tr>
</thead>
<tbody>
<tr>
<td>CXCR1</td>
<td>Sepsis, atherosclerosis</td>
</tr>
<tr>
<td>CXCR2</td>
<td>Sepsis, atherosclerosis</td>
</tr>
<tr>
<td>CXCR3</td>
<td>Multiple sclerosis, rheumatoid arthritis, transplant</td>
</tr>
<tr>
<td>CXCR4</td>
<td>AIDS, cancer</td>
</tr>
<tr>
<td>CCR1</td>
<td>Multiple sclerosis, rheumatoid arthritis, transplant, asthma, nephritis</td>
</tr>
<tr>
<td>CCR2</td>
<td>Multiple sclerosis, rheumatoid arthritis, transplant, asthma, atherosclerosis</td>
</tr>
<tr>
<td>CCR3</td>
<td>Asthma</td>
</tr>
<tr>
<td>CCR4</td>
<td>Asthma</td>
</tr>
<tr>
<td>CCR5</td>
<td>AIDS, multiple sclerosis, rheumatoid arthritis, transplant, asthma, nephritis, inflammatory bowel disease</td>
</tr>
</tbody>
</table>

**Figure 12.1** (a) The interactions of HIV-1 gp120, CD4, and a chemokine coreceptor, such as CXCR4 or CCR5, leading to the cell entry of the virus. (b) The binding of a chemokine ligand (such as SDF-1α) to its receptor (such as CXCR4) blocks the association of HIV-1 gp120/CD4 with CXCR4, thus preventing the viral entry. On the other hand, the chemokine receptor–ligand interaction can activate normal signaling and biological functions (such as chemotaxis).
(Alkhatib et al., 1996; Deng et al., 1996; Dragic et al., 1996). The crucial importance of CCR5 in HIV-1 transmission and feasibility of CCR5-specific agents to inhibit viral infection were further demonstrated by the observation that individuals with CCR5 mutations appear to be both healthy and highly resistant to HIV-1 infection (Samson et al., 1996). However, in 40–50% of HIV-1-infected individuals, T-cell (T)-tropic strains that predominantly use CXCR4 eventually replace M-tropic strains, leading to rapid disease progression (Cheng-Mayer et al., 1988; Schellekens et al., 1992; Tersmette et al., 1989). CXCR4 can also be used by human immunodeficiency virus type 2 (HIV-2) strains adapted to replication in CD4-negative cell lines (Endres et al., 1996). Dual-tropic strains are those HIV-1 isolates that are capable of using both CXCR4 and CCR5. In addition, an HIV-1 envelope can bind CXCR4 independently, and this interaction is enhanced by the presence of CD4 (Bandres et al., 1998).

Two theories have been proposed for the mechanism by which natural chemokines of CXCR4 or CCR5 can prevent chemokine receptor-dependent HIV-1 entry. First, natural chemokines of CXCR4 or CCR5 can inhibit HIV-1 infection (Bleul et al., 1996a; Oberlin et al., 1996) by blocking the binding of HIV-1 gp120 to CXCR4 or CCR5. The mere occupancy of HIV-1 coreceptors by chemokines even in the absence of G-protein-mediated signaling is known to be sufficient for inhibition of HIV-1 infection (Arenzana-Seisdedos et al., 1996; Oravecz et al., 1996). Alternatively, natural chemokines can inhibit HIV-1 entry by inducing receptor downregulation from the cell surface, thereby removing the essential coreceptors (Amara et al., 1997; Förster et al., 1998). In fact, these studies suggested that coreceptor internalization contributes to efficient chemokine inhibition of HIV-1 entry.

The discovery of distinct chemokine coreceptors for different HIV-1 strains has important implications for understanding HIV-1 pathogenesis and developing new therapeutic strategies. Among many new directions in AIDS research opened by this discovery, an important area of investigation is the biochemical and biophysical characterization of the interactions of chemokine receptors with HIV-1 and natural as well as synthetic chemokine ligands. The understanding of the precise structure–function relationship of chemokine receptors and ligands will be essential for providing the structural and biochemical basis to address basic questions of (a) mechanisms of HIV-1 viral entry, host cell tropism, and resistance to infection; (b) mechanisms of receptor recognition and signal transduction by natural chemokines and synthetic mimics and their role in the normal biological functions; (c) any common and/or distinct regions/sites on the receptors for HIV-1 and chemokine recognition that are important for devising selective intervening strategies; and (d) any general insights into the structure–function and mechanism of other membrane proteins that belong to the G-protein-coupled receptor (GPCR) superfamily.

To define precise regions involved in the molecular interactions of chemokine receptors with their natural ligands and HIV-1 gp120, an approach combining structural biology, molecular biology, and synthetic chemistry is being used. As described in the following section, molecular modeling techniques are applied to propose plausible structural models for chemokine receptors and their complexes with ligands. Such models reveal potential sites of interaction between a receptor and a ligand. To experimentally test the putative functional roles of such sites, peptide analogues targeting the functional epitopes on chemokine ligands are synthesized and
tested in biological assays. In addition, site-directed mutagenesis of the chemokine receptor is carried out to test the predicted sites for ligand binding and signaling. These studies provide structural and chemical insights into receptor–ligand interactions and form the basis for the design of small molecular agents that can prevent HIV-1 entry.

### 12.3 STRUCTURE–FUNCTION RELATIONSHIP OF CHEMOKINE RECEPTORS

The information about the detailed three-dimensional structure of a chemokine receptor and its complex with a ligand is critical for understanding the role and mechanisms of the function of chemokine receptors and ligands in Human immunodeficiency virus type 1 (HIV-1) pathogenesis. However, no crystal structure is available for chemokine receptors or any other G-protein-coupled receptor (GPCR) proteins (Kobilka, 1992; Strader et al., 1994). Since the structure of the seven transmembrane (TM) helices in a similar protein, bacteriorhodopsin (bR), has been determined by electron microscopy (Henderson et al., 1990), many members of the GPCR superfamily have been modeled by computational techniques based on the structure of the TM segments of bR (Kontoyianni and Lybrand, 1993; Strader et al., 1994). Our laboratory extended the computational methodologies used for modeling other GPCR proteins to the family of chemokine receptors. A general approach for modeling the structure and interaction of chemokine receptors has been developed from our previously published study of interleukin-8 receptor type β, a member of the chemokine receptor family (Luo et al., 1997).

As shown in Fig. 12.2, the structure for the TM helices of a chemokine receptor is constructed based on that of bR. To further define the three-dimensional structure of the extracellular segments of the chemokine receptor and understand how they interact with a ligand, high-temperature molecular dynamics simulations are performed. The three-dimensional structure of the ligand, as determined by NMR experiment, is used as the geometric constraint in dynamics simulations. A nanosecond simulation is conducted for the receptor–ligand complex, and 500 structures are extracted from the trajectory of the simulation. These structures are clustered by the similarity in their backbone structures. Plausible structural models for the receptor–ligand complex are selected based on the examination of available experimental data and binding energy of the complex. Using this approach, we proposed a hypothetical structure of CCR5 in complex with a chemokine ligand such as macrophage inflammatory protein (MIP)-1β (Fig. 12.3a), which was useful for guiding subsequent site-directed mutagenesis studies of CCR5 to identify some key residues on CCR5 important for ligand binding and signal transduction (Zhou et al., 2000a). In a related study, the structure of CXCR4 in complex with stromal cell-derived factor (SDF)-1α was also proposed by using a similar method (Fig. 12.3b).

In conjunction with molecular modeling, chimeras and site-specific mutant studies were performed to study the role of CXCR4 and CCR5 in HIV-1 entry. These studies demonstrated that the amino (N)-terminus, the second extracellular loop (ECL2), and the third extracellular loop (ECL3) of CXCR4 (Brelot et al., 1997,
Identity the transmembrane helices in the sequence of the chemokine receptor

Modeling of the three-dimensional arrangement of the transmembrane helices based on the structure of bR

Generate the structure of extracellular segments of the chemokine receptor using high-temperature molecular dynamics

Computer modeling of the binding interaction of a chemokine ligand with the chemokine receptor

Analysis of binding sites on both the chemokine receptor and ligand for the obtained structure of the receptor-ligand complex

Figure 12.2 The general procedure for the modeling studies of a chemokine receptor–ligand complex.

2000; Chabot and Broder, 2000; Chabot et al., 1999; Doranz et al., 1999; Kajumo et al., 2000; Lu et al., 1997; Picard et al., 1997a,b; Wang et al., 1998; Zhou et al., 2001) or CCR5 (Alkhatib et al., 1997; Atchison et al., 1996; Bieniasz et al., 1997; Gosling et al., 1997; Rucker et al., 1996; Zhou et al., 2000a) are required for HIV-1 coreceptor activity. For instance, the previous structure–function studies of CXCR4 showed that there is a significant overlap between HIV-1 and chemokine functional sites on the extracellular regions of CXCR4 (Brelot et al., 1997, 2000; Chabot et al., 1999; Doranz et al., 1999; Kajumo et al., 2000; Zhou et al., 2001). Although there have been some discrepancies among these studies, the extracellular regions of the coreceptor have been the focus, with each extracellular region being implicated in coreceptor function. These studies also indicated that multiple extracellular domains of CXCR4 are required for chemokine interactions and receptor signaling (Alkhatib et al., 1997; Atchison et al., 1996; Bieniasz et al., 1997; Brelot et al., 1997; Chabot and Broder, 2000; Doranz et al., 1997; Lu et al., 1997; Rucker et al., 1996). These chimeric and mutational studies were used to validate the accuracy of a two-site model that was initially developed for the C5a chemoattractant and its receptor (Siciliano et al., 1994). According to the model, the motif composed of amino acids 12–17 of SDF-1 α, RFFESH loop, first docks on the N-terminal domain of CXCR4, and this contact allows the subsequent interaction of the flexible N-terminus of SDF-1 α with the receptor groove formed by TM domains and/or extracellular loops, thus triggering the receptor function (Crump et al., 1997; Gupta et al., 2001; Siciliano et al., 1994). The N-terminus of SDF-1 α, being relatively flexible and unstructured in solution.
(Crump et al., 1997), has been proven to be essential for CXCR4 recognition and signal transduction (Crump et al., 1997; Dealwis et al., 1998; Heveker et al., 1998). Taken together, these findings have provided structural and chemical insights into receptor–ligand interactions and formed the basis for the design of small-molecule agents that prevent HIV-1 entry.

Despite the previous studies, a number of areas remained largely unexplored, for which further investigations using structural modeling and site-directed mutagenesis of the receptors are critical, such as (a) structural and functional roles of other residues, especially those near the important residues already identified in above studies; (b) structural and functional roles of residues and sites located near or within the TM helices that have not been studied, yet may play an important role in signaling and controlling the function of extracellular receptor surfaces; (c) the mechanism of action of natural chemokines needs to be further elucidated in terms of their binding and signaling, whereas many of previous studies focused on HIV-1 interaction; and (d) novel synthetic agonistic or antagonistic ligands developed by us and others, whose specific binding sites on CXCR4 or CCR5 are unclear, require thorough characterization in order to understand their mechanism of action and develop them into more potent and specific therapeutics.

For instance, it has yet to be determined whether residues and sites located near or within the TM helices of CXCR4 might play any structural and functional roles for ligand interactions and, if they do, whether such roles are differentiating for different ligands (i.e., HIV-1 gp120 and SDF-1α). Also, based on our previous report that three residues in the ECL2—Arg^{183}, Tyr^{184}, and Tyr^{190}—are not involved in SDF-1α binding but play an important role in HIV-1 coreceptor activity (Zhou et al., 2001), we hypothesized that CXCR4–gp120 interaction might be potentially disrupted without impairing its natural chemokine receptor activity. To address these questions, we constructed a panel of mutations at residues near or within the TM helices and several mutants of the ECL2 residues to investigate the role of TM and ECL2 in HIV-1 entry versus natural ligand binding and signaling.
According to this study (Tian et al., 2005), many residues throughout CXCR4 TM and ECL2 domains are specifically involved in interaction with HIV-1 gp120, because most of these sites did not play a role in either SDF-1α binding or signaling (Fig. 12.4). As many of these distinct sites for either HIV-1 or SDF-1α function are located on the upper part of the TM barrel close to the extracellular side or in the extracellular loops (ECL1 and ECL2).

**Figure 12.4** Distinct functional sites for HIV-1 gp120 and SDF-1α on a hypothetical structural model of CXCR4. The TM residues, Phe87 and Phe292, required for SDF-1α binding only are highlighted in lighter color and represented as ball-and-stick with an asterisk. Asp171, the only residue found to be involved in both HIV-1 gp120 and SDF-1α interactions, is highlighted in the lighter color with an asterisk. The TM residues that are involved in HIV-1 coreceptor activity only are highlighted in the darker color. Only the TM domains of CXCR4, with side (A) and top (B) views, are shown for simplicity. (C) Schematic illustration of the locations of residues found to be important for HIV-1 or SDF-1α on CXCR4 TM and ECL2 domains. The residues required for SDF-1α binding are shown as white spots, while those that are involved in HIV-1 coreceptor activity are shown as black spots. Asp171 is highlighted as a white spot. See color plates.
ECL2, their role is likely to be involved with direct interactions with different ligands. On the other hand, some residues predicted to be located deep inside the TM barrel and thus unlikely to contact gp120 directly were also able to reduce the coreceptor activity of CXCR4 without impairing the binding and signaling activities of SDF-1α. It is possible that these mutations presumably change the conformations of CXCR4 that are important for HIV-1 but not SDF-1α. Further work is needed to understand the mechanism of how the function of HIV-1 or SDF-1α is mediated differently by distinct sets of residues and possibly by conformational states. Nevertheless, the findings from the present study provide a basis for the development of new inhibitory agents that modulate the functional sites or conformations of CXCR4 for the purpose of reducing or avoiding the limitations and side effects caused by nonselective inhibitors of this important coreceptor.

The important functional roles of TM residues in CXCR4 are consistent with other lines of evidence previously reported in other GPCRs. For instance, studies of human neurokinin peptide receptor (NK1R) (Huang et al., 1995) and melanin-concentrating hormone (MCH) receptor (Macdonald et al., 2000) suggested that some residues in the TM regions contribute to ligand binding and activation. A recent study of CCR5 chimera also suggested that the N-termini of MIP-1α and RANTES (regulated on activation, normal T-cell expressed and secreted) mediate receptor activation by interacting with the TM helix bundle (Blanpain et al., 2003). Furthermore, according to a two-site model of the CXCR4-SDF-1α interaction, the motif consisting of amino acids 12–17 of SDF-1α first docks on the N-terminal domain of CXCR4, which allows its flexible N-terminus to interact with the receptor groove formed by TM domains and/or extracellular loops of CXCR4, thus triggering the receptor function (Crump et al., 1997; Gupta et al., 2001). When these results and models published by others are combined with the findings from the present study, one can hypothesize that the N-terminus of SDF-1α or certain flexible functional determinants of gp120 may readily interact with some residues in the TM domains of CXCR4. The interaction of the TM domains with HIV-1 gp120 should involve an extensive set of TM residues that are different from those residues required for the interaction with SDF-1α.

12.4 STRUCTURE–FUNCTION RELATIONSHIP OF CHEMOKINE LIGANDS

The only known natural ligand of CXCR4 is the stromal cell-derived factor (SDF)-1α, a CXC chemokine that plays critical roles in the migration, proliferation, and differentiation of leukocytes (Bleul et al., 1996b; Oberlin et al., 1996). CXCR4 can also be recognized by an antagonistic ligand, the viral macrophage inflammatory protein (vMIP)-II encoded by the Kaposi’s sarcoma-associated herpes virus (Moore et al., 1996). vMIP-II displays a broader spectrum of receptor activities than any mammalian chemokine, because it binds with high affinity to a number of both CXC and CC chemokine receptors, including CXCR4, CCR5, and CCR2, and inhibits cell entry of HIV-1 mediated by these receptors (Boshoff et al., 1997; Kledal et al., 1997). Studies with knockout mice of CXCR4 and SDF-1α have demonstrated that
these molecules play important roles in immunomodulation, organogenesis, hematopoiensis, and derailed cerebellar neuron migration (Ma et al., 1998; Nagasawa et al., 1996; Zou et al., 1998). The CC chemokine ligands of CCR5 include macrophage inflammatory protein (MIP)-1β, MIP-1α, and RANTES (regulated on activation, normal T-cell expressed and secreted). Among these CCR5 ligands, RANTES and MIP-1α can bind to other CC chemokine receptors, while MIP-1β is known to be most specific for CCR5 (Wells et al., 1996). The structures of several chemokines that bind CXCR4 or CCR5 have been determined by NMR or X-ray technique, including those of SDF-1α (Crump et al., 1997; Dealwis et al., 1998), vMIP-II (Fernandez et al., 2000; Liwang et al., 1999), MIP-1β (Lodi et al., 1994), and RANTES (Chung et al., 1995). Based on these structures, the structure–function relationship of these chemokines has been investigated mainly by a chemical approach of synthetic chemokine mutants and peptides.

The studies of the structures of chemokine receptors and their interactions with chemokine ligands and gp120 are critical for the development of new therapeutic agents that block the interaction between the HIV-1 envelope and chemokine receptors. Several CCR5 receptor antagonists have been reported to block HIV-1 infection (Arenzana-Seisdedos et al., 1996; Nardese et al., 2001; Simmons et al., 1997). For instance, aminoxypentane (AOP)-RANTES, a chemically modified analogue of RANTES, is a CCR5 antagonist that is effective in inhibiting CCR5-dependent HIV-1 strains (Simmons et al., 1997). However, since RANTES binds to a number of chemokine receptors in addition to CCR5, treatment with RANTES analogues that block the natural chemokines from functioning in these receptors could lead to potential problems. The development of more specific inhibitors of CCR5 clearly depends on further studies of chemokine ligand binding to CCR5 and other receptors. Regardless, coreceptor-based therapeutic strategies (Amara et al., 1997; Donzella et al., 1998; Doranz et al., 1997; Fenard et al., 2001; Heveker et al., 1998; Huang, 2000; Loetscher et al., 1998; Luo et al., 1999a,b, 2000; Murakami et al., 1997; Sachpatzidis et al., 2003; Schols et al., 1997; Zhou et al., 2000b, 2002) have the appeal of targeting relatively invariant host determinants, in contrast with anti-HIV-1 agents directed against components of the rapidly mutating virus population. The feasibility of this approach to inhibit viral infection is further demonstrated by the observation that individuals with CCR5 mutations appear to be both healthy and highly resistant to HIV-1 infection (Berger et al., 1999; Samson et al., 1996).

As for CXCR4, several groups reported that peptides and organic compounds (Donzella et al., 1998; Doranz et al., 1997; Fenard et al., 2001; Sachpatzidis et al., 2003; Zhou et al., 2000b, 2002), which were discovered much earlier through random screening and are unrelated to natural chemokines, have antagonistic activity in CXCR4. For instance, they include ALX40-4C (Doranz et al., 1997), T22 (Murakami et al., 1997), and AMD3100 (Schols et al., 1997). However, due to their high positive charge and unclear mechanism of action, the CXCR4 specificity and side effects of these molecules are still uncertain. Furthermore, there is cause for concern regarding undesired side effects of blocking the normal CXCR4-SDF-1α function, since knockout mice lacking either CXCR4 (Tachibana et al., 1998; Zou et al., 1998) or SDF-1α (Nagasawa et al., 1996) die during embryogenesis, with
evidence of hematopoietic, cardiac, vascular, and cerebellar defects. As a result, it would be desirable to develop compounds that can target specific regions of CXCR4 that are selective for HIV-1 coreceptor function only, but not the normal function of SDF-1α. In fact, AMD3100 and ALX40-4C have entered clinical trials, but further development of AMD3100 was discontinued due to cardiac toxicity (Doranz et al., 2001). Even if adverse effects are not observed during clinical trials, any CXCR4 antagonist has a potential to do just that if it would be used chronically as highly active antiretroviral therapy.

An alternative route to designing specific CXCR4 inhibitors is to use natural chemokine ligands as design templates (Heveker et al., 1998; Huang, 2000; LaBarge and Blau, 2002; Loetscher et al., 1998; Luo et al., 1999a,b, 2000; Sachpatzidis et al., 2003; Zhou et al., 2000b, 2002) by mimicking specific regions of a chemokine ligand in order to study the structure–function relationship of the native molecule and to develop novel agonists or antagonists of chemokine receptors. This approach was first attempted on SDF-1α, the only known natural ligand of CXCR4 (Heveker et al., 1998; Loetscher et al., 1998; Luo et al., 1999a,b). The structure of SDF-1α consists of three major domains: the disordered amino (N)-terminus with mostly extended conformation, a central positively charged core region of three antiparallel β-strands, and the overlying carboxyl (C)-terminal α-helix (Crump et al., 1997; Dealwis et al., 1998). Peptides derived from the N-terminus of the SDF-1α sequence were proven to be essential for CXCR4 recognition, signal transduction, and antiviral activity, but they were less potent than native SDF-1α (Heveker et al., 1998; Loetscher et al., 1998). The exchange of K1 and P2 with other amino acids resulted in the generation of potent CXCR4 antagonists (Heveker et al., 1998). The functional role of the C-terminus helix of SDF-1α, which by itself had no activity in receptor binding and signaling, was also studied by joining the N- (residues 1–4) and C-terminus (residues 56–67) with a four-glycine linker (Luo et al., 1999a). The attachment of the C-terminus dramatically increased the effect of the N-terminal fragment by inducing chemotaxis and intracellular Ca2+ influx, but its activity was completely abolished by the addition of heparin, suggesting that the C-terminal fragment of the peptide binds glycosaminoglycans (GAGs) (Luo et al., 1999a). These findings demonstrated a way to enhance biological activity of CXCR4 antagonists, possibly by attaching a GAG-binding motif (Luo et al., 1999a). Furthermore, the experimental support for the role of the highly positive β-sheet region of SDF-1α in interacting with negatively charged residues in the extracellular domains of CXCR4 is also available (Luo et al., 1999b). The attachment of positively charged residues to the N-terminal peptide sequence of SDF-1α was found to enhance the ability of the peptides in CXCR4 binding and inhibiting CXCR4-mediated T-cell (T)-tropic HIV-1 entry (Luo et al., 1999b). Taken together with the observation that peptides and organic compounds, such as AMD3100 (Schols et al., 1997), T22 (Murakami et al., 1997), and ALX40-4C (Doranz et al., 1997), have high positive charges and affinity for CXCR4, the study indicates that the electrostatic interaction may play a role in CXCR4 recognition (Luo et al., 1999b).

CXCR4 can also be recognized by vMIP-II, which displays a broader spectrum of receptor activities than any mammalian chemokine, because it binds with high affinity to a number of both CC and CXC chemokine receptors (Boshoff et al., 1997;
Kledal et al., 1997). The overall topology of vMIP-II is very similar to that of SDF-1α; it consists of a disordered N-terminus, followed by a helical turn, which leads into the first strand of a three-stranded antiparallel β-sheet. Following the sheet is a C-terminal helix (Liwan et al., 1999; Sozzani et al., 1998). Recently, synthetic peptides derived from the N-terminus of vMIP-II were studied by our laboratory to elucidate the mechanism of vMIP-II interaction with its coreceptors. It was discovered that the N-terminus of vMIP-II is the major determinant for CXCR4 recognition (Luo et al., 2000). Only V1 peptide (1–21 residues) from the N-terminus of vMIP-II showed CXCR4 binding, whereas all other peptides derived from other regions of vMIP-II did not have any activity in CXCR4. On the other hand, none of these peptides showed binding activity for CCR5. Further studies showed that V1 selectively prevents CXCR4 signal transduction and coreceptor function in mediating the entry of T- and dual-tropic HIV-1 isolates (Luo et al., 2000).

Realizing that modified analogues based on the V1 peptide may lead to a new class of compounds with higher potency and stability, an all-D-amino acid analogue of V1 peptide, designated as DV1 peptide, was also synthesized (Zhou et al., 2002). Despite dramatically different conformations of side-chain groups as a result of being a mirror image of each other, DV1 displayed higher binding affinity toward CXCR4 than V1 (Zhou et al., 2002). Additional peptides composed of mixed D- and L-amino acids showed CXCR4 binding, further demonstrating the remarkable stereochemical flexibility of the CXCR4–peptide interface (Zhou et al., 2002). In addition, acting as CXCR4 antagonists, DV1 showed significant activity in inhibiting the replication of CXCR4-dependent HIV-1 strains (Zhou et al., 2002). Unnatural D-peptides can be highly desirable and advantageous over natural L-peptides for therapeutic development, because they are highly stable and resistant to proteolytic degradation (Zhou et al., 2002). As to CCR5 binding determinants, from studies of vMIP-II we hypothesized that CCR5 binding has a strict structural requirement and involves multiple sites other than the N-terminus of vMIP-II. This is consistent with the more recent report from another laboratory that peptides from the N-loop and first β-strand of RANTES showed CCR5 binding, and that this binding depends on a dimeric structure (Nardese et al., 2001).

More recently, Sachpatzidis et al., (2003), understanding the potential adverse effects from the use of antagonistic inhibitors of CXCR4 in HIV-1 therapy, identified RSVM and ASLW as novel allosteric agonists that are insensitive to the CXCR4 antagonist AMD3100 by screening a semi-randomized 17-mer library in a yeast strain expressing a functional CXCR4 receptor (Sachpatzidis et al., 2003). In chemotaxis assays, RSVM behaves as a partial agonist, while ASLW behaves as a superagonist, displaying a chemotactic index that is greater than the maximum observed with SDF-1α (Sachpatzidis et al., 2003). Allosteric agonists may be therapeutically useful in combination with small-molecule antagonists for anti-HIV therapy since they could maintain essential receptor function. The data also illustrated that other binding sites may exist for nonphysiological agonists (Sachpatzidis et al., 2003). Despite their early stage of development, allosteric agonists and other potent CXCR4 peptide antagonist or agonists discussed previously could serve as new leads for the development of new therapeutic agents for HIV-1 infection and other immune system diseases.
12.5 IMPORTANCE OF CHEMOKINES AND CHEMOKINE RECEPTORS IN OTHER PATHOLOGICAL PROCESSES

12.5.1 Chemokine Receptors and HIV-Associated Dementia

HIV-associated dementia (HAD) is a cognitive and motor dysfunction observed after infection with HIV-1 (Kaul et al., 2001). Although about one-half of children and one-fourth of adults infected with HIV-1 eventually develop the dementia (Lipton and Gendelman, 1995; McArthur et al., 2003), there is no specific treatment for HAD. In recent years, chemokine receptors have been found to play important roles in the pathogenesis of HIV-1 infection, including HAD. In a plausible model, HIV-1 enters target cells through a direct fusion process in which HIV-1 gp120 binds to CD4, the main receptor on the target cell surface, and to one of the two chemokine receptors, usually CXCR4 or CCR5 (Berger et al., 1999; Poignard et al., 2001). Natural chemokines that bind CXCR4 and CCR5 can inhibit HIV-1 infection (Bleul et al., 1996a; Oberlin et al., 1996), probably by blocking common binding sites on the chemokine receptors that are also required for gp120 interaction and/or by inducing receptor internalization (Amara et al., 1997; Förster et al., 1998). Although CXCR4 is expressed in virtually all tissues, including neurons, microglia, and astrocytes in the brain (McManus et al., 2000; Sanders et al., 1998), CCR5 is the primary receptor by which the cells in the nervous system get infected (He et al., 1997; Rottman et al., 1997).

There are two major mechanisms by which gp120-induced neuronal apoptosis occurs. The predominant pathway is an indirect process that results in the release of toxic factors from infected cells such as macrophages and microglia. HIV-1 can stimulate or infect microglia and macrophages by interacting with the chemokine receptors in conjunction with CD4. Macrophages and microglia can also be activated by factors, such as shed gp120, released from infected cells (Gartner, 2000; Kaul et al., 2001). The other mechanism is a direct interaction of gp120 with neurons via the chemokine receptors (Meucci et al., 1998). In fact, picomolar concentrations of CXCR4 (or X4)-preferring or CXCR4/CCR5 (or dual)-tropic gp120 are known to induce neurotoxicity through CXCR4 despite the fact that neurons do not express CD4 on the cell surface (Hesselgesser et al., 1998; Kaul and Lipton, 1999; Ohagen et al., 1999). However, the previous studies demonstrating that gp120-induced neuronal damage in rodent cultures was prevented by anti-gp120 antibodies but not by anti-CD4 antibodies implied that CD4 is not necessary for neurotoxicity (Dreyer et al., 1990; Kaiser et al., 1990). Whether or not neuronal apoptosis occurs via an indirect or a direct pathway, the chemokine receptors are involved in the development of HAD, which suggests that therapeutic interventions at the level of chemokine receptors may reverse or attenuate the symptoms associated with HAD.

12.5.2 Chemokine Receptors and Multiple Sclerosis

The etiology of multiple sclerosis (MS) remains a mystery and is passionately debated by clinical investigators. It is believed that MS pathology results from the
transmigration and accumulation of leukocytes in the central nervous system (CNS) that consequently attack glia cells. Chemokines are involved in the pathogenesis of immune-mediated inflammation of the CNS, both in controlling leukocyte migration across the brain endothelium and in the activation and movement of cells within the brain parenchyma. The biological activities of chemokines are mediated by interactions with their corresponding chemokine receptors (Zlotnik, 2000). The ß-chemokine receptors CCR3 and CCR5 have been detected in the normal control human CNS tissue associated with microglia (Xia et al., 1998) and are expressed by cultured microglia (Bonecchi et al., 1998). Investigations have shown that glial cells within the CNS have the capacity to express both chemokines and chemokine receptors during a variety of inflammatory and degenerative conditions (Simpson et al., 1998; Xia et al., 1998). Increased expressions of CCR3 and CCR5 have been detected in the CNS and their ligands in MS (Simpson et al., 1998). CCR2 has also been identified on mononuclear phagocytes. These findings imply that ß-chemokines, such as RANTES, monocyte chemoattractant protein (MCP)-1α, -2, -3, and -4, and macrophage inflammatory protein (MIP)-1α and MIP-1β, could all potentially activate and recruit both resident glia and infiltration haematogenous cells to sites of CNS inflammation. Cytokine-stimulated glial cells have also been shown to express chemokines and chemokine receptors in vitro (Jiang et al., 1998).

MS is an autoimmune disease of the CNS characterized by relapse and remittance of clinical signs and the presence of inflammation and demyelination in the white matter. Accumulating evidence has suggested that these findings represent the first stage of the disease followed by a chronic irreversible phase due to degeneration of the myelin sheath and the underlying axon. As the suppression of the relapse during the course of MS with immunotherapy is extremely important to prevent the development of the subsequent irreversible stage of the disease, it is essential to clarify the mechanism of disease relapse using an animal model, that is, Encephalomyelitis (EAE). EAE is a T-cell-mediated inflammatory demyelinating disease of the CNS that functions as a mouse model for human MS. Upon immunization of susceptible mice with myelin proteins (McRae et al., 1992; Trotter et al., 1987), EAE neuroinflammation follows with the subsequent invasion of leukocytes into the CNS through the blood–brain barrier. Leukocytes in the EAE-affected brain derived primarily from infiltrating the blood monocytes and lymphocytes actively mediate tissue destruction in EAE by destroying the myelin sheath, thus leading to neurological disease similar to MS. MCP-1, a natural chemokine, has been identified not only in MS, but also in EAE (Miyagishi et al., 1997). CCR2, a receptor for MCP-1, is a seven-transmembrane spanning protein that is functionally linked to downstream signaling pathways through heterotrimeric G-proteins (Charo, 1999). In MS, CCR2 is expressed on microglia and macrophages in chronic active lesions and on perivascular mononuclear cells in both white matter lesions and unaffected cortex (Simpson et al., 2000). Furthermore, CCR2 expression on circulating CD4 T-cells is significantly elevated during MS relapse (Misu et al., 2001; Sorensen, 2001). In an animal model of MS, CCR2 knockout mice showed reduced mononuclear infiltration and decreased susceptibility to EAE (Fife et al., 2000; Izikson et al., 2000). Studies have been conducted in MS lesions, infiltrating cells, and cerebrospinal fluid (CSF) and the interrelationship among the levels of chemokines in CSF, expression of
chemokine receptors on lymphocytes, MS activity, and the clinical character of MS have been examined (Balashov et al., 1999; Mahad and Woodroofe, 2002; Sorensen et al., 2002). The elevated expression of CCR2 in the CNS suggests that this β-chemokine receptor and its ligand MCP-1 play a major role in the pathogenesis of MS (Simpson, 2000).

Of particular interest are (a) the CCR2 receptor expressed on peripheral blood monocytes and activated T-cells (Gu L, 1999) and (b) MCP-1 secreted by a variety of cell types in response to proinflammatory stimuli. Mahad and Woodroofe (2002) indicated that expression and cellular localization of CCL2/MCP-1 and CCR2 in MS have been described in the brain, cerebrospinal fluid (CSF), and blood. Evidence from descriptive, transgenic, knockout, and neutralizing studies of EAE points toward a nonredundant role of CCL2 and CCR2 in the recruitment of inflammatory infiltrate into the CNS. Hence, CCL2 and CCR2 may be targets for specific and effective treatment of MS (Mahad and Woodroofe, 2002). Some evidence has suggested that MCP-1 and CCR2 are both essential for the induction and progression of EAE and has demonstrated that MCP-1 mRNA is significantly upregulated in EAE during the acute phase of the disease (Juedes et al., 2000; Kennedy et al., 1998). Vaccination of mice with naked MCP-1 DNA, resulting in the generation of anti-MCP-1 antibodies, inhibited the induction of EAE (Youssef et al., 1998). Moreover, MCP-1 knockout mice exhibited a reduction in clinical and histopathological EAE after immunization with significantly reduced accumulation of macrophages in the CNS (Huang et al., 2001). Furthermore, two independent studies of CCR2 knockout mice by Fife et al. (2000) and Izikson et al. (2000) have found a complete resistance to EAE induction and failure of monocyte/macrophage accumulation in the CNS, suggesting that CCR2 may play a specific nonredundant role in monocyte recruitment, and that the inhibition of the CCR2:MCP-1 axis may ameliorate EAE neuropathology.

12.5.3 Chemokine Receptors and Stem Cell Migration to the Sites of CNS Injury

Migration toward pathology is the first critical step in stem cell engagement during regeneration. Neural stem cells (NSCs), both mouse and human, have a capacity for precise migration to even widespread and distant areas of pathology in a number of experimental models of CNS disease, including at ages where extensive migration has conventionally been deemed to be limited (Imitola et al., Imitola et al., 2003, 2004; Park et al., 2002a,b; Snyder et al., 1997). Because stem cell engagement with a degenerating environment (LaBarge and Blau, 2002; Snyder et al., 1997) is the first critical step in regeneration, realizing the therapeutic promise of the NSC depends in part on understanding the mechanisms underlying its mobilization during injury. In this regard, Imitola et al. previously showed that human NSCs (hNSCs) migrate in vivo toward an infarcted area (a representative CNS injury), where local astrocytes and endothelium upregulate the inflammatory chemoattractant stromal cell-derived factor (SDF)-1α. NSCs express CXCR4. Exposure of hNSCs to SDF-1α and the consequent induction of CXCR4-mediated signaling triggered a series of intracellular processes associated with fundamental aspects of survival, proliferation, and, importantly, migration. These steps include rapid and sustained phosphorylation of
p38MAPK kinase [implicated in regulating cytokine-induced cell migration (Cara et al., 2001; Dechert et al., 2001; Sharma et al., 2003)], phosphorylation of the ribosomal S6 kinase (p90RSK) [implicated in phosphorylation of cytoskeletal molecules involved in neurite outgrowth (Suzuki et al., 2001; Wong et al., 1996)], phosphorylation of c-Jun [a kinase involved in migratory responses (Li et al., 2003)], rapid activation of extracellular response kinase (involved in proliferative responses), and rapid phosphorylation of paxilin [a scaffold molecule critical for migration (Huang et al., 2003)]. These data implicate SDF-1α/CXCR4, representative of the inflammatory milieu characterizing many pathologies, as a pathway that activates NSC molecular programs during injury and suggest that inflammation may be viewed not simply as playing an adverse role, but also as providing stimuli that recruit cells with a regenerative homeostasis-promoting capacity. The detailed molecular mechanisms of how SDF-1α/CXCR4 and potentially other chemokine/receptor pairs mediate neural stem cell migration and repair in CNS pathologies remain to be elucidated.

12.5.4 Chemokine Receptors and Rheumatoid Arthritis

Rheumatoid arthritis (RA) is an autoimmune disease characterized by chronic inflammation and massive infiltration of leukocytes into the inflamed joints. In RA, a large number of CD4+ memory T-cells infiltrate the inflamed synovium (Kohem, 1996; Morimoto et al., 1988; Morita et al., 1998), where they get activated as they express cytokines and activation makers (Cohen et al., 1995; Dolhain et al., 1996; Grom et al., 1996; MacDonald, 1997; Nanki et al., 2000a,b; Smeets et al., 1998). Inflammatory cytokines [including IL-1, IL-6, IFN-γ, and tumor necrosis factor (TNF)-α] and activation markers are known to play central roles in the pathogenesis of RA (Nanki et al., 2000a). Although the development of biological drugs targeting these cytokines has produced promising results in the clinical therapy of RA patients, the design of other drugs that are independent of cytokine’s function is needed for the improvement of RA chemotherapy. Previously, Nanki et al. reported that the memory T-cells express a chemokine receptor CXCR4 at a high level and that the concentration of SDF-1α, an endogenous ligand of CXCR4, is extremely high in the synovium of RA patients. It is also found that SDF-1α stimulates the migration of memory T-cells and inhibits T-cell apoptosis, suggesting that the SDF-1α-CXCR4 interaction plays an important role in the accumulation of T-cells in the RA synovium (Nanki et al., 2000b). Since SDF-1α (Nagasawa and Kishimoto, 1994; Tashiro et al., 1993) is independent of the other inflammatory cytokines in terms of its expression and function, the development of inhibitors that blocks the SDF-1α-CXCR4 interaction might produce promising leads for RA therapy.

Several recent publications support the involvement of SDF-1α-CXCR4 interaction in memory T-cell migration in the inflamed RA synovium. For instance, several CXCR4 antagonists, including T22 (an 18-mer peptide) (Murakami et al., 1997), T140 (a 14-mer) (Tamamura et al., 1998), and FC131 (a cyclic pentapeptide) (Fujii et al., 2003), which have inhibitory activities against the entry of T-cell line-tropic (X4-) HIV-1 into the target cells (Murakami et al., 1997, 1999; Xu et al., 1999), are known to have anti-RA activities. De Clercq’s group also reported that another
CXCR4 antagonist, AMD3100, has both anti-HIV-1 and anti-RA activities (Hendrix et al., 2000; Schols et al., 1997). AMD3100 is a bicyclam derivative first described for its potent activity against HIV-1 infection (De Clercq et al., 1992) and is currently under investigation for clinical applicability in AIDS patients (De Clercq et al., 1994). However, the clinical applications of AMD3100 are limited by the fact that this compound is rapidly cleared from circulation, and that the treatment schedule in mice requires the use of osmotic minipumps (Datema et al., 1996).

12.5.5 Chemokine Receptors and Allergic Asthma

Allergic asthma is often characterized by airway inflammation of different cell types, airway hyper-responsiveness (AHR), mucus production, and a variable airflow (Busse, 2001; O’Byrne and Parameswaran, 2001). The initiating phase is characterized by IgE and mast cells. In the propagation phase, Th2-polarized T-lymphocytes and eosinophils are guided to a chronic inflammatory state by infiltrating the airways. The effector phase is characterized by the production of spasmogenic substances, AHR, and mucus hypersecretion. Anti-cytokine therapy has been explored as a potential therapeutic approach in the treatment of allergic asthma, as the cytokines IL-4, IL-5, and IL-13 are of fundamental importance in its pathophysiology. The initial study using monoclonal antibodies (MAbs) against IL-5 (Leckie et al., 2000) in human asthmatics demonstrated a decrease in the number of eosinophils, though no major change in the AHR was observed. More recent studies showed the importance of CCR3 for the recruitment of eosinophils into the lung tissues and AHR in a murine model of allergic asthma. Ma et al. (2002) showed that epicutaneously allergen-sensitized CCR3 knockout mice were protected from AHR after allergen challenge. This protection was accompanied by the absence of CCR3-positive cells infiltrating the lung, especially eosinophils and intraepithelial mast cells (Humbles et al., 2002; Ma et al., 2002). Eosinophils of the CCR3 knockout mice accumulated along the blood vessels of the lung and migrated through the endothelium (Humbles et al., 2002). However, the cells were not able to transfer the lamina elastica of the venules and failed to migrate into the lung tissues (Humbles et al., 2002). The inflammatory infiltrate in allergic diseases consists of eosinophils, Th2 cells, basophils, and mast cells. All these cells are known to express CCR3 on their surfaces constitutively or upon activation by cytokines (Ponath et al., 1996; Sallusto et al., 1997). Among these cells, eosinophils express the highest amount of CCR3, which is the dominant chemokine receptor on their surfaces (Elsner et al., 2000). Several ligands have been described for CCR3, and they include eotaxin, eotaxin-2, eotaxin-3, RANTES, MCPs, and HCC-1. In addition to the induction of actin polymerization, chemotaxis, and intracellular calcium flux in eosinophils via CCR3, they also release toxic ROS upon activation of CCR3. After binding of eotaxin or related chemokines to CCR3, the receptor is internalized. Receptor recycling seems to be important for the activation of leukocytes by chemokines. As such, blocking CCR3, which is a principal receptor for eosinophil migration in allergic asthma, could be a promising therapeutic tool in the management of asthma. In fact, this is substantiated by the development of MAbs, antagonists derived from modified chemokines (peptide chemokine receptor antagonists), or small molecules targeting CCR3 (Schwarz, 2002).
12.6 SMM-CHEMOKINES: A NEW CLASS OF UNNATURAL SYNTHETIC MOLECULES AS CHEMICAL PROBES OF CHEMOKINE RECEPTOR BIOLOGY AND LEADS FOR THERAPEUTIC DEVELOPMENT

12.6.1 Concept and Development of SMM-Chemokines

Chemokine ligands and receptors are hotly investigated areas in biomedical research. The dissection of the biological roles of specific ligands and receptors has been challenging because of the lack of selectivity in chemokine ligand–receptor interactions, given that 50 chemokine ligands and 19 functional receptors have been described to date (Proudfoot, 2002). This also greatly limits direct applications of natural chemokines in the treatment of various diseases such as (HIV-1) infection. To overcome the limitation of natural chemokines, various approaches have been taken to develop synthetic chemokine analogues that have higher efficacy and improved properties, such as the synthetic modifications of various chemokines conducted in a number of elegant works by Clark-Lewis and colleagues (Crump et al., 1997; Jones et al., 1997; Loetscher et al., 1998) and the recent success in the application of medicinal chemistry to the generation of RANTES analogues reported by Offord and colleagues (Hartley et al., 2004). The lack of selectivity among chemokine ligands is best exemplified by viral morphage inflammatory protein (vMIP)-II, which recognizes a variety of CC and CXC chemokine receptors, including CXCR4, CCR5, and CCR2 (Moore et al., 1996). Although the potential benefits of chemokine receptor inhibitors for AIDS and other diseases have been demonstrated, the lack of selectivity in the chemokine system has made it problematic to use natural, nonspecific chemokines in clinical applications due to their potential side effects. Consequently, the development of new inhibitors engineered with higher selectivity for targeted receptors and reduced toxicity is clearly desirable, because they can be used both in clinical applications and as specific probes of receptor biology to study the role of a particular ligand or receptor.

There has been intensive work in the development of new chemokine analogues by other groups in the field. For instance, Clark-Lewis and his colleagues have studied extensively the structure–function relationship of interleukin 8 and other chemokine receptors using chemically synthesized chemokine analogues (Crump et al., 1997; Jones et al., 1997; Loetscher et al., 1998). Also, on the basis of these previous works by others, we have been working toward the development of a systematic strategy based on chemokine structures to synthesize a new family of unnatural chemokines that, unlike natural chemokines, have higher receptor binding selectivity. We recently reported our progress in developing such a strategy by employing SMM (synthetically and modularly modified)-chemokines as a potential general method for the \textit{de novo} design of novel ligands selective for any chemokine receptor of interest (Fig. 12.5). For this approach, the word “synthetically” refers to the use of total chemical synthesis to incorporate an almost unlimited range of unnatural amino acids and chemical modifications at any specific site(s). The word “modularly” refers to the changes of short important sequence modules, rather than the entire sequences, to achieve efficiency and be cost-effective in the generation of specificity and diversity.
The word “modified” refers to the ability through various chemical modifications to improve biological properties (e.g., low toxicity, high receptor selectivity, and affinity) that are more desirable in clinical applications and in basic research of receptor biology. A proof of the concept was shown by applying this strategy to transform vMIP-II, a very nonselective chemokine, into new analogues with significantly enhanced selectivity and potency for CXCR4 or CCR5, two principal coreceptors for HIV-1 entry, through modifying only a small Amino (N)-terminal module of 10 residues (Kumar et al., 2006). Two representative SMM-chemokines, RCP168 and RCP188, selective for CXCR4 and CCR5, respectively, showed similar or significantly enhanced binding affinities for their corresponding target receptors but drastically decreased or even completely abolished cross-binding activities for other receptors (Table 12.3).

In addition to high receptor selectivity, another important biological property of these de novo-designed ligands is signaling activity. RCP168, a vMIP-II analogue with its N-terminal (1–10) residues replaced with p-amino acids, did not trigger either Ca$^{2+}$ signaling or receptor internalization, which is distinct from the natural ligand of CXCR4, stromal cell-derived factor (SDF)-1α (Figs. 12.6a and 12.6b). More interestingly, RCP168 did not interfere with the Ca$^{2+}$ signaling induced by SDF-1α at its effective CXCR4 binding concentration and only showed its effect at over 20 times

**Figure 12.5** General design concept for SMM-chemokines. The N-terminal (1–10) sequence module of vMIP-II was modified by introducing unnatural D-amino acids or specific sequences grafted from other chemokines to chemically engineer receptor selectivity and/or signaling property. See color plates.
higher concentrations. RCP168 also potently inhibits HIV-1 entry, in contrast to its much weaker activity in interfering with SDF-1α signaling. Thus, these disparate inhibitory activity profiles of RCP168 in differentiating HIV-1 coreceptor function versus the normal function of CXCR4 may prove to be advantageous in clinical applications, because RCP168 may not induce unwanted Ca^{2+} signaling or interfere with SDF-1α signaling important for the normal physiological functions at the concentrations used for inhibiting HIV-1 infection. The mechanistic basis for the disparate activities of RCP168 was investigated and shown by the mutational mapping analysis of binding sites of RCP168 and other D-amino acid-containing SMM-chemokines on CXCR4, revealing that RCP168 binding sites on CXCR4 overlap significantly with HIV-1 but differ from SDF-1α (Choi et al., 2005).

The CCR5-specific SMM-chemokine, RCP188, with its grafting of the N-terminal module of a CCR5 agonist macrophage inflammatory protein (MIP)-1β, was able to mimic the signaling activity of MIP-1β in activating Ca^{2+} release in CCR5-expressing 293 cells (Fig. 12.6c). This demonstrates that CCR5-selective

### Table 12.3 Binding Affinity and Selectivity of SMM-Chemokines for CXCR4 and CCR5

<table>
<thead>
<tr>
<th>Analogue</th>
<th>CXCR4 Binding (nM)</th>
<th>CCR5 Binding (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCP111</td>
<td>22</td>
<td>4</td>
</tr>
<tr>
<td>RCP112</td>
<td>&gt;2700</td>
<td>146</td>
</tr>
<tr>
<td>RCP168</td>
<td>5</td>
<td>43</td>
</tr>
<tr>
<td>RCP169</td>
<td>141</td>
<td>&gt;2700</td>
</tr>
<tr>
<td>RCP188</td>
<td>&gt;2700</td>
<td>6</td>
</tr>
<tr>
<td>RCP189</td>
<td>&gt;2700</td>
<td>106</td>
</tr>
</tbody>
</table>

![Figure 12.6](image)

**Figure 12.6** Signaling activities of SMM-chemokines. All the data shown are representative of at least three independent experiments. (a) Ca^{2+} influx in Sup T-1 cells was measured in response to 0 nM, 100 nM, and 1 μM RCP111, which was followed by adding 100 nM SDF-1α. (b) The changes in [Ca^{2+}]_{i} of Sup T-1 cells in response to RCP168 were measured. After 5 min of incubation, 100 nM SDF-1α was added. (c) The Ca^{2+} signals induced by RCP188 in CCR5-transfected 293 cells were followed by 100 nM MIP-1β.
binding and signaling activities are encoded in the imported 10-residue sequence from MIP-1β. As such, we are able to apply the concept of SMM-chemokines to generate specific agonists and antagonists for either CCR5 or CXCR4, thus demonstrating that both receptor selectivity and signaling property can be designed at will into our synthetic ligands. While antagonists are often used for therapeutic applications, agonists can also be of great interest as molecular probes to study a particular receptor-mediated signaling pathway. The results from the first generation of SMM-chemokines reported here seem remarkable and promising considering that only the N-terminal 10-residue module was modified, yet a significant effect on binding or signaling selectivity was achieved. This illustrates the simplicity and robustness of this method and at the same time raises the possibility of its continued refinement and sophistication.

Furthermore, the high-resolution crystal structure of RCP168 (unpublished data) provided the structure-based mechanism for the selective interaction between the ligand and chemokine receptors, which guided the design and synthesis of even more selective inhibitory analogs. The high-resolution crystal structure of RCP168, when compared with that of vMIP-II previously reported by others (Fernandez et al., 2000; Liwang et al., 1999), revealed that the enhanced selectivity of RCP168 was associated with the structural changes at the N-terminus due to the D-amino acid modification and, surprisingly, the 30’s loop due to a conformational change propagation from the distal N-terminus through a disulfide bridge. The observed structural changes in the N-terminal and the 30’s loop regions of RCP168 in association with the enhanced selectivity of RCP168 suggest that these regions are important for the selectivity of chemokine–receptor interactions. Based on this structural insight, new analogues containing modifications at the 30’s loop were designed by replacing the 30’s loop of RCP168 with the corresponding region of SDF-1α, the only natural ligand of CXCR4, with the aim of generating higher CXCR4 selectivity. Indeed, this analogue, RCP303, exhibited more CXCR4 selective binding profile than the parent molecule, because the 30’s loop substitution led to a substantial reduction in binding to both CCR5 and CCR2, further confirming the role of the 30’s loop in affecting the receptor binding selectivity. In light of this finding, one may rationalize a structural basis for the conformational change cascade in chemokine–receptor interactions, which might include (1) the initial binding of the N-terminus of a chemokine to the receptor, (2) the resulting conformational changes in the N-terminus (including the N-loop) and subsequently the 30’s loop as facilitated by the disulfide bridge, and finally (3) the triggered recognition between the 30’s loop and the receptor, leading to the multipoint (at least including the N-terminus and the 30’s loop) contact between the chemokine and its receptor.

In addition to suggesting a potential mechanism for ligand-receptor selectivity based on the conformation changes of the ligand at the N-terminus and the 30’s loop, this study unveiled an interesting property of CXCR4—that is, the flexibility of the CXCR4 surface in recognizing ligands of different chiralities or, more precisely, different conformations. We previously found that all-D-amino acid peptides derived from the N-terminus of vMIP-II selectively bind to CXCR4 (Zhou et al., 2002). The results of RCP168 provided further support to the notion that CXCR4 is capable of interacting with diverse conformations of a ligand. The feature that CXCR4 can recognize both D- and L-amino acid-containing ligands is unprecedented among the
G-protein-coupled receptors (GPCR) superfamily, because we are not aware of other similar cases reported in the literature. The subsequent studies to map the receptor binding sites using site-directed mutants of CXCR4 have shown distinctive regions involved in D- versus L-ligand binding (Choi et al., 2005). Interestingly, D-amino acid-containing ligands such as RCP168 recognize sites on CXCR4 shared by HIV-1 gp120, but not by SDF-1α. In light of these results, the unusual flexibility of the CXCR4–ligand binding surface raises an intriguing question about whether HIV-1 gp120 might exploit this feature of CXCR4–ligand interaction for its entry into the target cell. It is well known that gp120 can frequently mutate itself, especially in the V3 loop region that is thought to recognize the coreceptor. The sequence mutations presumably bring about changes in the conformation of gp120, which could serve as a strategy for the virus to evade the recognition of HIV-1-neutralizing antibodies. However, it is puzzling that despite these changes, gp120 retains the ability to recognize the host cell receptors such as CXCR4. Our findings may shed some new light on a possible mechanism of how the varied structures of gp120 could be accommodated by the flexible CXCR4 ligand binding surface.

RCP168 is a prototype molecule for a novel family of SMM-chemokines that are highly selective and potent CXCR4 receptor inhibitors as compared with the natural chemokines. An important goal of the SMM-chemokine approach is to engineer de novo receptor selectivity into nonselective, natural chemokines. This was clearly demonstrated by the generation of RCP168. In contrast to the broad activity of vMIP-II, RCP168 has higher selectivity and affinity in CXCR4 binding. As a result, RCP168 is more potent in inhibiting HIV-1 infection than SDF-1α. The anti-HIV activity of RCP168 is also comparable to that of T-20. These results strongly suggest that RCP168 may serve as a prototype molecule for the development of highly selective and effective anti-HIV agents to be used in combination with other currently available drugs such as T-20 and/or drugs targeted to HIV-1 protease or reverse transcriptase. In this regard, it is encouraging to note that the structural basis for RCP168 activity as revealed by its high-resolution crystal structure has led to the design of a new generation of analogues such as RCP303 that showed even higher CXCR4 selectivity than RCP168, demonstrating that the combined structural and chemical biology approach can be used to develop chemokine analogues with desired high potency and selectivity.

12.6.2 Use of SMM-Chemokines to Probe Mechanism of Receptor Signaling

In another study (Dong et al., 2005), we applied the concept of SMM-chemokines containing D-amino acids to study the chemical basis of CXCR4–ligand interactions. Specifically, we asked the question of whether the binding and signaling functions of a chemokine ligand via its receptor are mediated by the same or distinct domains on the receptor. For this purpose, we used SDF-1α as a model to study the chemical mechanism of chemokine ligand–receptor binding and signaling. By replacing the N-terminal (1–8) residues of SDF-1α with all D-forms of (1–10) sequence module of vMIP-II, we found different stereochemical requirements in binding and signaling of the ligand, because the D-amino acid replacement only affected its signaling activity...
but not CXCR4 binding. In addition, the new compound was shown to be an effective HIV-1 inhibitor, indicating that its receptor binding, but not signaling, contributes to its anti-HIV activity. This study provided new insight into the stereochemical basis of CXCR4–ligand binding and signaling and suggests that further studies of the chemical basis of other chemokine ligand–receptor interactions may be of interest for understanding the mechanism of action of this important family of small protein ligands.

12.6.3 Application of SMM-Chemokines to Probe Mechanism of CXCR4–Ligand Interactions

In a recent work accepted for publication (Tian et al., 2005), we found that many residues throughout CXCR4 transmembrane (TM) and extracellular loop 2 (ECL2) domains are specifically involved in interaction with HIV-1 gp120, because most of these sites did not play a role in either SDF-1α binding or signaling. These results provided the direct experimental evidence for the distinct functional sites on CXCR4 for HIV-1 and the normal ligand SDF-1α. To further understand the CXCR4–ligand interaction and to develop new CXCR4 inhibitors to block HIV-1 entry, we used CXCR4-targeting SMM-chemokines, derived from the native sequence of SDF-1α or vMIP-II, as receptor probes and conducted ligand binding site-mapping experiments on a panel of site-directed mutants of CXCR4 (Choi et al., 2005). This study provided the first experimental evidence demonstrating that SMM-chemokines interact with many residues on CXCR4 TM and extracellular domains that are important for HIV-1 entry, but not SDF-1α binding or signaling (Fig. 12.7). The preferential overlapping in the CXCR4 binding residues of SMM-chemokines with HIV-1 over SDF-1α illustrates a mechanism for the potent HIV-1 inhibition by these SMM-chemokines and suggests that these chemically engineered molecules have interesting and unique receptor binding mechanisms distinct from those of the natural chemokines and may be used to selectively disrupt the coreceptor activity of CXCR4. The discovery of distinct functional sites or conformational states influenced by these receptor sites mediating different functions of the natural ligand versus the viral or synthetic ligands has important implications for drug discovery, since the sites shared by SMM-chemokines and HIV-1 but not by SDF-1α can be targeted for the development of selective HIV-1 inhibitors devoid of interference with the normal SDF-1α function.

12.6.4 Application of SMM-Chemokines to Probe Signaling Pathways Involved in HIV-Associated Dementia

To understand the mechanism of CXCR4 or CCR5 signaling in neuronal apoptosis associated with HIV-associated dementia (HAD), we have also applied SMM-chemokine analogues derived from natural SDF-1α or vMIP-II as chemical probes of the mechanism(s) whereby these SMM-chemokines prevent or promote neuronal apoptosis. Because of the profound activities of chemokine receptors in HAD, developing selective, potent inhibitors of chemokine receptors and understanding the physiological or pathological processes of HAD are crucial in devising novel strategies for clinical interventions.
We previously demonstrated how this SMM-chemokine approach can be applied to convert the nonselective vMIP-II into highly selective ligands of CXCR4 or CCR5 in terms of their binding, signaling, and/or antiviral activities (Kumar et al., 2006). We also used a similar strategy to modify biological and pharmacological properties of SDF-1a, which is of particular importance because SDF-1a, the only natural ligand of CXCR4, is neurotoxic with or without gp120 (Kaul and Lipton, 1999). In this unpublished work (Fig. 12.8), using rodent cerebrocortical cultures (RCCs) that not only contain the type and proportion of cells

Figure 12.7  Distinct functional sites for SDF-1α and SMM-chemokines highlighted on a hypothetical structural model of CXCR4. The residues involved in both SDF-1α and SMM-chemokine binding are highlighted in the lighter color and represented as ball-and-stick, whereas those selectively involved in SMM-chemokine binding are highlighted in the darker color. Only the TM domains, with side (a) and top (b) views, are shown for simplicity. (c) Schematic illustration of the locations of residues important for ligand binding on CXCR4 TM and extracellular domains. The residues involved in the binding activities of both SDF-1α and SMM-chemokines are highlighted in white spots, whereas those selectively involved in SMM-chemokine binding (most of which overlap with HIV-1 binding) are highlighted in black spots. Such overlapping sites between HIV-1 and SMM-chemokines may serve as a potential target recognized by new selective anti-HIV inhibitors. See color plates.

We previously demonstrated how this SMM-chemokine approach can be applied to convert the nonselective vMIP-II into highly selective ligands of CXCR4 or CCR5 in terms of their binding, signaling, and/or antiviral activities (Kumar et al., 2006). We also used a similar strategy to modify biological and pharmacological properties of SDF-1α, which is of particular importance because SDF-1α, the only natural ligand of CXCR4, is neurotoxic with or without gp120 (Kaul and Lipton, 1999). In this unpublished work (Fig. 12.8), using rodent cerebrocortical cultures (RCCs) that not only contain the type and proportion of cells
normally found in brain, i.e., neurons, astrocytes, and macrophages/microglia, but also express CXCR4, CCR5, and other chemokine receptor homologues (Heesen et al., 1998; Hesselgesser et al., 1997; Lavi et al., 1997; Rottman et al., 1997) that, like the human chemokine receptors, are capable of mediating HIV-1 infection via gp120 binding (Parolin et al., 1998; Plesoff et al., 1997), we demonstrated that neurotoxic SDF-1α (Kaul and Lipton, 1999) and vMIP-II can be modified to protect neurons from the neuronal apoptosis induced by CXCR4-preferring gp120 IIIB without activating the Ser/Thr kinase Akt, extracellular regulated kinase (Erk), or c-JUN-N-terminal kinase (JNK). We also found that the inhibition of CCR5 by antagonistic SMM-chemokines leads to neurotoxicity, which is in sharp contrast to the neuroprotection induced by natural CCR5 ligands (Kaul et al., 2001; Kaul and Lipton, 1999), by activating p38 MAPK. Furthermore, the simultaneous treatment of RCCs with a neuroprotective CXCR4 ligand and a neurotoxic CCR5-selective SMM-chemokine/CCR5-preferring gp120BAL led to neurotoxicity, suggesting the predominant role of steric hindrance in the neuroprotective mechanisms of antagonistic CXCR4-specific SMM-chemokines. These results illustrate a new strategy of using novel chemically engineered inhibitors of chemokine receptors to study the molecular mechanisms and intervention methods of the neuronal cell death caused by HIV-1 infection. One important lesson from HAD is that the development of new therapeutics for HAD is likely to have significant impact on several other important neurodegenerative diseases, because pathways to neuronal protection or damage may, at least in part, be common to other CNS disorders including stroke, spinal cord injury, and Alzheimer’s disease.

Figure 12.8  A model for the molecular mechanisms of action of neuroprotective CXCR4-selective SMM-chemokines versus neurotoxic CCR5-specific SMM-chemokines. Antagonistic CXCR4-selective SMM-chemokines prevent gp120IIIB-induced neuronal apoptosis most likely via steric hindrance without activating Akt, ERK, or JNK, whereas natural CXCR4 agonist SDF-1α (and HIV-1 gp120) activates the pathway involving p38 MARK. Antagonistic CCR5-selective SMM-chemokines lead to neuronal apoptosis via the p38 MAPK pathway, whereas natural CCR5 agonist MIP-1β activates a different pathway involving Akt.
12.7 CONCLUSIONS AND FUTURE PROSPECTS

Specific SMM (synthetically and modularly modified)-chemokines for a designated chemokine receptor can be potentially used in many areas of basic research and therapeutic development. For instance, one can use these SMM-chemokines as highly specific molecular probes to study the biology of chemokine receptors at the cellular level—that is, to characterize specific biological roles of a receptor in a physiological or pathological process out of the complicated and potentially overlapping receptor network. Also, the mechanism of receptor–ligand interactions at the molecular or atomic level can be investigated by making various artificial changes in ligand probes and testing how the receptor responds. This was demonstrated in another study in which the SMM-chemokines as reported here were used as chemical biology probes to discover new functional sites on CXCR4 important for the selective interactions with (HIV-1) but not with the normal physiological ligand of CXCR4, SDF-1α, providing a new basis for the development of more selective antiviral therapies than the conventional drugs that interrupt both pathological and physiological pathways (Choi et al., 2005). As chemokine receptors belong to the superfamily of G-protein-coupled receptors (GPCRs), which represent the largest class of drug targets and yet are still poorly understood in terms of their structure-function relationship and ligand-receptor interactions, the development of chemical biology probes such as SMM-chemokines may have general implications for the study of these important membrane receptors.

In addition to being valuable research tools, receptor-selective SMM-chemokines are promising leads for developing highly effective pharmaceutics. For clinical applications, SMM-chemokines can be highly desirable and advantageous for their greater selectivity, higher potency, and reduced toxicity compared with their natural counterparts (Kumar et al., 2006). In fact, biological studies of RCP168 in the blockade of HIV-1 entry demonstrated its potency and selectivity better than SDF-1α or viral macrophage inflammatory protein (vMIP)-II. Most importantly, not only was RCP168 found, in the study of drug-nonresistant HIV-1 strains, to be comparable in efficacy to some of the most well-known and commonly used HIV-1 drugs such as AZT, which targets HIV-1 reverse transcriptase as shown in this study, or T-20, which is the only marketed drug targeting HIV-1 gp41 protein and its mediated viral entry, but also RCP168 directed at an HIV-1 coreceptor expressed on the host cells was capable of overcoming the drug-resistance problem in drug-resistant viral isolates, including AZT- and 3TC-resistant HIV-1 isolates. This highlights the advantage of targeting a nonviral protein over virus-encoded proteins in terms of tackling the drug-resistance issue. Since drug resistance is a major problem in the treatment of not only AIDS but also other viral infections, the results shown in this study may have broader implications for the development of therapeutic strategies for diseases caused by other viruses.

ACKNOWLEDGEMENTS

We thank all former and current members of the Huang laboratory who have contributed to these projects and collaborators at Raylight Corporation/Chemokine Pharmaceutical Inc. and other institutions. This work has been supported by grants from the National Institutes of Health.
REFERENCES


REFERENCES

synovia of patients with juvenile rheumatoid arthritis and juvenile spondylarthropathy. *Arthritis Rheum* 39, 1703–1710.


CHAPTER 12 BASIC AND TRANSLATIONAL RESEARCH OF CHEMOKINE LIGANDS


13.1 INTRODUCTION

A. Ciechanover and A. Hershko, Israeli scientists, and I. Rose, an American scientist, received the 2004 Nobel prize in chemistry because they discovered the mechanism of ubiquitin-regulated proteolysis. Their work indicates the research direction of proteolysis, helps people understand more about their immune systems, and is of great significance in DNA repair and regulation and in treatment of human diseases.

Actually, protein ubiquitylation is a type of protein post-translational modification. Protein post-translational modification refers to the chemical modification process on specific amino acid residues that occurs after mRNA has been translated into protein. Protein post-translational modification plays an important role in organisms. The accomplishment of the Human Genome Project is one of the greatest scientific and technological achievements of our time. Upon close inspection of the first complete draft of the human genome, it is surprising to find that only about 30,000–50,000 genes have been found (Venter et al., 2001), which is only about 3–5
times those in the eelworm or the drosophila. It is far from enough to regulate such a complex life process depending only on such a small number of genes. Therefore, the protein post-translational modification process is extremely important. It makes the protein obtain more complicated structures, perfect functions, more accurate regulations, and more specific operations. Functions of many intercellular proteins are regulated by dynamic post-translational modification, and many physiological functions of cells, such as cell response to environment (Guo et al., 2004), are carried out by dynamic post-translational modification. Human complexity is not simply a result of the direct protein products of genes. It is the protein post-translational modification that allows one gene to correspond to not only one protein, which therefore, to some extent, grants our human life its diversity.

13.2 PROTEIN POST-TRANSLATIONAL MODIFICATIONS

There are more than 20 types of protein post-translational modifications in the eukaryote cell. The most common modifications include ubiquitylation, phosphorylation, glycosylation, lipodation, methylation, and acetylation. Recently, as the development of human genomics and proteomics studies, research on protein post-translational modifications has made great progress.

13.2.1 Ubiquitylation

Regulation of cell functions via proteolysis of its proteome members was neglected for a long time. The discovery of ubiquitin and thereafter the proteasome completely revolutionized the thinking of the scientific community. Ubiquitin is a highly conserved 76-amino-acid polypeptide and is expressed in all eukaryotes. Protein tagged with ubiquitin can be specifically recognized and then degraded by the proteasome, which is a common pathway for degradation of short-life or aberrant proteins. Unlike the protein degradation processing in the enteron, this proteolytic device also requires much of the cell’s energy to function. The ubiquitin combining with protein requires energy, and so does the proteolysis process itself (Hilt and Wolf, 2004). It is necessary to realize that the ubiquitin proteasome system is a regulatory system, vital to all eukaryotic cells.

13.2.1.1 Ubiquitin Proteasome System

The ubiquitin proteasome system is the regulatory mechanism in eukaryotes (Fang and Weissman, 2004). In the 1970s and 1980s, the enigma of ubiquitin-mediated proteolysis was uncovered (Fig. 13.1). Ubiquitylation occurs as a result of the sequential action of three classes of enzymes: ubiquitin activating enzyme (E1), ubiquitin conjugating enzyme (E2), and ubiquitin protein ligase (E3) (Hershko and Ciechanover, 1998; Pickart, 2001; Weissman, 2001). Substrate specificity is largely determined by the E3. The work of E2s and E3s can be reversed by the action of DUBs (Fig. 13.2). Two major classes of DUBs have been described so far: UCHs and UBPs, which are both cysteine hydrolases (Wilkinson, 2000). In general, UCHs hydrolyze
primarily C-terminal ester and amide bonds of ubiquitin and are capable of cleaving ubiquitin precursors to generate active ubiquitin. UBPs cleave and disassemble polyubiquitin chains.

13.2.1.2 Role of Ubiquitylation in Organisms

Ubiquitylation plays an essential role in ultimately all basic cellular functions—for instance, cellular differentiation, organelle biogenesis, apoptosis, DNA repair, new protein synthesis, regulated cell proliferation, protein transportation, antigen processing, and stress responses.

Bence et al. (2001) reported that protein aggregation directly impaired the function of the ubiquitin–proteasome system. Transient expression of two unrelated aggregation-prone proteins causes nearly complete inhibition of the ubiquitin–proteasome system. Because of the central role of ubiquitin-dependent proteolysis in regulating fundamental cellular events, a potential mechanism linking protein aggregation leads to cellular dysfunction and death.
Intraneuronal inclusions containing ubiquitylated filamentous protein aggregates are a common feature of many major human neurodegenerative disorders, including Alzheimer’s and Parkinson’s diseases (Spillantini et al., 1997; Engelender et al., 1999). Engelender et al. (1999) reported the interaction between synphilin-1 and ubiquitin protein ligase E3 SIAH-1 and SIAH-2 in the Lewy body. The synphilin-1/SIAH complex cannot be hydrolyzed, which leads to a high level of ubiquitylated cellular inclusion. As in the case of synphilin-1, in intact cells SIAH-2 interacts with α-synuclein and makes it monoubiquitylated. Monoubiquitylated proteins cannot be specifically degraded by a proteasome, which leads to the formation of cellular inclusion. Layfield et al. (2003) gave a new explanation regarding the presence of ubiquitylated protein in inclusion through close examination of ubiquitin-positive inclusions at the protein level. The presence of ubiquitin within inclusions may indicate a secondary, presumably protective cellular response, rather than a primary dysfunction of ubiquitin-mediated proteolysis. Ubiquitylation is probably a following process. Within this model, other factors are likely to initiate in inclusion biogenesis.

Ubiquitylation is also an important modification on histones and mainly occurs on histones H2A and H2B. Recent studies have uncovered the enzymes involved in histone H2B ubiquitylation (Robzyk et al., 2000; Hwang et al., 2004; Wood et al., 2003) and a “cross-talk” between H2B ubiquitylation and histone methylation (Sun and Allis, 2002). Wang et al. (2004a) reported the purification and functional characterization of an E3 ubiquitin ligase hPRC1Lf. It is found that hPRC1L is specific for histone H2A and monoubiquitylated nucleosomal histone H2A at Lys\(^{119}\). In addition, Lys\(^{123}\) on the S. cerevisiae H2B C-terminal domain is the substrate of E3 Rad6, and this modification is of great importance in meiosis and mitosis. The TBP-associated complex, TAF II250, mediates monoubiquitylation of histone H1, which may contribute to transcription (Pham and Sauer, 2000).

Activator ubiquitylation is essential for transcriptional activation. In 2001, Salghetti et al. (2001) proposed that ubiquitylation regulated TAD function by serving as a dual signal for activation and activator restriction and that VP16 TAD signals ubiquitylation through the E3 Met30.

In the studies on p53, it is found that HAUSP is identified as a DUB for p53 in vitro and in vivo and stabilizes p53 even in the presence of excess Mdm2 (Li et al., 2004), while COP1 serves as an E3 ubiquitin ligase for p53 in vitro and in vivo. p53 was targeted for degradation by the ubiquitin proteasome system, which inhibits p53-dependent transcription and apoptosis (Dornan et al., 2004).

### 13.2.2 Phosphorylation

Phosphorylation is the process in which the phosphoryl group is transferred from ATP onto specific sites of proteins by protein kinase. Actually, most cellular processes are regulated by reversible phosphorylation, and at least 30% of proteins have such modifications (Ficarro et al., 2002; Krupa et al., 2004). Phosphorylation always occurs on Ser, Thr, and Tyr residues of proteins. The shape and function of cells will be changed under phosphorylation regulation.

Reversible phosphorylation is related to almost all physiological and pathological processes, including cellular signal conduction, tumor occurrence, metabolism,
nervous activity, muscle contraction and proliferation, and development and differentiation of cells (Fig. 13.3). In 1992, E. H. Fischer and E. G. Krebs received the Nobel prize in physiology or medicine because of their discoveries concerning protein reversible phosphorylation as a biological regulatory mechanism.

In the process of cellular signal conduction, incretion and cytokine acting as cellular signals combine with membrane or cytoplasmic receptors and are activated by kinase. Incretion and cytokine are phosphorylated as the phosphorylation of kinase and result in intracellular signaling.

In the study of cancer, it is found that tubulin phosphorylation may lead to the occurrence of cancer. The cell possesses an LCP apoptotic step in the form of assembly of nitrotyrosinated α-tubulin onto microtubules. This leads to microtubule dysfunction and ultimately apoptosis. Phosphorylation of TTL is postulated to cause escape from LCP apoptosis. Cells that escape this LCP apoptotic step develop into cancer (Idriss, 2004).

In the study of DNA metabolism, it is found that cellular DNA damage causes the N-terminus of the 32-kDa subunit of human RPA to become hyperphosphorylated, which helps regulate DNA metabolism and promote DNA repair. Current data indicate that hyperphosphorylation causes a change in RPA conformation that down-regulates activity in DNA replication but does not affect DNA repair processes (Binz et al., 2004).

Cabrejos et al. (2004) investigated the role of phosphorylation by vertebrate protein kinase CK2 on the activity of the general transcription factors TFIIA, TFIIE,
TFIIF, and RNAPII. Data showed that the largest subunits of TFIIA, TFIIE, and TFIIF were phosphorylated by the CK2 holoenzyme. Also, RNA polymerase II was phosphorylated by CK2 in the 214,000- and 20,500-Da subunits. Phosphorylation of TFIIA, TFIIF, and RNAPII increases the formation of complexes on the TATA box of the Ad-MLP promoter. Phosphorylation of TFIIF increases the formation of transcripts, whereas phosphorylation of RNA polymerase II dramatically inhibits transcript formation.

In the studies of phosphorylation on histones, Maile et al. (2004) discovered that histone H2B was phosphorylated at evolutionarily conserved Ser33 (H2B-S33) by the CTK of the Drosophila general transcription factor TFIID subunit TAF1. Phosphorylation of H2B-S33 at the promoter of the cell-cycle regulatory gene string and the segmentation gene giant coincides with transcriptional activation.

In the studies of the HCV NS5A, Evans et al. (2004) reported that the interaction between HCV NS5A and hVAP-A was required for efficient RNA replication. Further analyses revealed that adaptive mutations suppressed NS5A hyperphosphorylation and promoted hVAP-A binding. NS5A hyperphosphorylation negatively regulates viral RNA replication.

Jones et al. (2004) described a new model for regulation of SHP-1 involving phosphorylation of its C-terminal Ser$^{591}$ by associated protein kinase C$_A$. Phosphorylation of SHP-1 negatively regulated the activity of SHP-1 as demonstrated by a decrease in the in vitro ability of SHP-1 to dephosphorylate Vav1 on tyrosine, which leads to an increase in the tyrosine phosphorylation status of its substrates.

Li and co-workers (Luo et al., submitted; Luo et al., 2004) demonstrated that phosphorylation may influence local configuration of protein through hydrogen-bond interaction. In studies of the C-terminal domain of c-Fos protein, it is found that phosphorylation on S$_{362}$ causes alteration in local configuration and further brings instability to turn structure (Li et al., 2004). From their studies, it is concluded that the MALDI-TOF MS is a reliable method to analyze the phosphorylation site on peptides or proteins (Huang et al., 2004)

### 13.2.3  Glycosylation

Protein glycosylation refers to the process in which oligosaccharides combine with specific amino acid residues on protein in the form of indican.

Protein glycosylation can be divided into four main categories, mainly depending on the linkage between the amino acid and the sugar: O-linked glycosylation, N-linked glycosylation, C-mannosylation, and GPI anchor attachments (Blom et al., 2004).

$O$-linked glycosylation mainly occurs on Ser and Thr nearing Pro and mostly on the $\beta$ configuration of protein. $O$-glycans are built up in a stepwise fashion with monosaccharides added one at a time to a growing branch. At present there is no acceptor motif defined for $O$-linked glycosylation. $O$-linked glycosylation may happen at two cellular locations in the cell. One is in the Golgi; the other is in the nucleus and cytoplasm of cells (Hart, 1997). Those taking place in the Golgi are initiated by the addition of various reducing terminal linkages such as $N$-acetylgalactosamine, $N$-acetylglucosamine, mannose, and fucose. $O$-glycosylation of secreted and membrane-binding proteins is a post-translational event, taking place
after N-glycosylation and folding of the protein (Asker et al., 1995), in the cis-Golgi compartment (Roth et al., 1994). O-glycosylation occurring in the nucleus and cytoplasm of cells is characterized by the attachment of a monosaccharide, N-acetylglucosamine, to a Ser or Thr residue (Blom et al., 2004). The best-known form of O-glycosylation in mammals is the addition of GalNAc, catalyzed by GalNAc-transferases, to Ser or Thr residues of secreted and cell surface proteins and further addition of Gal, GalNAc, or GlcNAc moieties.

O-GlcNAc residues are found in two conformations attached to the polypeptide backbone: O-α-GlcNAc and O-β-GlcNAc (Blom et al., 2004). This O-GlcNAcylation has three key features: O-GlcNAcylation occurs at the sites similar to those modified by protein kinases; O-GlcNAcylation is reciprocal to phosphorylation and may have the reverse function of phosphorylation on many proteins; and O-GlcNAcylation is highly dynamic, with rapid cycling in response to cellular signals (Hart, 1997).

Oligosaccharides attached to Asn residues of secreted or membrane-binding proteins are described as N-linked. This process is catalyzed by glycotransferase on endoplasmic reticulum. The sequence motif Asn-Xaa-Ser/Thr (Xaa is any amino acid except Pro) has been defined as a prerequisite for N-glycosylation (Gavel and von Heijne, 1990). Although rare, the sequence motif Asn-Xaa-Cys has also been shown to act as an acceptor site (Blom et al., 2004).

C-mannosylation is the attachment of an α-mannopyranosyl residue to the indole C2 of Trp via a C-C link and occurs on the first Trp in the motif W-X-X-W (or in some cases, W-X-X-C and W-X-X-F). The GPI anchor refers to glycoposphatidylinositol groups attached to a protein chain near the C-terminal, which anchors the protein to the cell membrane (Blom et al., 2004; Hart, 1997).

13.2.3.1 Analysis of Glycosylation Sites

The identification of glycosylation sites in proteins is possible through a combination of proteolytic digestion, separation, MS, and MS/MS. LC combined with MS/MS has been a reliable method for detecting glycopeptides in digestion mixtures and for assigning glycosylation sites. Krokhin et al. (2004) identified and confirmed the glycosylation site by MALDI and had set out to develop the software for automatically identifying glycopeptide precursors.

13.2.3.2 Synthesis of Glycoprotein

The synthetic challenges associated with generating homogeneous populations of selectively modified proteins have hindered detailed studies of the effects of these modifications on protein structure and function.

Zhang et al. (2004) demonstrated an approach to the cotranslational synthesis of selective glycoproteins in which the modified amino acid is genetically encoded. This approach should be generally applicable to the protein with other post-translational modifications. Also, it provides a potent strategy for producing homogeneous glycoforms of therapeutic glycoprotein.

In a recent study, the bacterial N-linked glycosylation pathway is discovered in the human gastrointestinal pathogen Campylobacter jejuni. In addition, an O-linked
glycosylation pathway has been identified and characterized in *C. jejuni* and the related species *Campylobacter coli*. These provide model systems for the elucidation and exploitation of glycoprotein biosynthesis. Both pathways show similarities to the respective *N*- and *O*-linked glycosylation processes in eukaryotes, and these modifications could play similar biological roles. Homologues of the genes in both pathways are found in other organisms. The complex glycans linked to the glycoprotein share common biosynthetic precursors (Szymanski et al., 2003).

### 13.2.3.3 Role of Glycosylation in Organisms

Protein glycosylation influences the function of protein (Cassey, 1995) and is of great importance in many cellular processes such as immunoprotection, virus replication, cell growth, intracellular adhesion, and inflammation. Known *O*-glycosylated proteins (Comer and Hart, 2000) include transcription factors, nuclear-pore and heat-shock proteins, RNA polymerase II, and the transcription product of carcinogen and enzymes (Fig. 13.4). Aberrant glycosylation always results in the occurrence of diseases.

Higher glycosylation on transferrin has been demonstrated in patients with Alzheimer’s disease, rheumatoid arthritis, and other diseases associated with a free radical etiology. Transferrin is a glycosylated metal-carrying serum protein. One of the biological functions of glycosylation is to stabilize the transferrin and indirectly influence iron homeostasis (van Rensburg et al., 2004). Also, there are many studies concerning the relationship between glycosylation and rapid-developed muscular dystrophy. Muntoni et al. (2004) demonstrated a novel mechanism responsible for muscular dystrophy. Aberrant glycosylation of *α*-dystroglycan appears to be a common finding in all these cases.

![Glycoproteins in cell](image)

**Figure 13.4** Glycoproteins in cell.
Chen et al. (2004) synthesized the O-GlcNAcylated polypeptides and studied the effects of O-GlcNAcylation on the configuration and biological function of the polypeptides. They also studied the different regulatory mechanisms between O-GlcNAcylation and O-phosphorylation.

Since protein glycosylation is vital to fundamental cellular processes, targeting glycosylation for drug design is likely to be a necessity. Therapeutic strategies are to regulate relative enzymes. Imino sugars are monosaccharide mimics that have a nitrogen atom in place of the ring oxygen. Members of the imino-sugar family can inhibit several glycosylation enzymes, including ER α-glucosidases I and II and the ceramide-specific glucosyltransferase. Targeting ER α-glucosidases at a low level could be a potential strategy for treating virus infections without compromising the host cell. Imino sugar N-nonyl-DNI has antiviral activity to an animal model of hepatitis B virus, a model of neonatal calf diarrhea coronavirus, and an \textit{in vitro} model of HCV (Dwek et al., 2002).

13.2.4 Lipodation

Protein lipodation refers to the process in which long-chain aliphatic hydrocarbons combine with protein to form conjugates on the O or S atom. Usually in the process, Cys is acetylated by a palmitoyl group or alkylated by a farnesyl group on the S atom. These two kinds of aliphatic hydrocarbon chains always modify the same protein, and the protein is anchored onto the cell membrane by the compatibility between the aliphatic hydrocarbon chain and biological phospholipids membrane (Cassey, 1995).

Lipid proteins are a type of membrane-binding protein. The specific aliphatic-hydrocarbon modification helps them anchor on the cell membrane and further helps the proteins function. Recent studies in biophysics find that only when lipid proteins have anchored onto the membrane could they obtain the activity to participate in the biological process (Guo et al., 2004).

13.2.4.1 Synthesis of Lipoprotein

In the last 10 years, genetic engineering has been introduced to organic synthesis, and a new approach has been developed to synthesize lipoconjugates. Waldmann and co-workers (Bader et al., 2000) have made significant breakthroughs in this field. Enzymes and noble metals are involved as catalyzers, and solid-phase synthesis is used in synthesizing lipopeptides. In the reaction, an artificial connecting group is introduced, and MIC is used as a protecting group. The polypeptide is then combined with C-terminal interrupted protein obtained by genetic engineering through the reactions of the MIC and hydrosulphide groups on Cys. These proteins were tested \textit{in vitro} and have biological activities. SPR showed that lipoproteins had strong interactions with membranes. Using fluorescent microscopy to examine the distribution of these lipoproteins in the cell, it was also found that they exhibited exactly the same properties as that \textit{in vivo}. Waldmann and co-workers (Völkert et al., 2003) described a modular strategy for the assembly of farnesylated N-Ras heptapeptides carrying a photoactivatable BP group within the lipid residue. This strategy is based on the fragment condensation of an N-terminal hexapeptide synthesized on the
solid support with a Cys methyl ester modified with different farnesyl analogues. Using this strategy, 24 peptides were synthesized, incorporating farnesyl analogues with four different chain lengths. Two of these photoactivatable conjugates were ligated to oncogenic human N-RasG12V/C181. A cellular transformation assay revealed that the semisynthetic proteins retained their biological activity despite the photolabel.

### 13.2.4.2 Role of Lipodation in Organisms

Protein lipodation is vital to signal conduction in life. Lipid protein acts as the switch of signal conduction. In signal conduction in the human body, a series of processes is needed from cytokine to regulation of gene expression. Signals from cytokine receptors are transduced to SH2-containing domain adaptors and then to the guanine nucleotide exchange factor and the Ras protein. The Ras protein combines with GTPd, which acts as the switch of the entire signal conduction process. The circulation of Ras protein in life is shown in Fig. 13.5. Recently, Ras protein was proposed as an effective drug target.

Aberrantly modified lipid protein will influence signal conduction. Ras protein mutants are found in 30% of all human tumor tissues. Among these tumors, 80% are malignant (Peters et al., 2002). The cause of this aberrant modification is spot mutation on the Ras protein. It is now unclear whether chemical signals or gene mutations result in Ras mutation. Nevertheless, progress has been made in the research of targeting protein lipodation for drug design. Farnesyltransferase inhibitor has pronounced inhibitory activity (Thutewohl et al., 2003), while there is no cytotoxic effect on normal cells. Also, palmitoylethyltransferase inhibitor demonstrates antitumor activity and also has an antitumor effect on breast and prostate cancer.

### 13.2.5 Methylation

Protein methylation is a complicated post-translational modification process like phosphorylation and glycosylation and plays an important role in the regulatory process. Methylation on the Lys or Arg side-chain amido is catalyzed by methyltransferase. In addition, another type of methylation occurs on the Asp or Glu side-chain carboxyl to form a methyl ester. We will focus on the former methylation type here. Methylation will strengthen the steric hindrance and influence the formation of a hydrogen bond by substitution of a hydrogen atom on nitrogen. Therefore, methylation has the ability to regulate intermolecular interaction or interaction between small molecule and target proteins.

![Figure 13.5 Circulation of Ras protein in life.](image-url)
13.2.5.1 Classification of Arginine Methylation

In 1967, Paik and Kim discovered methylation on arginine and its role in cellular processes including signal conduction, transcription activation, and protein sorting. Arginine methylation may occur on many proteins. In eukaryotic cells, arginine methylation can be divided into three categories: N^G-MMA, N^G-N^G-aDMA, and N^G-N^G-sDMA. Arginines in different proteins are differently methylated, and many proteins may even have more than one type of methylation. Methylation on hnRNPs and other RNA-binding proteins always occur on the RGG motif. Also, all arginine methylations on the RGG motif are monomethylations or asymmetric dimethylations, but not symmetric dimethylations. On the RXR and RG motifs, only asymmetric dimethylation may occur. However, arginine on myelin basic protein can be not only monomethylated but also symmetrically dimethylated. Also, arginine methylations on SmD1 and SmD3 are symmetric dimethylations. Unlike methylation of hnRNPs, that of myelin basic protein, SmD1 protein, and SmD3 protein occurs on the GRG motif, which indicates that the difference in the former or latter position of arginine may influence the methylation form.

13.2.5.2 Methylation on Histone

Histone is vital to transcription. It anticipates cellular processes by modifications such as phosphorylation, acetylation, or methylation on its tail. Histone arginine and lysine methylation are responsible for transcription regulation and heterochromosome formation (Trievel, 2004). A high level of histone acetylation is related to enhanced transcription activity, while histone methylation works in a much more complicated manner because methylation can both enhance and inhibit transcription.

Histone Lysine Methylation

Histone lysine methylations occur on H3-K4, H3-K9, H3-K27, H3-K3, H3-K79, and H4-K20 N-terminals and sometimes on the H1 N-terminal. Methylation on H3-K9, H3-K27, and H4-K20 are related to chromosome inactivation and methylation on H4-K9 may lead to inhibition in the chromatin level on a large scale. Methylation on H3-K4, H3-K36, and H3-K79 are related to chromosome activation, and monomethylation on H3-K4 may act against gene restraint caused by methylation on H4-K9 (Sims et al., 2003).

In 2001, T. Hisashi et al. [cited in Tamaru and Selker (2001)] first proved the relationship between histone H3-K9 methylation and DNA methylation. In a gene scan of an N. crassa mutant with a DNA methylation defect, dim-5 containing the SET-domain gene was identified. Dim-5 is H3 HMTaes. H3-K9 methylation by Dim-5 may help combine DNA methyltransferase by direct or indirect affinity of the mediator to methylated H3-K9. Jackson et al. (2002) gave new evidence of this relationship by function analysis of Arabidopsis kryptonite mutant.

In Shi et al. (2004), it was found that a nuclear homologue of amine oxidases, LSD1 (KIAA0601), functions as a histone demethylase and transcriptional corepressor. LSD1 specifically demethylates histone H3 lysine 4. Lysine demethylation occurs via an oxidation reaction that generates formaldehyde. Importantly, RNAi inhibition of LSD1 causes an increase in H3 Lys^4 methylation and concomitant derepression of target genes, suggesting that LSD1 represses transcription via histone
demethylation. The results reveal dynamic regulation of histone methylation by both histone methylases and demethylases.

Histone Arginine Methylation  Histone arginine methylation occurs on H3-R2, H3-R4, H3-R17, and H3-R26, and all these modifications can enhance transcription.

Methyltransferases have already been discovered, but enzymes that demethylate histones have not been identified for years. Cuthbert et al. (2004) proposed the deimination process, which converted histone arginine to citrulline and antagonized arginine methylation. Later, Wang et al. (2004) showed that human PAD4 regulated histone arginine methylation by converting methyl-arginine to citrulline and releasing methylamine. It has been found that PADI4 (refers to PAD4 in Wang’s paper) specifically deiminates arginine residues R2, R8, R17, and R26 in the H3 tail. Deimination by PAD4 prevents arginine methylation by CARM1. Dimethylation of arginine prevents deimination by PAD4, although monomethylation still allows deimination to take place. In vivo targeting experiments on ecdogenous promoter demonstrate that PAD4 can repress hormone receptor-mediated gene induction. PAD4 is recruited to the pS2 promoter following hormone induction when the gene is transcriptionally down-regulated. These results define a novel mechanism for antagonizing the transcriptional induction mediated by arginine methylation (Cuthbert et al., 2004; Wang et al., 2004b).

Trojer et al. (2004) studied filamentous fungus Aspergillus nidulans and identified three distinct PRMTs. The existence of in vivo arginine methylation on histones as demonstrated by site-specific antibodies and the high level and specificity of PRMTs for individual core histones in A. nidulans suggest an important role of these enzymes for chromatin modulating activities. All these enzymes exhibit intrinsic HMTaes activity when expressed as GST fusion proteins. Two of these proteins, termed RmtA and RmtC, reveal significant sequence homology to the well-characterized human proteins PRMT1 and PRMT5, respectively. The second methyltransferase, named RmtB, is only distantly related to human/rat PRMT3. Native as well as recombinant RmtA is specific for histone H4 with arginine 3 as the methylation site. Furthermore, methylation of histone H4 by recombinant RmtA affects the acetylation by p300/ CBP and exhibits a unique substrate specificity in catalyzing the methylation of histones H4, H3, and H2A.

13.2.5.3 Role of Methylation in Organisms

Histone methylation plays a central role not only in epigenetic chromatin modification in eukaryote cells, but also in cell differentiation, development, gene expression, genome stability, and cancer research. Other types of protein methylations and methyltransferases are also vital to organisms. Aberrant protein methylation or mutation in methyltransferase always leads to disease. For instance, monomethylated arginine and asymmetric dimethylated arginine are NOS inhibitors. These kinds of modifications have already been discovered in many patients with heart disease. Arginine methylation is not reversible, which will influence NOS activity. Several methylated proteins are found in autoimmune diseases. For example, the antibody
against myelin basic protein was detected in patients with multiple sclerosis diseases. Chuikov et al. (2004) found a novel mechanism to regulate p53 function through lysine methylation by Set9 methyltransferase. Set9 specifically methylates p53 at one residue within the C-terminus regulatory region. Methylated p53 is restricted to the nucleus, and the modification positively affects its stability. Set9 regulates the expression of p53 target genes in a manner dependent on the p53-methylation site. The crystal structure of a ternary complex of Set9 with a p53 peptide and the cofactor product provides the molecular basis for recognition of p53 by this lysine methyltransferase.

### 13.2.6 Acetylation

Protein acetylation is another important type of protein post-translational modification in cells. Acetylation occurs on many proteins, such as histone.

#### 13.2.6.1 Dynamic Histone Acetylation

Recently, more research has been carried out on histone acetylation. Histone acetylation, discovered more than 40 years ago, is a reversible modification of lysine within the N-terminal domain of core histones. Histone acetylation is catalyzed by HATs, and deacetylation is catalyzed by HDs or HDACs. The N-terminal tails of the core histones are rich in lysines and are therefore positively charged under physiological conditions. It is assumed that this allows an intimate interaction either with the negatively charged backbone of the DNA and/or with adjacent nucleosomes, leading to a “tight” nucleosome formation and higher-order chromatin folding. Acetylation on histone may weaken its interactions with the DNA, resulting in an “open” chromatin conformation (Hansen et al., 1998; Walia et al., 1998). Such a conformation facilitates the access for transcriptional regulators and the combination to transcription factor and promotes gene transcription. It is believed that localized relaxation of chromatin caused by histone N-terminal domain acetylation is a necessary but not sufficient condition for processes that repackage DNA. Until now, several types of HATs and HDs have been discovered, and CBP/p300, which may interact with several transcription regulators, is probably the most important. The function of the rapid, continuous, or repetitive acetylation and deacetylation reactions with half-lives of just a few minutes remains unknown. It is still unclear whether turnover itself is linked to chromatin transcription beyond its contribution to rapid changes toward hyper- or hypoaecetylation of nucleosomes. However, recent experiments suggest that turnover may be linked directly to steps in gene transcription (Waterborg, 2002).

Hyperacetylation of normally silenced regions or deacetylation of normally actively transcribed regions can all lead to various disorders, including developmental and proliferative diseases such as leukemia, epithelial cancers, fragile X syndrome, and Rubinstein–Taybi syndrome. Timmermann et al. (2001) reviewed diseases induced by aberrant acetylation or deacetylation, so we will not focus on this here. In vivo V(D)J recombination is developmentally regulated by enhancer-dependent changes in the accessibility of chromosomal recombination signal sequences to the recombinase, but the molecular nature of these changes is unknown. McMurry
and Krangel (2000) reported that histone H3 acetylation was measured along versions of a transgenic V(D)J recombination reporter and the endogenous T-cell receptor A/Δ locus. Enhancer activity was shown to impart long-range and developmentally regulated changes in H3 acetylation, and H3 acetylation status was tightly linked to V(D)J recombination. H3 hyperacetylation is proposed as a molecular mechanism coupling enhancer activity to accessibility for V(D)J recombination.

Polyglutamine diseases are hereditary neurodegenerative disorders caused by the expansion of a CAG repeat in the disease gene. Recent studies show that the altered balance between protein acetylation and deacetylation may be a key process contributing to expanded polyglutamine-induced pathogenesis. The restoration of this balance is possibly made by the genetic or pharmacological reduction of the opposing enzyme group, for example, the HDACs. Recent progress in HDAC research has made the development of inhibitors of specific HDAC family proteins possible, and these compounds could be effective candidates for treatment of these devastating diseases (Bodai et al., 2003).

### 13.2.6.2 Acetylation on Other Proteins

In the studies of tubulin acetylation, Hubbert et al. (2002) discovered that the enzyme HDAC6 reversed the post-translational acetylation of tubulin. This finding provided evidence that reducing tubulin acetylation enhanced cell motility. It was also suggested that decreasing tubulin acetylation reduced microtubule stability. However, in Palazzo et al. (2003), it was found that alteration in cell motility observed by Hubbert et al. in cells overexpressing HDAC6 resulted not from changes in the formation of stable microtubules, but from alterations in the degree of tubulin acetylation.

Nguyen et al. (2004) provided evidence that pRb underwent acetylation upon cellular differentiation, including skeletal myogenesis, and demonstrated that acetylation regulated the differentiation-specific function(s) of the pRb. The pRb is known to induce growth arrest and cellular differentiation. Studies have shown that the co-activator p300 was found to acetylate pRb. The biological significance of pRb acetylation, however, remains unclear. The p300-associated factor (P/CAF) can mediate pRb acetylation as pRb interacts directly with the acetyltransferase domain of P/CAF in vitro and can associate with P/CAF in differentiated cells. Experiments showed that acetylation did not affect pRb-dependent growth arrest or the repression of E2F transcriptional activity. Instead, acetylation was required for pRb-mediated terminal cell cycle exit and the induction of late myogenic gene expression.

### 13.3 INTERACTION OF DIFFERENT TYPES OF PROTEIN POST-TRANSLATIONAL MODIFICATIONS

Different post-translational modification processes do not occur isolatedly in vivo. In cellular activities, cooperation between different modifications is greatly needed. In signal conduction, for example, outside-membrane receptors and corresponding effectors, commonly glycoproteins, bind proper ligands. The stimuli from surroundings are conducted to the membrane through glycoprotein, and then to the
membrane-binding lipoprotein; finally, they activate downstream proteins or kinases. Meanwhile, lipid protein always acts as the switch of phosphorylation on a series of proteins in almost the entire signal conduction. These phosphorylations are regulated by a unique kinase and are the main body of signal conduction.

A single protein can possess more than one modification. Different types of modifications influence each others’ functions and cooperate with each other. Phosphorylation is parallel with glycosylation in many cases, especially in dynamic characteristics and in its ubiquitous existence in cells. These two modifications are found in transcription factors, carcinogenic products, and enzymes. From the data obtained, the addition and elimination of $N$-acetylglucosamine (GlcNAc) may act as a regulatory mechanism. Besides the fact that phosphorylation and glycosylation have a similar process, the two types of modifications are reciprocal, which is the same as the concept of “Yin–Yang” in ancient China (Qian et al., 2003). Phosphorylation and glycosylation on RNA polymerase II, for example, play different roles in gene expression. The RNA polymerase C-terminal domain is hyperglycosylated. As polymerase enters into the nucleolus and interacts with the transcription factor, the protein is immediately and totally deglycosylated and at the same time phosphorylated, which indicates that RNA polymerase can possess a high level of GlcNAc only when nonphosphorylated. Phosphorylation and glycosylation occur in the inner and outer parts of the cell membrane, respectively. Thus, it is concluded that glycosylation shares the same sites with phosphorylation and influences the phosphorylation degree. For example, a microtubule-associated protein, tau, is multiglycosylated. There are more than 12 glycosylation sites, including about 4 GlcNAcyl-modified sites, in one molecule. In the brains of patients with Alzheimer’s disease, tau proteins form a kind of hyperphosphorylated double-helix tangle. The phosphorylated Ser$^{262}$ on tau protein, which should be glycosylated in its normal condition, blocks the interaction with microtubule. Therefore, the aberrant phosphorylation is induced by glycosylation defection (Qian et al., 2003).

Histone can be methylated and acetylated at the same time (Fig. 13.6) (Rice and Allis, 2001).

![Figure 13.6](image-url)  
**Figure 13.6** Post-translational modification on histones.
The major targets for post-translational acetylation and methylation are conserved lysine residues located in the N-terminal tails of histones H3 and H4. Histone acetylation occurs throughout the cell cycle, whereas histone methylation peaks in G2 phase and during heterochromatin assembly (Rice and Allis, 2001). Observations have suggested that certain combinations of acetyl and methyl modifications of lysines in histone tails may have antagonistic or cooperative biological effects. For example, hyperacetylated H4 from transcriptionally active chromatin preparations is a preferential target of histone H3 methylation, suggesting that these modifications may act synergistically to promote transcription in a way that remains unclear.

13.4 SUMMARY

Because protein post-translational modification is not directly decided by genes, studies on protein post-translational modification are of great significance for future research on proteomics. Thus, the post-translational modification proteomics is established and now becomes an intense focus of research in the world. Recent advances in these modifications are helping elucidate the role of these modifications in life and hold great promise for future pharmaceuticals. Targeting the mutantly modified protein in aberrant cells will help determine how post-translational modification regulates its function. Understanding what influences the post-translational modification will help uncover cellular processes and functions of protein networks at the molecular level and will enable more precision in directly targeting molecules for drug design. Post-translational modification mimics are set to dominate the next wave of protein therapeutics and become powerful medicinal tools in the twenty-first century (Benjamin, 2004).

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (Grant Nos. 20272032 and NSFCBIC 20320130046) and the Teaching and Research Award Program for Outstanding Young Teachers in Higher Education Institutions of MOE, China.

REFERENCES


14.1 INTRODUCTION

While new drug discovery remains a high priority in pharmaceutical industries, research on innovative drug-delivery methodologies is also drawing comparable attention. Successful delivery technology can ensure a full use of the drugs with higher efficiency, longer duration, and fewer side effects. Economic consideration is also a driving force in the research of high-performance drug-delivery systems. Now drug delivery is one of the fastest-growing segments of the pharmaceuticals market (Barbe et al., 2004). New drug-delivery methods aim at the development of capabilities of delivering precise quantities of a drug at the right time and as close as possible to the treatment site. Biomaterials suitable for local delivery are currently one of the most important topics in medical research (Meseguer-Olmo et al., 2002). A local release system of drugs has some advantages, such as decreasing the risks of systemic

*Corresponding author: weiyen@drexel.edu
toxicity and side effects associated with oral and parenteral therapies, and higher drug concentrations in the desired area could be achieved to improve the efficacy of the treatment (Falaize et al., 1999).

Interest in controlled release has grown greatly and rapidly. Controlled release is concerned with the control in time and space of the biological effects of drugs. The ideal controlled-release system is one that can release drugs in a controlled manner and that has a long enough time span to provide optimum effectiveness of the drugs, thus to minimize adverse reactions, or to maximize efficacy (Falaize et al., 1999). These objectives may be achieved by controlling diffusion, reaction rates, or other physicochemical parameters by using some rate-controlling materials. There are a number of mechanisms that provide timed release of drugs, such as microencapsulation, transdermal patches, and implants (Freiberg and Zhu, 2004). Sustained drug delivery provides the same medicinal effect with higher efficiency, longer duration, and fewer side effects than traditional delivery methods.

A variety of materials have been studied as drug carriers. The desired carriers should be able to reduce the toxicity of drugs, increase their absorption, and/or improve their release profile. In recent years, numerous polymeric carriers and controlled-release systems have been developed (Acemoglu, 2004). Biodegradable and bioerodible polymers with nontoxic degradation products have been well studied as drug-delivery systems (Grolleman et al., 1986; Laurencin et al., 1987; Leong et al., 1985). One drawback with these polymer delivery systems is that possible damage could be done to the loaded drug caused by the exothermic polymerization reaction. Sol–gel-derived glasses have also been studied extensively in the past decade as controlled-release materials (Ahola et al., 2000; Bottcher et al., 1998; Falaize et al., 1999; Kortesuo et al., 1999, 2001; Nicoll et al., 1997; Radin et al., 2002; Santos et al., 1999). The sol–gel materials could be used to host various synthetic or natural therapeutic agents for controlled-delivery applications. It was shown that these materials are biocompatible in vivo. The materials cause no adverse tissue reactions and degrade into Si(OH)$_4$, which is eliminated from the body through the kidneys (Kortesuo et al., 2000). The drug molecules are simply entrapped within the silica matrix with no chemical bonding and remain in a biologically active form. The release is achieved through the dissolution and diffusion of the drug molecules from the sol–gel matrix. Upon release, the drug can immediately exert its effect upon the body. It has been reported that the bactericidal effect of antibiotics is retained upon their release from silica materials (Aughenbaugh et al., 2001; Radin et al., 2001). Their biological effect is the same as that when used alone and could be predicted based upon their release rate. So far these materials have been found to be very promising in bone therapies (Arcos et al., 2003; Hall et al., 2003; Meseguer-Olmo et al., 2002). Because of poor blood circulation in bone tissue, systemic antibiotic administration often cannot provide efficient concentration (Arcos et al., 2003). It was found that sol–gel synthesis parameters, such as sol pH, water/alkoxide molar ratio, or drug concentration, have an effect on the drug release rate (Aughenbaugh et al., 2001; Kortesuo et al., 2001). The silica matrix delivery approach also provides a sustained delivery. The release of the drug to the physiological environment is extended over a longer period than if the drug is used in its native form. However, most of these sol–gel materials are microporous with typical pore sizes smaller than 1.5 nm and very broad pore size distribution.
In 1998, we reported a novel, low-cost, environmentally friendly, biocompatible, nonsurfactant templated method for the preparation of mesoporous or, more generally speaking, nanoporous materials via sol–gel reactions (Wei et al., 1998; for a recent review, see Wei and Qiu, 2004). By using nonsurfactant organic molecules as template, mesoporous materials can be prepared under mild conditions. In contrast to most surfactant templated synthesis of mesoporous materials, the template removal is achieved by solvent extraction at room temperature (Wei et al., 1999). The materials usually have high surface areas (e.g., 1000 m$^2$/g) and large pore volumes (e.g., 0.5–1.0 cm$^3$/g) as well as controllable pore sizes in the range of 2–12 nm with a narrow size distribution. The pore diameters are often tunable by simply varying the template concentration. One of the most unique advantages of our nonsurfactant approach to mesoporous materials is its biocompatibility. The entire preparation can be performed at room temperature and near neutral pH. The nonsurfactant templates could be sugar compounds, such as glucose, maltose, fructose, soluble starch, or cyclodextrins, which are highly biocompatible.

In general, the synthesis starts with the hydrolysis of inorganic precursors—for example, tetramethyl orthosilicate (TMOS) or tetraethyl orthosilicate (TEOS) for silica. The template is then added into the sol. Upon gelation and drying, the template–silica nanocomposites are obtained usually as transparent and monolithic solids. The template molecules can be completely removed by solvent extraction. The removal of the templates results in a mesoporous silica matrix with interconnected channels of regular diameters. These channels usually do not have any discernible packing order or orientation. Thus the mesoporosity is isotropic and accessible in all directions. The pore parameters, such as pore size, pore volume, and surface area, are adjustable to a certain extent by varying the concentration of template in the preparation. Generally, the values of pore parameters increase with the template concentration. The mesoporous materials prepared with nonsurfactant templates have been successfully used for enzyme and protein immobilization (Wei et al., 2000, 2001, 2002a,b). These studies showed that these macromolecules could retain their biological activities within the mesoporous silica matrices. Through proper design of the channel diameter, these incorporated macromolecules would not leach out of the matrices. These bioactive sol–gel materials could be used as biocatalysts and biosensors (Dave et al., 1994; Ellerby et al., 1992). More importantly, the characteristics of the tailorable physical (e.g., pore size and volume) and chemical (e.g., hydrophilicity and other functionalities) structures of the nanopores could provide a unique system for the stabilization of proteins against denaturation and foreign agents such as protease. The nanoporous matrices may also function like rigid matrix artificial chaperones to provide designable assistance in protein folding processes (Wei and Qiu, 2004). Recently, the single protein molecule that may function as a catalytic enzyme is thought to undergo constant unfolding-refolding conformational changes. The catalytic activity measured is the average of all conformations within the measurement period (Xie, 2001). The controlled biocompatible nanoporous materials containing a single catalytic protein may provide an excellent and a unique system for studying and establishing the effect of such a conformational change on the bioactivity of the protein molecules because the conformational change depends on available space around the protein molecules (Wei and Qiu, 2004).
In this work, sol–gel silica materials with designed pore parameters were synthesized to host various antibiotics for controlled-delivery applications. Various amounts of fructose were used as template to tailor the pore structure. The in vitro release properties were investigated in relation to the silica pore structures. The sol–gel synthesis parameters other than template concentration, such as pH of the sol, water/alkoxide molar ratio, and drug concentration, were more or less fixed in this study to avoid further complication in the pore structure variations. With certain pore sizes, an initial large concentration release followed by a long-term steady-state release was observed. This is desirable for some treatment modalities. We found that the amount of antibiotics released and the release rate were dependent on the silica nanostructure that was controlled by the amount of fructose template employed in the sol–gel reactions.

14.2 EXPERIMENTAL

14.2.1 Materials

Antibiotics hygromycin B, mycophenolic acid, and vancomycin (vancomycin-HCl) were purchased from Sigma. Tetramethoxysilane (TMOS, Aldrich), hydrochloric acid (HCl, Fisher), and D-fructose (98%, Aldrich) were used as received without further purification. Doubly distilled, deionized water was used to make all aqueous solutions.

14.2.2 Synthesis of Mesoporous Silica Materials Loaded with Antibiotics

Immobilization of antibiotics in silica gels was carried out by mixing the antibiotics solution with prehydrolyzed TMOS sol. For a typical example, at room temperature, to a homogeneous solution of 1.52 g of TMOS, 15 μL of 40-mM HCl, and 0.35 g of distilled and deionized H2O in a 10-mL glass vial, an appropriate amount of 50 wt% fructose aqueous solution was added under magnetic stirring. Amounts of 50 wt% fructose solution were designed to yield 0–60% by weight of template in the final dry gel products. The sol mixture was prepared in a relatively large quantity for the convenience of preparing a series of samples with different concentrations. Ten milligrams of antibiotics in 1.0 mL distilled and dionized water were added into each vial in a period of 10 min under stirring. Upon completion of the dissolution, the stirring was continued for another 3–5 min, and then the stirring bar was removed. The vial was then sealed with a piece of paraffin film. Upon gelation of the system within a few minutes to a few hours at room temperature, 2–4 holes were pinned in the paraffin film with a syringe needle to allow for the evaporation of solvent and reaction byproducts (i.e., methanol and water). After 7 days under fume hoods, the sample-containing vials were placed in a dynamic vacuum oven and dried to reach a constant weight at room temperature in about 1 week. The silica gel samples containing both the antibiotic and template molecules were obtained in the form of transparent, dry, hard, and glassy disklike monoliths.
14.2.3 Instrumentation and Characterization

Characterization of the microstructure parameters (e.g., pore size and pore size distribution, specific surface area, and pore volume) of the silica matrices was carried out with N2 adsorption–desorption isotherms at −196°C on a Micromeritics ASAP 2010 system (Norcross, GA) on the samples after completion of the releasing study. Prior to measurements, the samples were degassed at 100°C and 1 Pa for 6–7 h. Infrared spectra of KBr powder-pressed pellets were recorded on a Perkin–Elmer Model 1600 FTIR spectrophotometer. Thermal gravimetric analysis (TGA) was carried out with a TA Q50 Instrument under oxygen atmosphere at a programmed heating rate of 10°C/min.

14.2.4 Drug Releasing Study

A 0.01 mM phosphate buffered saline (PBS) solution with a pH of 7.4 was used for the releasing study following the similar protocols in the literature (US FDA, 1998). The antibiotics-containing silica disks were immersed separately in 3 mL of PBS in 5-mL glass vials and placed in a 37°C water bath. The disks usually cracked within a few minutes after immersion. The immersion solutions were exchanged after a designated time period. At the end of the time period, the solution was completely removed from each vial, and 3 mL of fresh PBS was immediately added. Immersion was continued until release stopped. To confirm the releasing pattern, the procedure was repeated with two more independent preparations.

14.2.5 Antibiotics Release Assay

The antibiotics concentration of each elution sample was analyzed using a UV/VIS spectrometer (Perkin Elmer, Lambda 2) at different wavelengths, namely 260, 250, and 280 nm, respectively, for hygromycin B, mycophenolic acid, and vancomycin. Antibiotics-PBS standard solutions were used for calibration under identical conditions.

14.3 RESULTS AND DISCUSSION

14.3.1 Materials Characterization

The as-synthesized sol–gel silica materials loaded with template and antibiotic molecules were transparent, glassy disks with typical dimensions of 8 mm in diameter, 3–8 mm in thickness, and 0.7–1.7 g in weight depending on the fructose content. The good optical transparency indicates that the template and antibiotics molecules were homogeneously distributed within the silica without macroscopic phase separations. The synthesis of the silica–template–antibiotics composite has been achieved under mild conditions such as at room temperature and low acidity, which should not cause denaturation and damage of bioactive molecules. When the silica biogels were immersed into PBS solution, they usually cracked into several smaller pieces, which is typical of bulk sol–gel materials.
Figure 14.1 shows some representative FT-IR spectra of the silica materials with increasing template content. FT-IR spectral series of the materials loaded with 10 mg of each antibiotics were similar to those in Fig. 14.1. The bottom spectrum in Fig. 14.1 is from the sample without fructose template but loaded with 10-mg antibiotics, which is the same as the IR spectrum of pure silica. Incorporation of a very small amount (i.e., 10 mg) of antibiotics did not show any appreciable effect on the spectral features. In contrast, as the template concentration is increased, the spectra exhibit some noticeable changes in the 2900–2800 cm\(^{-1}\), 1500–1400 cm\(^{-1}\), and 820–780 cm\(^{-1}\) regions. With the increasing content of template, these changes become more significant. These regions correspond to aliphatic C–H bending and stretching vibrations of the fructose molecules. Hence, these vibration bands are indicative of the presence and content of fructose template in the sample series.

Figure 14.2 shows the FT-IR spectra of the same silica materials as in Fig. 14.1, but after the drug release study. The spectral series do not show any discernible absorption in the 2900–2800 cm\(^{-1}\), 1500–1400 cm\(^{-1}\), and 820–780 cm\(^{-1}\) regions as observed earlier in Fig. 14.1. All the spectra show the typical absorption bands as those in pure silica, which indicates that all fructose template has been completely removed during the drug release studies. The complete removal of fructose template from the silica matrices is further evidenced by the fact that calcination of all the silica materials after drug release study to 750°C shows no appreciable weight loss in the TGA measurements. These observations indicate that during the drug release studies, the templates were completely extracted out of the template–silicate nanocomposites and the silica matrices were nanoporous.
Nitrogen adsorption–desorption isotherms of the nanoporous silicas after the removal of fructose template were recorded over the relative pressure ($P/P_0$) from 0 to 1 at $-196^\circ$C. The typical isotherms of silica samples prepared at various amounts of the fructose template from 0 to 60 wt% are shown in Fig. 14.3. Figure 14.3a shows the isotherms of the silica materials loaded with antibiotics mycophenolic acid. Isotherms of the silica loaded with the antibiotic hygromycin B are similar. Figure 14.3b shows the isotherms of the silica materials loaded with vancomycin. At a fructose content ranging from 15 to 60 wt% in the sample preparation, the isotherms have similar features. According to the IUPAC classification (Sing et al., 1985), these isotherms are typical type IV isotherms. They have a characteristic hysteresis loop at around 0.5–0.9 $P/P_0$, which is associated with capillary condensation in mesoporous structure. With increasing concentration of fructose, the hysteresis loops become greater in magnitude and shift to higher $P/P_0$.

However, without using fructose template, the isotherms of silica loaded with only antibiotics seem to differ when a different antibiotic compound was entrapped. This is demonstrated by the difference between the two bottom isotherms of Fig. 14.3a and Fig. 14.3b. The bottom isotherm in Fig. 14.3a has no hysteresis loop, which usually indicates that the material is microporous. The bottom isotherm in Fig. 14.3b resembles a type IV isotherm, indicating that the material might have mesoporosity to a certain extent. The difference could be attributed to the molecular size of the antibiotics. Vancomycin’s molecular structure is much larger than the other two antibiotics. Hygromycin B ($\text{C}_{20}\text{H}_{37}\text{N}_3\text{O}_{13}$) has a formula weight of 527.5 and mycophenolic acid ($\text{C}_{17}\text{H}_{19}\text{O}_6$) has a formula weight of 320.3, while vancomycin ($\text{C}_{66}\text{H}_{75}\text{Cl}_2\text{N}_9\text{O}_{24}$) has a formula weight of 1448.5. Mesopores might form with
vancomycin, probably because of its large size, but with the other two antibiotics, only micropores formed in the silica matrix.

The composition of the fructose-templated silica materials loaded with vancomycin and the pore parameters of the mesoporous silica after the release study are summarized in Table 14.1. The data of materials loaded with two other antibiotics were quite similar except the pore diameters for sample Fru-0, in which

![Figure 14.3](image-url)

Figure 14.3  Nitrogen adsorption–desorption isotherms at $-196^\circ\text{C}$ for porous silica samples prepared with various amounts of fructose and after the drug release study (i.e., the fructose was completely removed from the samples): (a) With antibiotics hygromycin B or mycophenolic acid; (b) with vancomycin.
no fructose template was added. For the Fru-0 sample of vancomycin, the pore diameter is about 5 nm. For the Fru-0 sample of the other two antibiotics, the pore diameter is about 3–4 nm. The compositions measured based on the total weight loss at 750°C from TGA are comparable with the designed feeding concentrations. The discrepancies could be attributed to incomplete sol–gel reactions. The silica materials exhibit a large pore volume up to 1.1 cm³/g, and the materials have tunable pore diameters of 5–12 nm. The effects of template concentrations on pore parameters resemble those we observed for other nonsurfactant compounds (Wei et al., 1998, 1999). As the fructose concentration increases, the pore volume and pore diameter tend to increase. The Brunauer–Emmett–Teller (BET) surface area increases with the fructose content increase up to 40 wt%; after that, it tends to decrease. As reported earlier, this is a result of competition of the effect of pore diameter increase over the effect of pore volume increase (Pang et al., 2000, 2001).

As shown in Fig. 14.4, the Barrett–Joyner–Halenda (BJH) pore size distribution curves from the desorption branches of the samples encapsulated with hygromycin B or mycophenolic acid (shown in Fig. 14.4a) and vancomycin (shown in Fig. 14.4b) exhibit relatively narrow distributions. In the case of the vancomycin sample (Fig. 14.4b), the pore diameter is in the range of 3–11 nm. A larger pore size is observed for a sample with higher fructose content. The Fru-0 curve exemplified in Fig. 14.4b is from the sample loaded with vancomycin. For the other two antibiotics, there is no identifiable distribution peak for samples in the absence of fructose template (i.e., Fru-0) as shown in Fig. 14.4a.

Isotherms analysis indicates that the removal of the template after the release study results in a mesoporous silica matrix with interconnected channels of regular diameters. This feature was also confirmed by transmission electron microscopic and X-ray diffraction studies. These interconnected channels allow the surrounding fluid

### TABLE 14.1 Composition and Pore Parameters of the Sol–Gel Silica Samples Prepared in the Presence of 0–60 wt% Fructose and 10 mg of Vancomycin After Release Study

<table>
<thead>
<tr>
<th>Sample Code</th>
<th>Fructose (wt%)</th>
<th>BET Surface Area (m²/g)</th>
<th>Pore Volume (cm³/g)</th>
<th>D&lt;sub&gt;BHT&lt;/sub&gt; (Å)</th>
<th>D&lt;sub&gt;BET&lt;/sub&gt; (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fru-60</td>
<td>60</td>
<td>345</td>
<td>1.07</td>
<td>116</td>
<td>124</td>
</tr>
<tr>
<td>Fru-50</td>
<td>50</td>
<td>336</td>
<td>0.72</td>
<td>82</td>
<td>86</td>
</tr>
<tr>
<td>Fru-40</td>
<td>40</td>
<td>360</td>
<td>0.61</td>
<td>66</td>
<td>68</td>
</tr>
<tr>
<td>Fru-30</td>
<td>30</td>
<td>329</td>
<td>0.48</td>
<td>57</td>
<td>58</td>
</tr>
<tr>
<td>Fru-15</td>
<td>15</td>
<td>228</td>
<td>0.31</td>
<td>57</td>
<td>55</td>
</tr>
<tr>
<td>Fru-0</td>
<td>0</td>
<td>111</td>
<td>0.15</td>
<td>58</td>
<td>54</td>
</tr>
</tbody>
</table>

*a* The numerical figure in the sample code denotes the template fructose concentration (wt%) in the materials before the release study.

*b* Values from the weight loss of the samples, which represent total content of volatile compounds at 750°C based on TGA measurement.

*c* Average pore diameter calculated from 4V/A, where V is the total pore volume and A is the BET surface area.

*d* Average pore diameter calculated from BJH desorption branch of the isotherms.
to penetrate and to dissolve and release the template molecules and the incorporated antibiotics molecules. Thus, the antibiotics could elute from the silica matrix through these channels. Because the diffusion rate of drug molecules should be related to the channel diameter and structures, the drug release could, therefore, be controlled by tailoring the pore parameters.

Figure 14.4 BJH pore size distributions obtained from the desorption branches of nitrogen sorption isotherms at $-196^\circ$C for porous silica samples prepared with various amounts of fructose after the drug release study: (a) With antibiotics hygromycin B or mycophenolic acid; (b) with vancomycin.
14.3.2 *In Vitro* Release

*In vitro* cumulative release of vancomycin is plotted in Fig. 14.5. The release study was repeated with two more independent sample preparations. The same release profile was obtained each time. The solutions were exchanged at the designated time points. All the samples showed a faster initial release followed by a slower, steady release. The shape of the release profile was found to be dependent on the template content. The results indicate that the release is faster when the template content is higher. The data show that the release was gradual and sustained when the amount of template was no more than 30%. From Table 14.1, one can see that when the template content is no more than 30%, the interconnected channels formed by the template removal have a similar pore diameter (~5 nm) to the Fru-0 sample. By increasing the template content, the channels are better connected (increased pore volumes). A larger pore volume allows more of the immersion solution to enter the silica matrix in a given period of time. Thus, vancomycin could be released faster than from the Fru-0 sample, although these materials have the same pore diameters. For the nontemplated Fru-0 sample, probably because of the limited interconnections of channels as implied from small pore volumes, the immersion solution could not readily penetrate through the silica matrix to the interior; therefore, the release of antibiotics from the silica matrix was difficult. Samples Fru-40 to Fru-60 have much larger pore diameters and volumes. The antibiotics can easily leach out through the interconnected channels. This is likely the reason for the observed fast release. The templated silica materials showed a very good recovery of loaded antibiotics. The amount of antibiotics released was more than 85% for all templated samples, while only about 25% of vancomycin was released after 15 days for the nontemplated Fru-0 sample. The
good recovery of loaded antibiotics indicates that there was no significant loss of the antibiotics during the synthesis and that the antibiotics were not chemically bonded to the materials.

The release kinetics were also dependent on the amount of template. When template was introduced into the silica matrix, the samples all showed a faster release compared to the Fru-0 sample. When the cumulative release versus time curve shows some linearity, that part of the process is considered as near-zero-order. The last part of the release is characterized by near-zero-order kinetics for all the samples as shown in Fig. 14.5. The cumulative release was also plotted versus the square root of time (in hours). A linear relationship shows in part of curves for samples Fru-0 to Fru-30 as shown in Fig. 14.6. This indicates that a first-order release occurred during part of the process. For sample Fru-0, the curve was linear up to 2 days of immersion with $r^2$ value 0.99. The fast first-order release was followed by a slow near-zero-order release from day 2 onward. For sample Fru-15, a three-stage release process was observed. That is, first an initial faster release stage in day 1, followed by a relatively fast stage with first-order kinetics until day 10 with $r^2$ value 0.99 and a subsequent slower release of near-zero-order. For sample Fru-30, a faster initial stage was observed for the first half day, followed by a first-order release up to day 5 with $r^2$ value 0.97 and a subsequent slow near-zero-order release. The transition from the first- to the zero-order process (5 and 10 days for samples Fru-30 and Fru-15, respectively) occurred when about 80% of vancomycin had been released. But for samples Fru-40 to Fru-60, a much faster initial release occurred in the first day, and the kinetics could not be determined. The release kinetics were dependent on the template content. By varying the template content, the release kinetics of silica materials could be tailored with respect to specific therapeutic goals.

In contrast, for the other two antibiotics—that is, hygromycin B and mycophenolic acid—the release profiles did not show a controlled and sustained release. Figure 14.7 shows, for example, the release profile of hygromycin B. About 90% of the two antibiotics were released within 3 days. The release kinetics were not template content-dependent. The release rate and amount do not have a relation to the silica structure parameters, such as pore diameters. Such a significant difference from vancomycin could be attributed to the size effect. For the two antibiotics hygromycin B and mycophenolic acid, their molecular sizes are smaller than the diameters of the interconnected channels with the smallest amount of template (15 wt%) employed in the study. They can easily leach out of the silica matrix. Therefore, no sustained release was observed.

### 14.4 CONCLUSIONS

Nanoporous silicas containing antibiotics have been successfully prepared via the nonsurfactant-templated sol-gel reactions of tetraalkoxysilane in the presence of fructose as the template and the antibiotics, including hygromycin B, mycophenolic acid, and vancomycin, followed by removal of the template by extraction during the drug release. The pore parameters, such as pore diameter, pore volume, and specific area, can be tailored by varying the template concentration. In the case of vancomycin,
the biomaterials show a long-term controlled release. The amount of antibiotics released and the release rate are dependent on the silica pore parameters. From the samples prepared with the template, most of the antibiotics (>85%) were released in a short time period. In contrast, for the samples prepared with less or no template, only about 25% of the antibiotics were released by day 15. Since the template provides isotropically distributed interconnected channels inside the silica matrix with tunable pore diameters, the release rate and amount are greatly improved. Variation of the

Figure 14.6 Part of cumulative release (in mg) versus the square root of time (in hours) shows linearity in samples Fru-0 (a), Fru-15 (b), and Fru-30 (c).
template content is a very simple, effective technique for the control of the rate and amount of drug release. With a high antibiotic recovery and with the release kinetics that can be tailored with respect to the requirements for a particular therapeutic treatment, porous silica materials with tunable pore diameters would make a good controlled drug release system.

**ACKNOWLEDGMENTS**

We are grateful to the National Institutes of Health (Grant No. R01-DE09848), the Nanotechnology Institute of Southeastern Pennsylvania, and the U.S. Army Research Office for their support of this work. We thank A. A. and G. B., respectively, for insightful suggestions and intellectual encouragement.

**REFERENCES**


---

**Figure 14.7** Release profile of hygromycin B in silica materials with various template content.


PART III

BIOLOGICAL APPROACHES AND TRANSLATIONAL RESEARCH IN DRUG DISCOVERY
CHAPTER 15

RETINOIC ACID AND ARSENIC TRIOXIDE TREATMENT IN ACUTE PROMYELOCYTIC LEUKEMIA: A MODEL OF ONCOPROTEIN TARGETED THERAPY

Jian-Hua Tong, Sai-Juan Chen, and Zhu Chen*

15.1 INTRODUCTION

15.2 CYTOGENETIC CHARACTERISTICS OF APL CELLS AND THEIR SENSITIVITY TO ATRA

15.3 RETINOIC ACID SIGNALING PATHWAY IN APL CELL MATURATION

15.4 ARSENIC TRIOXIDE IN THE TREATMENT OF APL

15.5 CONCLUSIONS

ACKNOWLEDGMENT

REFERENCES

15.1 INTRODUCTION

Acute promyelocytic leukemia (APL) is a distinct subtype of acute myeloid leukemia (AML), and it accounts for 10–15% of AML cases in adults. According to the FAB (French–American–British) nomenclature of acute leukemia, APL is classified as M3 subtype, characterized by the block of differentiation of clonal hemopoietic precursors at the promyelocytic stage of the myelopoiesis and the presence of a specific cytogenetic marker. In addition to the common clinical characteristics of acute leukemia (such

*Corresponding author: zchen@stn.sh.cn
as anemia or fever), most APL patients present a severe bleeding syndrome, often accompanied by hypofibrinogenopenia and a disseminated intravascular coagulation (DIC), resulting in relatively high early mortality (Avvisati et al., 2001).

The therapeutic approach of APL has been greatly improved over the last three decades. In the late 1970s and early 1980s, conventional cancer chemotherapy, consisting of an anthracycline and cytosine arabinoside, was mainly used in the treatment of APL. Although APL patients are usually sensitive to chemotherapy, and a relatively high complete remission (CR) rate (60–70%) is observed, use of high-dose chemotherapy drugs can produce severe side effects—such as bone marrow suppression, immunological dysfunction, gastrointestinal reaction, hemorrhage exacerbation, and even a secondary tumor occurrence—by killing normal cells. In the mid-1980s, the application of all-trans retinoic acid (ATRA), first performed at the Shanghai Institute of Hematology (SIH), for the treatment of APL achieved a significant breakthrough in cancer differentiation induction therapy (Huang et al., 1988). Instead of killing the malignant cells as chemotherapy does, ATRA can induce APL cells to undergo terminal differentiation, with which a high CR rate of 85–90% is achieved by rapidly ameliorating the severe coagulopathy of APL patients and reducing the cytotoxic effects and multiple drug resistance caused by chemotherapy. However, although very effective in remission induction of APL patients, the treatment with ATRA alone is not sufficient to maintain a long disease-free remission, and relapse is inevitable if remission is not consolidated with chemotherapy. Moreover, in the majority of cases, relapsed patients often lose their sensitivity to ATRA (Degos and Wang, 2001). Since the 1990s, arsenic trioxide (As2O3), an ancient traditional Chinese drug, has been used as an effective drug in the treatment of APL, not only in newly diagnosed cases, but also in relapsed and/or refractory patients after ATRA or chemotherapy (Shen et al., 1997; Sun et al., 1992). The principal side effects of As2O3 are dermatological and gastrointestinal symptoms, which can be tolerated and/or can disappear rapidly with symptomatic treatment. Recently, another arsenic derivative, As4S4, has also proven to be highly effective and safe for both remission induction and maintenance in all stages of APL (Lu et al., 2002). Besides, it has been reported that a dual treatment with ATRA and As2O3 for APL can further improve the CR rate as well as the overall survival of patients (Shen et al., 2004). This ATRA/As2O3 dual treatment may open a new era for the real cure of APL, a previously life-threatening subtype of leukemia.

With the revolutionary achievement in the treatment of APL, significant progress has also been made in understanding this malignancy during the last decades.

15.2 CYTOGENETIC CHARACTERISTICS OF APL CELLS AND THEIR SENSITIVITY TO ATRA

APL is distinct from other subtypes of leukemia because of its two biological features: (1) its exclusive cytogenetic marker and (2) its responsiveness to ATRA.

In the 1970s, it was discovered that APL is associated with a disease-specific karyotype. In the great majority of APL patients, abnormal promyelocytes harbored a reciprocal chromosomal translocation t(15;17)(q22;q21) (Rowley et al., 1977). Two
genes involved were subsequently cloned by virtue of the development of molecular biological techniques. The retinoic acid receptor α (RARα) gene on chromosome 17 and the promyelocytic leukocyte (PML) gene on chromosome 15 are fused to generate a PML-RARα chimeric protein (Melnick and Licht, 1999). It is noteworthy that APL patients with t(15;17) exhibit an exquisite sensitivity to ATRA and that complete remission can be induced (Huang et al., 1988). Molecular studies on a large series of APL cases revealed a great heterogeneity of PML-RARα transcripts due to variable breakpoint cluster regions within PML genes as well as alternative splicing (Pandolfi et al., 1992; Tong et al., 1992). Three main PML-RARα isoforms—type L, type S, and type V—were found to exist in APL patients. Clinical data showed that the patients who contained type S PML-RARα had a high likelihood of early death or relapse after being treated with ATRA (Pandolfi et al., 1992).

In addition to the PML-RARα rearrangement, four other APL-associated translocations have been characterized to date at the molecular level. They are t(11;17)(q23;21), t(11;17)(q13;q21), t(5;17)(q35;q21), and del(17)(q11q21), which fuse RARα to promyelocytic leukemia zinc finger (PLZF), nuclear mitotic apparatus (NuMA), nucleophosmin (NPM), and STAT5b genes, respectively, and lead to expression of PLZF-RARα, NuMA-RARα, NPM-RARα, and STAT5b-RARα fusion proteins (Arnould et al., 1999; Chen et al., 1993; Redner et al., 1996; Wells et al., 1997). The leukemic blasts from the patients having variant translocations present similar morphology, but they differ on their sensitivity to ATRA. It was reported that the patients with NuMA-RARα or NPM-RARα could achieve remission after ATRA/chemotherapy, whereas those with PLZF-RARα or STAT5b-RARα had a generally poor response to ATRA (Zelent et al., 2001). Nevertheless, it should be noticed that the breakpoints in the RARα locus of all these translocations are localized in its second intron, and common features of all these fusion proteins are invariable associations with the B through F regions of RARα that contain its DNA and ligand-binding domains. These observations indicated that the translocation-induced impairment of RARα signaling was probably responsible for both leukemogenesis and sensitivity to ATRA.

15.3 RETINOIC ACID SIGNALING PATHWAY IN APL CELL MATURATION

15.3.1 Retinoic Acid and Its Receptors

Retinoic acids (RAs) are a group of derivatives from vitamin A that regulate a wide range of complex physiological events in embryonic development and homeostasis maintenance in vertebrates. Moreover, RAs are capable of triggering differentiation not only in normal cells, but also in cancer cells.

The multiple physiological functions of RAs are generally mediated by their cognate receptors. So far at least two sorts of receptors are found to be able to bind RA: cellular RA binding proteins (CRABPs) and nuclear RA receptors.

The CRABPs are small-size proteins that belong to the β-clamp protein family binding small hydrophobic ligands. It has long been known that CRABPs are
implicated in the biosynthesis, storage, metabolism, and transport of RA. By modulating the concentration of free RA available to the nuclear receptors, CRABPs can indirectly influence the gene expression required for “RA response.” However, several recent studies demonstrated that CRABP II could also be associated with nuclear RA receptors in mammalian cells and that this protein complex could bind the DR5 responsive element of RA-target genes. In the presence of RA, CRABP II could enhance the transactivation of nuclear RA receptors, strongly suggesting a potential role of CRABP II as a novel ligand-dependent transcriptional regulator involved in the RA signaling pathway in eukaryotic cells (Bastie et al., 2001; Delva et al., 1999).

The nuclear receptors for RA include two classes: retinoic acid receptors (RARs) and retinoid X receptors (RXRs). Both belong to the steroid/thyroid hormone receptor superfamily, which act as nuclear ligand-inducible transcription factors. In general, RARs function by heterodimerization with RXRs; the RAR/RXR complex then binds to specific retinoic acid response elements (RAREs) present in the promoters of their target genes. In the absence of RA, the heterodimer binds to the corepressor (CoR) complex containing SMRT, NcoR, and mSin3, which in turn interacts with histone deacetylase, leading to the silencing of target genes. In the presence of RA, the conformation of RAR/RXR dimer changes, resulting in dissociation of the corepressors and recruitment of the coactivators (CoAs), such as CBP/P300, ACTR, and P/CAF, which contain histone acetylase activity. It is believed that the histone acetylation can decondense the chromatin and stimulate the gene transcription (Chambon, 2005; Rosenfeld and Glass, 2001).

### 15.3.2 Dominant-Negative Effect of PML-RARα in APL Cells

In the case of APL, the fusion protein PML-RARα due to t(15;17) retains the DNA and ligand-binding domains. A series of data showed that PML-RARα might impair the RA signaling pathway by binding as homodimer or heterodimer with RXR to RARα target genes in competition with RARα or by binding as homodimers to a novel set of target genes. The identification of some PML-RARα-specific responsive elements further suggested a major gain of function of this fusion protein as a new transcription factor that probably controls the expression of certain genes responsible for APL cell proliferation (Kamashev et al., 2004). Besides, high levels of PML-RARα in APL cells could also sequester RXR and/or some RARα cofactors, acting as a dominant-negative inhibitor against normal RXR function in APL. RXR is not only an important component in the RA-responsive pathway, but also a promising heterodimerization partner for a great number of nuclear receptors such as the vitamin D3 receptor (VDR) and thyroid hormone receptor (TR) (Melnick and Licht, 1999; Rosenfeld and Glass, 2001).

Regarding transcriptional properties, by comparison with wild-type RARα, PML-RARα has an altered ability to bind the corepressors more tightly than wild-type RARα, requiring pharmacological doses of ATRA (10⁻⁶ mol/L) for the release, whereas RARα/RXR releases corepressors only at 10⁻⁹ mol/L ATRA (Lin et al., 2001). This can be one of the reasons of deregulation of some key myeloid genes’
expressions at physiological doses of ATRA, and can account for the need for high
doses of ATRA to induce differentiation in APL. These results are further supported
by the fact that RA combined with sodium butyrate or trichostatin A (TSA),
inhibitors of histone deacetylase, can accelerate APL cell differentiation (Kosugi
et al., 1999).

On the other hand, PML-RAR\alpha can yet exert its dominant effects on wild-type
PML. PML is a phosphoprotein tightly associated with nuclear matrix. More and
more studies indicate that PML has multiple biological activities in cell growth
suppression, apoptosis, and immune response pathways of many cytokines (Melnick
and Licht, 1999). The normal function of PML is closely related with its cellular
localization. The macro-speckled nuclear distribution, referred to as nuclear bodies
(NBs) or PML oncogenic domains (PODs), is one of the most important features of
PML. In normal cells, the NB structure can vary in number, size, and protein
composition depending on cell type, cell cycle, and external hormonal stimulation.
It is indeed a complex of multiproteins. Until now, more than 20 proteins with diverse
natures have been found to exist in PML-associated NB, including Sp100, SUMO-1,
pRB, CBP, NDP55, ISG20, PLZF, Daxx, and p53. This makes the NB hold a multi-
plicity in its function (Zhong et al., 2000a). It has been shown that PML can form not
only homodimers through its RING finger/B-box motif and coiled-coil motif, but also
multimeric complexes with some partners present in the NB. For example, Daxx is
identified as a molecule important for the cytoplasmic transduction of the Fas-
dependent proapoptotic stimulus; it can physically interact with PML in NB. But
such ability of Daxx to induce apoptosis is completely abrogated in PML\textsuperscript{–} cells,
suggesting PML and Daxx cooperate in a novel NB-dependent pathway for apoptosis
(Wang et al., 1998b; Zhong et al., 2000b). Also, the interaction between PML and
CBP may reflect the regulation effect of PML on gene transcription to some extent
(Melnick and Licht, 1999). SUMO-1 (small ubiquitin modifier) is an ubiquitin-like
protein that can covalently modify NB proteins such as PML and Sp100 (Boddy et al.,
1996; Seeler and Dejean, 2001). Some recent studies confirmed that sumoylation of
PML was a prerequisite for the proper formation of the NB, and PML might exert
some of its biological functions through the regulation of NB formation (Muller et al.,

In APL cells, PML-RAR\alpha and PML display an aberrant microspeckled nuclear
pattern, probably due to the ability of PML-RAR\alpha to heterodimerize with PML.
Besides PML, other NB components, such as Sp100, CBP, Daxx, and SUMO-1, are
also found to delocalize to nuclear micro-speckles from their usual sites (Melnick
and Licht, 1999; Zhong et al., 2000a). Such disruption of the NB may explain why the
fusion protein fails to inhibit cell growth and induce apoptosis. ATRA administration
of APL cells can gradually relocalize PML and other NB members into their wild-
type NB configuration. This relocalization is highly correlated with induction of
morphological differentiation in APL cells (Duprez et al., 1996). Mechanistic studies
show that RA-caused PML-RAR\alpha degradation is responsible for the reorganization
of PML-associated NB (Yoshida et al., 1996). The molecular details of this degrada-
tion involve the ubiquitin/proteasome system. Briefly, the AF-2 domain of
RA-activated PML-RAR\alpha could bind the SUG-1 component of the proteasome
19S complex, leading to its catabolism (Zhu et al., 1999a, 2002a). Proteasome
inhibitors can block RA-induced degradation, and PML-RAR\(\alpha\) remains micro-speckled, further confirming the above conclusion. It is also found that the degradation of PML-RAR\(\alpha\) not only restores the normal nuclear distribution of PML-associated NB, but also releases its heterodimerization partner RXR, thus stabilizing RAR/RXR and other nuclear receptor heterodimers required for RXR and recovering the regulation control of gene transcription in favor of APL cell maturation. However, the idea that the degradation of PML-RAR\(\alpha\) and the restoration of NB distribution are indispensable to APL cell differentiation is controversial. Because some reports demonstrate that the RXR-specific ligands can, through the RXR/RXR homodimer, synergize with cAMP to induce maturation of NB4 cells without PML-RAR\(\alpha\) degradation and NB reorganization (Benoit et al., 1999), implying the degradation of fusion protein may be simply a concomitant phenomenon during APL cell differentiation instead of a requirement.

Nevertheless, there is no doubt that the PML-RAR\(\alpha\) exhibits its dominant-negative activity on RAR\(\alpha\) and PML. The results that PML-RAR\(\alpha\) disturbs cell differentiation and apoptosis in hematopoiesis are further reinforced by the observation of in vitro transfection experiments and in vivo transgenic mice studies (Melnick and Licht, 1999; Wang et al., 1998a). Paradoxically, the exquisite sensitivity of APL cells to ATRA likely results from their ability to target PML-RAR\(\alpha\). This observation has been supported by the fact that a mutated PML-RAR\(\alpha\) confers ATRA resistance in APL cells (Duprez et al., 2000) as well as that the leukemias arising in PML-RAR\(\alpha\) m4 mutant transgenic mice do not differentiate in response to RA treatment (Kogan et al., 2000).

### 15.3.3 Cross-Talk Between RA and Other Signaling Pathways

It has long been accepted that cell maturation is a process with great complexity. Although numerous lines of evidences show that the pharmacological doses of ATRA can exert their differentiating effects on APL cells by binding to PML-RAR\(\alpha\), dissociating the nuclear corepressors complex from PML-RAR\(\alpha\) to relieve the transcriptional repression, triggering the degradation of PML-RAR\(\alpha\) oncoprotein and restoring PML-associated NB structure, several reports on the integration of signal cross talk in the context of APL cell differentiation should hold our attention (Benoit et al., 2001).

#### 15.3.3.1 RA and cAMP/PKA

The second messenger, cyclic adenosine monophosphophate (cAMP), plays an important role in the response to hormonal signals for cell proliferation, differentiation, and apoptosis, including that in hematopoietic development. The interplay between RA and cAMP signaling for granulocyte maturation is observed in their synergistic effects on APL cells. Some cAMP agonists can potentiate maturation of NB4 cells, an APL cell line expressing PML-RAR\(\alpha\) (Lanotte et al., 1991), and render the cells responsive even to physiological levels of retinoids, whereas disrupting the cAMP signal pathway with cAMP-PKA (protein kinase A) antagonist will result in a reversible resistance to pharmacological levels of retinoids (Quenech’Du et al., 1998; Ruchaud et al.,
1994), indicating that the cAMP-PKA pathway may be an integrated part of the RA signal network. This possibility is further reinforced by the report of cross-talk between RXR-specific agonist “rexinoid” and PKA signaling (Benoit et al., 1999).

Our recent studies showed a rapid increase in intracellular cAMP levels and PKA activities during ATRA-induced NB4 cell maturation, which was mediated through adenylate cyclase (Zhao et al., 2004). More strikingly, we found that the phosphorylation of RARα by cAMP-PKA signaling could enhance its transcripational response to RA, suggesting the RA receptor as a unique node of cross talk between cAMP and RA pathways. This result was then strengthened by the demonstration of de The’s group that cAMP could also reverse the normal silencing of the transactivating function of RXR when bound to RARα or PML-RARα (Kamashev et al., 2004). Thus, a cytoplasm-to-nucleus pathway can be constructed from the activation of adenylate cyclase to the phosphorylation of RA receptors by PKA, followed by reprogramming the cellular transcriptional machinery. However, the precise mechanism underlying the activation of the adenylate cyclase by ATRA, which is generally considered to be a nuclear hormone, remains unclear. The fact that the cAMP response occurs within a few minutes, which is not abrogated by actinomycin D, is strongly against the conventional concept that this cAMP response is due to the activation of nuclear RA receptors followed by gene transcription/translation regulation and then signaling from nucleus to cytoplasm. Is there any receptor/binding protein for RA existing on the cell surface? As a matter of fact, the activation of pleiotropic cytosolic signaling transduction dependent on rapid nongenomic action of steroid hormones has been indeed reported in the experiments concerning glucocorticoid (Chen and Qiu, 1999) and bile acid (Kawamata et al., 2003; Maruyama et al., 2002). Identification of a G-protein coupled receptor (GPCR) responsive to bile acid strongly suggests that there may be other new GPCRs for nuclear hormones (Im, 2004). The exploration of cytoplasmic-membrane-type receptor/binding protein for ATRA, which is distinct from the known nuclear RA receptors, would be helpful in understanding signal transduction of RA. Meanwhile, it would enrich and complete the theory regarding the mechanism of action of RA and provide interesting possibilities for drug discovery.

### 15.3.3.2 RA and Interferons

Another important signaling cross-talk in APL cell differentiation is characterized by synergism between RA and interferons (IFNs). IFNs are a family of cytokines mainly with tumor growth suppressive and anti-viral activities. By interacting with their cognate receptors, IFNs can activate Jak/STAT pathways through phosphorylation. The activated STAT proteins then translocate to the nucleus where they stimulate the transcription of a set of genes, leading to different signal cascades and various biological effects (Brivanlou and Darnell, 2002). The combination of ATRA and IFNs represents a useful therapeutic approach in the treatment of leukemias and other malignancies (Gallagher et al., 1987; Nason-Burchenal et al., 1996).

It has been found that IFNs and RA can synergistically inhibit growth and induce differentiation in myeloid leukemia cell lines, such as NB4, HL-60, and U937 (Chelbi-Alix and Pelicano, 1999; Gaboli et al., 1998; Garattini et al., 1998). How do
Figure 15.1  An interplay between cell membrane and nuclear signaling in RA/As$_2$O$_3$ therapy for APL.
these two agents cooperate at the molecular level? With the recent efforts made in experimental studies, the mechanisms involved in cross-talk between RA and IFNs are being gradually unveiled: (1) IFNs are found to be capable of modulating the expression of RA receptors in APL cells and thus ultimately control the expression of ATRA-regulated genes, including the levels and the state of activation of members of the Jak-STAT pathway. This process leads to cross-regulation of genes normally under the control of RA and IFNs respectively. (2) Some NB-associated proteins such as PML, Sp100, and ISG20 are targets of IFNs because the promoters of these genes contain ISRE and/or GAS sequences that bind to STAT proteins (Grotzinger et al., 1996; Regad and Chelbi-Alix, 2001; Stadler et al., 1995). IFNs can cooperate with ATRA to induce PML expression through induction and activation of STATs. Since PML is a growth suppressor, it may confer in part the antiproliferative activity of IFNs/RA combination. (3) It is reported that ATRA can induce IFNα synthesis and secretion and increase the expression of STAT1 in NB4 cells, but not in ATRA-resistant cells NB4-LR1 and NB4-LR2 (Pelicano et al., 1997, 1999). Thus, numerous genes that have been isolated as “retinoic acid-induced genes” are indeed IFN-targeted genes, such as the RIG-G gene identified in our previous work from ATRA-induced NB4 cells (Yu et al., 1997). A database search shows that RIG-G, also called IFI60 (de Veer et al., 1998), belongs to an IFN-induced gene family. In fact, while RIG-G mRNA is dramatically upregulated in NB4 cells after 72 h of ATRA treatment, it is obviously induced after 4 h of exposure to IFNα (Xiao et al., 2006). Our current studies demonstrated that overexpression of RIG-G could significantly inhibit cell proliferation by accumulating cells in G1/S transition, probably through upregulation of p27 and p21 levels. We therefore hypothesize that induction of RIG-G might represent one of the key molecular nodes in which the synergistic effect of RA and IFN signaling pathways can be converged in the course of APL cell differentiation.

To sum up, no single signal is likely to work optimally. Responses apparently triggered by a single signal indeed result from cross-talk of diverse signaling pathways, just as in the case of ATRA-induced NB4 cell maturation, in which the potential role of cAMP/PKA or IFNs is masked to some degree (see Fig. 15.1). Besides the above-mentioned signaling cross-talks, there are many other arguments sustaining the notion that multiple signaling pathways are required for hematopoietic differentiation. For example, combination of RA with certain interleukins, G-CSF, vitamin D3, and TPA also generate exploitable synergistic responses (Benoit et al., 2001). The recent finding that Cobalt chloride-mimicking hypoxia favors retinoic acid-induced differentiation suggests a potential link between HIF-1α protein and leukemic cell differentiation (Huang et al., 2003; Jiang et al., 2005). Use of TSA to target chromatin remodeling, or interfering with DNA methyltransferase by azacytidine to restore chromatin structure (Fazi et al., 2005), can potentiate the effect of RA, developing a novel transcriptional/differentiation therapy of leukemia.

15.3.4 Limitations of Retinoic Acid in APL Therapy

Despite the great advantage of ATRA in both high rate of CR and low cytotoxic reaction, certain side effects of ATRA have emerged in APL therapy, mainly those
including retinoic acid syndrome (RAS) during treatment and retinoic acid resistance following treatment.

15.3.4.1 Retinoic Acid Syndrome

Retinoic acid syndrome (RAS), also called hyperleukocytosis, is observed in about one-third of APL patients during treatment with ATRA alone. The major clinical features of RAS combine fever, respiratory distress, weight gain, lower-extremity edema, pleural or pericardial effusions, and sometimes renal failure, frequently preceded by an increase in white blood cell counts in peripheral blood. The pathophysiology of RAS is still not clearly understood and seems to be related to the induction of G-CSF, IL-1β, IL-6, IL-8, and TNFα secretion by APL cells during ATRA treatment. In addition, ATRA-induced modification of the adhesive properties of APL cells might also play a role in RAS. At the present time, prevention of RAS involves using chemotherapy in combination with ATRA and adding high-dose corticosteroids (Degos and Wang, 2001; Fenaux et al., 2001).

15.3.4.2 Retinoic Acid Resistance

Acquired retinoic acid resistance occurs in most, if not all, APL patients who are treated with ATRA alone and achieve CR for only a short time (Degos and Wang, 2001). This is because the differentiation induction therapy is unable to eliminate leukemic cells completely. Furthermore, these residual cells often exhibit unresponsiveness to ATRA, explaining why relapsed patients are usually resistant or refractory to ATRA. In order to reveal the molecular mechanisms of APL cells’ resistance to ATRA, numerous studies were made over the last decade. Several defects at different levels have been described to account for the resistance of APL cells to ATRA:

1. A significant decrease of ATRA levels in plasma due to ATRA-induced upregulation of cytochrome P450 and CRABP II implicated in the metabolism of ATRA (Kizaki et al., 1997), thereby reducing ATRA concentration in the nucleus as well as reducing the sensitivity of APL cells to ATRA (Chen et al., 1996a). However, this deduction is likely doubtful because of the data that the principal ATRA metabolites are as active as ATRA at inducing NB4 cell maturation (Idres et al., 2001, 2002; Marill et al., 2003).

2. No PML-RARα expression or expression of a dominant-negative PML-RARα mutant without ligand-binding activity. In NB4-LR2 cells (Duprez et al., 2000), a truncated PML-RARα due to a nonsense mutation in the E domain is found to block RARα and wild-type PML-RARα transcriptional activities, rendering the cells resistant to ATRA. However, this resistance can be reversed by cooperation of an RXR/RXR-specific pathway and cAMP signaling (Benoit et al., 1999). A similar phenomenon is also observed in other ATRA-resistant cells, such as HL-60R (Robertson et al., 1992) and NB4-306 (Dermime et al., 1993).
3. Increased affinity of fusion protein to nuclear corepressor complex. In a few APL patients with type V PML-RAR\(\alpha\) isoform, some intron 2 sequences of RAR\(\alpha\) gene are found to insert between the truncated PML and RAR\(\alpha\) mRNA fusion partners, providing potential sites for binding CoR so that these patients exhibit less sensitivity to a pharmacological dose of ATRA (Gu et al., 2002). Deletion of these “inserted” sequences can restore the sensitivity to ATRA, suggesting that the aberrant binding of nuclear CoR mediates RA resistance (Gu et al., 2002). This observation is in accordance with the case of APL patients bearing t(11;17), in which the presence of a second CoR binding site on PLZF moiety resists the effect of ATRA (Guidez et al., 1998; He et al., 1998; Hong et al., 1997).

4. Failure of cooperation between ATRA and cAMP signaling pathways. NB4-LR1 cells, because of uncoupling RA and cAMP action, display no maturation in the presence of RA alone. cAMP is required for triggering RA-primed NB4-LR1 cells to undergo terminal maturation (Ruchaud et al., 1994). The fact that the distinct changes of intracellular cAMP concentration in NB4 and NB4-LR1 cells upon ATRA treatment are attributed to the role of adenylate cyclase suggests that the defaults in the activation of adenylate cyclase may contribute to the resistance of NB4-LR1 cells to ATRA-induced maturation (Zhao et al., 2004).

5. Deregulation of telomerase activity in APL cells probably involved in an eventual novel mechanism of retinoic acid resistance (Pendino et al., 2002).

Since the retinoic acid resistance is inevitable and the mechanism involved is so complex, it is necessary to explore new approaches for APL therapy. Currently, intravenous administration of As\(_2\)O\(_3\) has become the first-line approach of therapy for APL patients with relapses or refractory disease.

### 15.4 ARSENIC TRIOXIDE IN THE TREATMENT OF APL

#### 15.4.1 Clinical Practice of Arsenic in APL Therapy

Arsenic has long been known to act as a carcinogen, involved in human skin, lung, liver, and kidney tumors. Paradoxically, it has also been demonstrated to contain antitumor activity in some cases. In traditional Chinese medicine, arsenic compounds have been used for more than 500 years in the treatment of asthma and certain skin diseases. Use of arsenic for the treatment of chronic myeloid leukemia can be traced to the middle of the nineteenth century. But chemotherapeutic agents later replaced this treatment because of the toxicity of arsenic (Zhang et al., 2001; Zhou et al., 2005). In the 1970s, the discovery of the therapeutic effect of As\(_2\)O\(_3\) in APL revived this ancient drug. A Chinese group reported in 1992 for the first time that a solution containing As\(_2\)O\(_3\) and a trace mount of mercury chloride could induce CR in APL patients; the CR rate, the 5-year survival rate, and the 10-year survival rate were 65.5%, 50.0%, and 18.8%, respectively (Sun et al., 1992). In 1996, SIH was able to conduct a clinical trial using a pure As\(_2\)O\(_3\) solution at a dose of 0.16 mg/kg/day to treat
a group of relapsed APL patients previously treated with ATRA and chemotherapy. A complete remission was obtained after 28–54 days among 14 out of 15 cases (Chen et al., 1996b; Shen et al., 1997). This suggested that As$_2$O$_3$ displayed no obvious cross-resistance with ATRA. Many groups from China and western countries then confirmed these results (Dombret et al., 2002; Niu et al., 1999; Soignet et al., 1998). Morphologically, after 3–4 weeks of As$_2$O$_3$ treatment, the leukemic promyelocytes decrease, while a lot of myelocyte-like cells and degenerative cells appear in the bone marrow and peripheral blood of patients (Chen et al., 1997). Furthermore, recent studies on the clinical efficacy of low-dose As$_2$O$_3$ demonstrated that low-dose As$_2$O$_3$ seemed to have the same effect as the conventional dosage (Shen et al., 2001). The major side effects of arsenic, including abnormal increase in hepatic enzymes, hyperleucocytosis, skin reaction, and gastrointestinal disturbances, are moderate and tolerable and are less frequent in low-dose-treated patients. No significant bone marrow suppression is so far observed. These results suggest that arsenic is an effective and relatively safe drug in APL treatment, especially for relapsed and refractory patients (Shen et al., 2001).

15.4.2 Molecular and Cellular Mechanisms of Arsenic in APL Therapy

The clinical success of As$_2$O$_3$ in APL treatment has been stimulating the research on its mechanisms of action. Many significant advances in this regard have been made during the past few years.

The ability of As$_2$O$_3$ to induce APL cells to undergo apoptosis was first disclosed in some in vitro studies. It was found that 1 μM of As$_2$O$_3$ could trigger apoptosis of NB4 cells with downregulation of Bcl-2 gene expression (Chen et al., 1996b). Mechanistic studies showed that this As$_2$O$_3$-induced apoptotic effect was associated with the collapse of mitochondrial transmembrane potentials due to thiol-group oxidation and with the activation of caspase 3 (Cai et al., 2000). It was further found that As$_2$O$_3$ could also modulate the microspeckled pattern of PML nuclear distribution in NB4 cells by degradation of PML-RAR$\alpha$ (Zhu et al., 1997). Yet, unlike retinoic acid, which basically targets the RAR$\alpha$ moiety of the fusion protein, As$_2$O$_3$ targets its PML moiety. As$_2$O$_3$-induced PML sumoylation on a specific lysine residual within PML moiety is required for the recruitment of the 11S proteasome regulatory complex, which is responsible for PML-RAR$\alpha$ degradation (Lallemand-Breitenbach et al., 2001). It is now considered that As$_2$O$_3$ exerts its competence of apoptosis induction in APL cells through rapidly degrading PML-RAR$\alpha$. In fact, in contrast to ATRA, As$_2$O$_3$ can degrade both PML-RAR$\alpha$ fusion protein and wild-type PML (Zhu et al., 2001).

Aside from apoptosis-inducing activity, dose effect of As$_2$O$_3$ indicates that at low concentration (0.1–0.25 μM) it can also induce a certain degree of differentiation of NB4 cells (Chen et al., 1997). However, this differentiation is far from terminal because most cells cannot mature beyond the myelocyte–metamyelocyte stage and the NBT reduction test remains negative. Moreover, this partial differentiation can be observed only after long (>10 days) exposure of the cells to the drug. Since the low-dose As$_2$O$_3$ exerts no obvious effect on modulation of the interaction between CoR

SOFTbank E-Book Center Tehran, Phone: 66403879,66493070 For Educational Use.
and RAR/RXR or PML-RARα, it is unlikely that the partial differentiation induction of low-dose As$_2$O$_3$ in NB4 cells is mediated directly through the RARα pathway. The mechanisms underlying this As$_2$O$_3$-induced incomplete or partial differentiation remain to be explored.

On the other hand, pharmacokinetics analysis of APL patients receiving As$_2$O$_3$, whether in standard doses or low doses, reveals that the plasma As$_2$O$_3$ level is rapidly eliminated after a peak level superior to 1 μM caused by intravenous drip. During the treatment, the plasma As$_2$O$_3$ levels fluctuated mostly between 0.1 and 0.5 μM, and the continuous administration of As$_2$O$_3$ did not alter its pharmacokinetic behaviors (Shen et al., 1997, 2001). According to this finding and the morphological observation on bone marrow smears of As$_2$O$_3$-treated APL patients, it can be thus speculated that APL cells are in vivo primarily exposed to a drug concentration in favor of differentiation rather than apoptosis. Moreover, the apoptotic effect of As$_2$O$_3$ can be observed in a large range of cell types, covering not only leukemic cells but also some solid tumor cells, whereas the differentiating activity is relatively limited to APL (Chen et al., 2002; Zhu et al., 1999b).

The discrepancy between the differentiation states observed in vivo and in vitro suggests some differentiating factors present in bone marrow microenvironments of APL patients cooperating with As$_2$O$_3$ to induce APL cell maturation in vivo, including cytokines and other signal molecules. A significant increase in IL-1β and G-CSF expression is observed in arsenic-treated APL cells (Jiang et al., 2003). Our recent data showed existence of a strong synergy between cAMP and the low-dose of As$_2$O$_3$ in fully inducing differentiation of NB4, NB4-LR1, and fresh APL cells (Zhu et al., 2002b). We found that cAMP could facilitate the As$_2$O$_3$-induced degradation of PML-RARα and significantly induce G1 arrest and cell-cycle exit of NB4 and NB4-LR1 cells by modulating several major players in G1/S transition regulation. The cooperation might result from the concomitant PML-RARα degradation induced by As$_2$O$_3$ and the cell growth inhibition by cAMP. Our findings concerning this novel signaling cross-talk of As$_2$O$_3$ and cAMP was soon confirmed by the demonstration that in vivo activation of cAMP signaling was beneficial in mouse models of APL as well as in an RA/As$_2$O$_3$-resistant APL patient (Guillemin et al., 2002). These results may deepen our understanding of As$_2$O$_3$-induced differentiation in vivo and increase the potency of other differentiation inducers in vivo.

### 15.4.3 Dual Treatment with ATRA and Arsenic Trioxide in APL Therapy

Though the reports on the combined effects of ATRA and As$_2$O$_3$ in APL cell lines were conflicting (Gianni et al., 1998; Jing et al., 2001; Shao et al., 1998), our latest clinical data showed that the ATRA and As$_2$O$_3$ combination for remission/maintenance therapy of APL manifested much better results in newly diagnosed APL than either of the two drugs used alone in terms of the high quality of remission and the status of the disease-free survival (Shen et al., 2004). This observation was reminiscent of the results previously obtained in mice transplanted with PML-RARα transgene cells—that the association of ATRA with As$_2$O$_3$ could lead to a much faster decrease in the leukemic population and cure the induced leukemia through...
enhanced differentiation and apoptosis (Lallemand-Breitenbach et al., 1999). In order to understand the complexity and dynamics of in vivo synergy between RA and As$_2$O$_3$ in APL, we identified a number of pathways with temporospatial relationships by conducting an investigation incorporating advanced technologies of genomics, proteomics, and computational biology (Zheng et al., 2005). These pathways are regulated in concert and harmony, forming a network. It seems that RA exerts its effects on APL cells mainly through nuclear-receptor-mediated transcriptional regulation, whereas As$_2$O$_3$ exercises its impact through targeting multiple pathways/cascades at the levels of proteome, transcriptome, and probably metabolome. The finding of a strong activation of the ubiquitin/proteasome system strongly suggests that the synergistic effects of ATRA and As$_2$O$_3$ on degradation of PML-RAR$\alpha$ oncoprotein and on differentiation/apoptosis contribute to the mechanism of superior efficacy of RA/As$_2$O$_3$ dual treatment in APL. Although the protocol of ATRA/As$_2$O$_3$ combination in APL treatment still needs to validate and the long-term effect of this treatment on APL patients remains to be illustrated, such an inspiring discovery will furnish insights for new therapeutic strategies in the treatment of APL. With the ATRA/As$_2$O$_3$ dual treatment followed by obligatory consolidating therapy for maintenance, it is worth expecting that APL will become a curable disease in the near future.

15.5 CONCLUSIONS

Altogether, in the case of APL, PML-RAR$\alpha$ fusion protein due to a specific translocation t(15;17) has been considered as a dominant-negative receptor that blocks myeloid differentiation most likely through the impairment of retinoid response. In order to restore retinoid-signaling pathways, two strategies can be taken into account: (1) a targeted degradation or inactivation of PML-RAR$\alpha$ and (2) an activation of some alternative signaling pathways bypassing PML-RAR$\alpha$. During the last two decades, studies on the mechanisms of differentiation and apoptosis induction therapy of APL have significantly impacted understanding the pathogenesis of APL and identifying new approaches in the treatment of cancer. It is now recognized that PML-RAR$\alpha$ fusion protein is a molecule basis of both APL pathogenesis and therapy. ATRA and As$_2$O$_3$, though the mechanisms and the sites of action of the two drugs are widely divergent, both target the same fusion protein PML-RAR$\alpha$. Accordingly, it should be considered that the APL treatment, whether with ATRA, As$_2$O$_3$, or both, is a PML-RAR$\alpha$ oncoprotein targeted therapy. Establishment and realization of molecule target-based cancer therapy in another leukemia model beyond APL is our desire and research orientation.

ACKNOWLEDGMENTS

The authors were supported in part by the Chinese National Key Basic Research Project 973 (2002CB512805, 2004CB518600), the Chinese National High Tech Program 863 (2002BA711A04, 2004AA227050), the National Natural Science Foundation of China (90209007, 30271466, 30570778, 30670882), the Shanghai Municipal Commission for
Science and Technology, and the Shanghai Municipal Commission for Education. Jian-Hua Tong was also supported by the “Shu Guang” Program of Shanghai Municipal Commission for Education. The authors appreciate Professor Zhen-Yi Wang and Professor Zhi-Xiang Shen from SIH, and Dr. Michel Lanotte from Hospital St. Louis, Paris, France, for their continuous support. The authors are also grateful to Dr. Ji Zhang for help in reviewing the manuscript.

REFERENCES


trans


REFERENCES


16.1 INTRODUCTION

The discovery and development of effective chemotherapeutic agents to treat patients with cancer have been major endeavors of academic, government, and industrial laboratories for more than 50 years. Without question, the greatest progress to date in chemotherapy has been made against the leukemias and lymphomas, although significant advances are being made in the solid tumor areas as well. The ultimate goal of treatment approaches in cancer chemotherapy is the elimination of cancer cells while sparing normal cells. A better understanding of how fundamental cancer biology can be exploited to effectively eliminate cancer cells has emerged with the concept that cancer is more complex than simply a disorder of cell growth regulation, but involves the regulation of cell death and critical factors in the tumor–host microenvironment. The majority of drugs developed for cancer chemotheraphy have been DNA damaging agents or those that inhibit cell division and consequently work best in rapidly dividing

*Corresponding author: dcarson@ucsd.edu

Drug Discovery Research: New Frontiers in the Post-Genomic Era, Edited by Ziwei Huang
Copyright © 2007 John Wiley & Sons, Inc.
cancers. However, many tumors have large fractions of cells that are not dividing for extended periods and are thus difficult to treat selectively with these cell-cycle-specific agents. In these tumor types, the most effective agents will be those that can destroy the cancer cells whether they are actively dividing or resting. The successful development of such a drug, 2-chlorodeoxyadenosine (2CDA) for lymphoproliferative diseases, is a specific example of this concept and is the subject of this review.

The discovery of 2CDA as a useful chemotherapeutic agent is unique in that it was not the result of the random screening of large numbers of candidate compounds or high-throughput screening of compound collections such as combinatorial libraries or a member of a large virtual screening database of compounds. Rather, it was the product of a rational attempt to exploit the biochemical basis for lymphocyte depletion in a rare immunodeficiency disease in children caused by the functional lack of the enzyme adenosine deaminase (ADA). The development of 2CDA resulted from the insights into the mechanism by which ADA deficiency leads to immunodeficiency (Carson et al., 1977). The key observation was that children deficient in ADA due to an inborn error of metabolism suffered from severe combined immunodeficiency disease (Giblett et al., 1972; Ochs et al., 1973) and that this was caused by an accumulation of 2′-deoxyadenosine triphosphate (dATP) to levels that proved to be lethal to lymphocytes (Cohen et al., 1978). This lymphospecific toxicity in ADA deficiency results from the high deoxycytidine kinase (dCK) to 5′-nucleotidase ratio found in lymphocytes that favors the production and accumulation of the phosphorylated forms of deoxyadenosine (Carson et al., 1983; Kawasaki et al., 1993). Thus, the possibility became attractive that one might exploit this newfound sensitivity of lymphocytes to dATP exposure for applications in lymphoproliferative and autoimmune diseases. This insight served as a foundation for two different therapeutic strategies: (1) direct pharmacological inhibition of ADA and (2) use of an analogue of deoxyadenosine that might mimic the lethal effects observed in ADA deficiency. An example of the first strategy is the development of ADA inhibitors such as deoxycoformycin, which has met with limited success (Escoda et al., 1990; Keating 1993). The second strategy requires a consideration of the structure–activity aspects of possible candidate compounds in that they should be (1) a close analogue of deoxyadenosine and serve as a substrate for activation, (2) stable to degradation by functional ADA, and (3) stable to phosphorolysis by cellular purine nucleoside phosphorylase. 2CDA was found to be an excellent candidate in every respect. Thus, the simple replacement of the hydrogen by a chlorine atom at the 2-position of the adenine ring in 2CDA renders this nucleoside resistant to degradation by ADA while maintaining its substrate activity for phosphorylation by deoxycytidine kinase. It was this ADA resistance that initially qualified 2CDA to be included with other nucleoside analogues, known to be deaminase-resistant, for evaluation in vitro against a number of malignant cell lines and in vivo against L1210 in mice in hopes of finding analogues more potent and selective than the endogenous dATP. 2CDA was thus identified as the most potent among the initial group of 25 deoxyadenosine analogues tested (Carson et al., 1980).
16.2 CHEMISTRY AND METHODS OF PREPARATION

2CDA is 2'-deoxyadenosine that bears a chlorine atom at the 2-position (Fig. 16.1). 2CDA has been prepared by a number of methods involving (1) transformation of preformed nucleosides and (2) glycosylation procedures, both chemical and enzymatic. A brief description of these methods is presented here.

16.2.1 Classical Synthetic Methods

2CDA was first reported in 1963 as an intermediate in the synthesis of 2'-deoxyadenosine via a “cyclonucleoside” of adenosine (Ikehara and Tada 1963, 1965), an example of transformation of a preformed nucleoside. Early glycosylation procedures used in deoxyribonucleoside synthesis include the coupling under fusion conditions of a suitably protected deoxyribose with a versatile aglycone, as well as solution phase procedures (Ts’o, 1974; Walker et al., 1979). The first reported synthesis of 2CDA utilizing a glycosylation strategy was accomplished via fusion coupling of 2,6-dichloropurine and 1,3,5-tri-O-acetyl-2-deoxy-D-erythro-pentofuranose followed by amination at the 6-position with concomitant deprotection of the sugar ester groups as shown in Fig. 16.2 (Christensen et al., 1972). This glycosylation procedure is not stereospecific and unavoidably results in the formation of anomers of both α- and β-configuration, which must be separated to provide the desired β-anomeric product. Thus, the yield of the N-9 β-anomer is low, usually less than 20% overall for both steps.

16.2.2 Modern Stereospecific Methods

In view of the difficulties encountered by the formation of anomeric mixtures in the classical glycosylation procedures, we developed a stereospecific method in which 1-chloro-2-deoxy-3,5-di-O-toluoyl-D-erythro-pentofuranose (Hoffer, 1960) is reacted with the sodium salt of the chloroheterocycle, in this case 2,6-dichloropurine, resulting in the formation of the desired N-9 β-isomeric product in 59% yield (Kazimierczuk et al., 1984) (Fig. 16.3). No α-anomeric product was observed. This

Figure 16.1 Structure of 2CDA. 2CDA is 2'-deoxyadenosine substituted with a chlorine at the 2-position.
Figure 16.2 Classical preparation of 2CDA starting with 2,6-dichloropurine and fusing with esterified 2-deoxy-\(\alpha\)-ribose, followed by amination-deprotection with methanolic ammonia.

Figure 16.3 Modern stereospecific preparation of 2CDA utilizing the sodium salt of 2,6-dichloropurine and a 1-\(\alpha\)-chloro-esterified 2-deoxy-\(\alpha\)-ribose in an \(S_N\)2 reaction.
stereospecificity was achieved due to the fact that the starting chloro sugar has the α-configuration in the solid state, and the exclusive formation of the protected β-nucleosides is viewed as the result of a direct Walden inversion ($S_N2$) at the C_1 carbon by the anionic heterocyclic nitrogen. Thus, the incoming heterocycle approaches from the β-face of the sugar as the chloro group is leaving from the α-face, resulting in the nucleoside product having the desired β-configuration. This strategy is still used today for the manufacture of clinical grade 2CDA.

Variations of this sodium salt procedure have recently been used to prepare 2CDA under phase transfer conditions (Kazimierczuk and Kaminski, 2000) and from 2-chloroadenine directly (Gerszberg and Alonso, 2000). Other recent methods include the transformation of 2'-'deoxyguanosine (Jana et al., 2003) and heterocyclic modifications to avoid the formation of unwanted N-7 isomeric product (Gupta and Munk, 2004).

16.2.3 Enzymatic Methods

The first human clinical trials were conducted using 2CDA that had been prepared by enzymatic methods. Thus, Carson et al. (1980) prepared 2CDA by transglycosylation using a bacterial (Lactobacillus helveticus) enzyme with 2-chloroadenine as a substrate and thymidine as the deoxyribose donor. Later, other enzymatic transglycosylations were reported, for example, with 2-chloroadenine as substrate and using whole cells of E. coli BMT-1D/1A as the biocatalyst and 2'-deoxyguanosine as the glycosyl donor (Mikhailopulo et al., 1993), as well as using 2'-deoxyadenosine as a co-substrate with 2-chloroadenosine in an E. coli system (Blank et al., 1992). In addition, a purine nucleoside phosphorylase enzyme system that utilizes 1-phosphorylated sugars and nucleobase derivatives has been reported (Komatsu et al., 2001).

16.3 PHARMACOLOGY

Although 2CDA was originally described in 1963, no report of its biological properties appeared until 1970 (Simon et al., 1970), wherein Robins and co-workers showed it to be a weak inhibitor of calf intestine ADA. As mentioned, the presence of a chlorine atom at the 2-position prevents 2CDA from being deaminated by ADA. In addition, 2CDA is not a substrate for mammalian purine nucleoside phosphorylase (Bzowska and Kazimierczuk, 1995), the enzyme that catalyzes the cleavage of the sugar moiety from the nucleobase. This resistance to enzyme-catalyzed degradation thus allows for the subsequent formation of cellular intermediates essential for biological activity.

16.3.1 Mechanism of Drug Activation

Following transport into the cell, 2CDA must be phosphorylated eventually to the triphosphate form to be biologically active. Phosphorylation of 2CDA to its monophosphate is mediated by deoxycytidine kinase (Carson et al., 1980) and mitochondrial
deoxyguanosine kinase (Arner, 1996) while dephosphorylation is catalyzed by cytosolic 5'-nucleotidase. Cells with a high dCK to 5'-nucleotidase ratio are particularly sensitive to deoxyadenosine analogue toxicity, including 2CDA and fludarabine, while cells with lower ratios are resistant to these agents (Dumontet et al., 1999). Transport of 2CDA into the cell by one of several specific nucleoside transporter systems, although not strictly an activation mechanism, is essential for subsequent activation to take place (Beck et al., 2000). Nucleoside permeation across mammalian cell membranes is complex with at least four distinct transporters known. Two of these (es and ei) are equilibrative (facilitated diffusion) carriers that have been studied in considerable detail. The other two (cif and cit) are concentrative, Na\textsuperscript{+}-dependent carriers. Studies appear to indicate that 2CDA is transported into the cell by a concentrative cif carrier while the equilibrative es carrier is responsible for efflux of the drug (Crawford et al., 1990).

### 16.3.2 Mechanism of Action

The cytotoxicity of 2CDA mainly depends on the accumulation of 2CDA-triphosphate after phosphorylation of 2CDA-monophosphate by nucleoside monophosphate kinase and nucleoside diphosphate kinase. The triphosphate form of 2CDA has a number of activities. In dividing cells, it inhibits ribonucleotide reductase (Griffig et al., 1989) and DNA polymerase \(\alpha\). It can be incorporated into DNA resulting in impairment of DNA synthesis (Chunduru et al., 1993). In resting cells, DNA single strand breaks occur and accumulate since cells are unable to properly repair their DNA due to the inhibition of DNA polymerization and ligation by excess 2CDA-triphosphate (Seto et al., 1985). The enzyme poly (ADP-ribose) polymerase is activated, which then results in depletion of NAD and ATP pools and disruption of cellular metabolism. At the same time, Bcl-2 (the oncogene known to enhance cell survival) is down-regulated (Petersen et al., 1996). The scene is thus set for cell death by apoptosis. The molecular details of the 2CDA-mediated apoptotic process in chronic lymphocytic leukemia (CLL) cells have recently been reported from our laboratory (Leoni et al., 1998; Genini et al., 2000). A scheme depicting the activation of the apoptotic pathway by deoxyadenosines such as 2CDA is shown in Fig. 16.4.

The deoxynucleotide analogue, in this case 2CDA, enters the cell and is converted progressively to its active 2CDATP. The 2CDATP causes DNA strand break formation, activating poly (ADP-ribose) polymerase and p53, with resultant depletion of NAD and adenine nucleotides and a concomitant increase in oxidative stress. In untreated cells, the concentrations of ATP and cytochrome \(c\) in the cytoplasm are insufficient to trigger the caspase cascade. However, the binding of APAF-1 to caspase-9 in the presence of 2CDATP and small amounts of released cytochrome \(c\) leads to the cleavage of caspase-3, converting it to an active autocatalytic protease in a process analogous to the blood-clotting and complement cascades. The active caspase-3 then stimulates, in turn, the caspase-activated DNase (CAD) endonuclease, which irreversibly degrades DNA. Moreover, we found that ADP appears to act as a physiological intracellular inhibitor of the cytochrome \(c\) and APAF-1-mediated caspase pathway in both normal and CLL cells (Genini et al., 2000).
Thus, 2CDA is distinguished from other chemotherapeutic agents that affect purine metabolism in that it is cytotoxic to both actively dividing and quiescent lymphocytes and monocytes, inhibiting both DNA synthesis and repair.

### 16.3.3 Preclinical Evaluation

Preclinical studies showed that both B- and T-lymphoblastoid cell lines were sensitive to 2CDA and that prolonged exposure of resting peripheral blood lymphocytes to this agent in vitro resulted in greater lymphotoxicity than did brief incubations (Carson et al., 1984). This observation led to the selection of a continuous intravenous infusion schedule for the initial clinical trials of 2CDA. Reviews of the preclinical and clinical pharmacokinetic studies of 2CDA using various schedules and routes of administration have appeared (Saven and Piro, 1994; Liliemark, 1997). The pharmacokinetics of 2CDA best fit a two-compartment, first-order-elimination model. Recently, detailed \textit{in vivo} metabolism studies of 2CDA in mice, monkeys, and humans have been reported (Wu et al., 2004).

Figure 16.4  Mechanism of action of 2CDA: Activation of the apoptotic pathway. 2CDA enters the cell and is converted to its active 2CDATP. The 2CDATP causes DNA strand break formation, activating poly(ADP ribose) polymerase and p53, with resultant depletion of NAD and adenine nucleotides and a concomitant increase in oxidative stress. In untreated cells, the concentrations of ATP and cytochrome \textit{c} in the cytoplasm are insufficient to trigger the caspase cascade. However, the binding of APAF-1 to caspase-9 in the presence of 2CDATP and small amounts of released cytochrome \textit{c} leads to the cleavage of caspase-3, converting it to an active autocatalytic protease. The active caspase-3 then stimulates, in turn, the caspase-activated DNase (CAD) endonuclease that irreversibly degrades DNA.
16.4 CLINICAL ACTIVITY

Relative to traditional antimetabolites, 2CDA is unique in that it is equally active against dividing and resting lymphocytes, which is essential for the observed activity of 2CDA in indolent lymphoproliferative disorders. However, 2CDA has also shown significant activity in certain solid tumors and autoimmune conditions.

16.4.1 Hematologic Malignancies

Of the lymphoproliferative conditions that have been treated with 2CDA, hairy cell leukemia (HCL) has shown the most dramatic response. It is administered over a 7-day period by intravenous infusion at a dose of 0.09 mg/kg. Overall response rates in clinical studies have ranged from 80% to 100%, with a large majority of these being complete remissions. 2CDA is approved by the U.S. FDA for the treatment of active HCL and is indicated for symptomatic patients, those with significant neutropenia, leukocytosis, anemia, thrombocytopenia, splenomegaly, constitutional symptoms due to HCL, recurrent serious infections, and painful lymphadenopathy. A review of the guidelines for clinical use of 2CDA in HCL has recently appeared (Goodman et al., 2003). 2CDA also has activity in CLL (Juliusson et al., 1992; Saven, 1995), non-Hodgkin lymphoma (Kay et al., 1992; Hickish et al., 1993), Waldenstrom macroglobulinemia (Dimopoulos et al., 1993), cutaneous T-cell lymphoma (Saven et al., 1992), acute myelogenous leukemia (AML) (Carson et al., 1984; Santana et al., 1992), and chronic myelogenous leukemia (CML) (Beutler, 1993; Saven et al., 1994).

A particularly difficult-to-treat lymphoproliferative disorder is CLL, a disease caused by the proliferation and accumulation of B lymphocytes. It is the most common form of leukemia in North America and Europe. The alkylating agent chlorambucil, with or without prednisone, is the standard initial treatment (Foon et al., 1990). 2CDA has also been used as initial therapy for CLL patients. Patients who received purine nucleoside analogues, such as 2CDA or fludarabine, as their initial therapy and achieved long-lasting response can be successfully retreated with the same agent. Purine nucleoside analogues administered in combination with other chemotherapeutic agents and/or monoclonal antibodies may produce higher response rates, including complete response, compared with purine nucleoside analogs alone or other treatment regimens (Robak, 2005a,b). A study of the activity of 2CDA was conducted in 90 CLL patients, mostly Binet stage C disease who had failed alkylator therapy, and resulted in 4% complete responses and 40% partial responses (Saven, 1996). Here, the ratio of dCK to cytoplasmic 5'-nucleotidase was predictive of 2CDA responsiveness (Kawasaki et al., 1993). In combination studies in fludarabine-resistant CLL patients, pretreatment with bryostatin 1 has been shown to potentiate the efficacy of 2CDA by increasing the ratio of dCK to 5'-nucleotidase activity (Ahmad et al., 2000; Beck et al., 2004). However, a study in 20 previously untreated CLL patients who were administered 2CDA as a 7-day continuous infusion for four courses achieved 25% complete responses and 60% partial responses for an overall response rate of 85% (Saven et al., 1995). In another study, oral 2CDA was evaluated in 126 previously untreated CLL patients and compared with 40 previously treated patients (Karlsson et al., 2002). The overall responses were 77% and 39% in these conditions.
groups, respectively. Interestingly, in these studies the achievement of complete remission was not a prerequisite for long-term survival. Combination preclinical and clinical studies using 2CDA with other cytotoxic agents and/or monoclonal antibodies show synergistic actions (Robak et al., 2005a,b). These and other studies establish the significant activity of 2CDA, used alone or in combination with other agents, in patients with previously treated and untreated CLL (Robak, 2005a,b).

Patients with low-grade non-Hodgkin lymphoma have been responsive to 2CDA as a single agent (Kay et al., 1992; Hickish et al., 1993) and in combination with mitoxantrone (Saven et al., 1996; Rummel et al., 2002) or cyclophosphamide (Van Den Neste et al., 2000; Robak et al., 2001). 2CDA is also active in advanced, untreated low-grade lymphomas, but myelosuppression is a cumulative adverse effect and limits the number of courses of treatment that can be given (Canfield et al., 1997).

In a study of 15 patients with cutaneous T-cell lymphoma treated with 2CDA after having failed standard topical and/or systemic therapy, the overall response rate was 47%. Three of 15 patients (20%) achieved complete responses, and four of 15 patients (27%) achieved partial responses (Saven et al., 1992). The median duration of response was 5 months.

Several studies conducted with 2CDA as a single agent in patients with Waldenstrom macroglobulinemia, a B-cell neoplastic condition, showed good overall response rates from 55% to 100%, with the higher response rates noted in the previously untreated patient groups (Dimopoulos et al., 1993; O’Brien et al., 1996; Fridrik et al., 1997; Liu et al., 1998).

Modest activity has been seen with 2CDA in myeloid leukemias. In CML, one study showed an overall response rate of 47% in patients previously treated with combination therapies (Dann et al., 1998). In another study, 2CDA was also active in chronic-phase CML, inducing complete hematologic responses in these patients, but the absence of cytogenetic responses and severe immunosuppression limit its clinical use (Saven et al., 1994). In AML, 2CDA has single-drug activity and may enhance the formation of the active metabolite (ara-CTP) of cytosine arabinoside (ara-C) in combination (Juliusson et al., 2003). In fact, this combination along with idarubicin showed an overall complete response rate of 62% with 30% long-term survival (>2 years). In relapsed/refractory AML, however, 2CDA is ineffective (Gordon et al., 2000). The mechanism of activity in myeloid leukemias is unclear since only low levels of dCK are expressed in granulocytes (Carson et al., 1980, 1983).

In other studies, 2CDA was devoid of activity in patients with multiple myeloma (Dimopoulos et al., 1992) but has shown promise in the treatment of Langerhans histiocytosis (Dimopoulos et al., 1997) and B-cell prolymphocytic leukemia, inducing high overall and complete response rates (Saven et al., 1997).

### 16.4.2 Solid Tumors

Because 2CDA and other similar purine nucleoside analogs are active antimetabolites in both proliferating and resting lymphoid cells, it was thought that these agents might have significant activity against certain solid tumors as well. A dose-escalation study of 2CDA in patients with various solid tumors was conducted to determine the maximum-tolerated dose (MTD) and define its toxicity profile at higher doses (Saven...
et al., 1993). In these studies, two of seven patients (28.6%) with malignant astrocytomas obtained partial responses with a median duration of 8 months. Interestingly, there was no clinical correlation with dCK expression and response to 2-CdA. Another phase 1 clinical trial was conducted to determine the MTD in 18 patients with solid tumors (non-small-cell lung and colorectal) given 2CDA by 1-hour intermittent infusion, repeated daily for 5 days. No responses were observed (Kobayashi et al., 1994). There is some in vitro evidence that 2CDA could be useful against hepatocellular carcinomas (Graziadei et al., 1998), but no clinical trials for this solid tumor type have been reported. With limited exceptions, 2CDA as a single agent has not shown significant activity in solid tumors.

### 16.4.3 Multiple Sclerosis and Other Nonmalignant Diseases

It is not surprising, given the antilymphocyte activity of 2CDA, that the use of this agent to suppress an inappropriate immune response has been investigated. Multiple sclerosis (MS) is characterized by clinical signs and symptoms of CNS demyelination. The disease can be chronic and progressive and have relapsing and remitting stages. Although MS is thought to be a lymphocyte-dependent autoimmune disease, the specific antigens and triggering agents involved in the disease are unknown. The rationale for the use of 2CDA in MS is that autoantigen-specific T-lymphocytes are thought to be activated peripherally before migrating to the CNS, where they mediate damage to myelin (Hafler and Weiner, 1989). This suggests that 2CDA-induced reduction in the number of lymphocytes may help slow the progression of the disease. The use of 2CDA in MS has been studied, and its potential use in this disease has been reviewed (Langtry and Lamb, 1998). In a 2-year placebo-controlled, double-blind, crossover study to evaluate 2CDA in the treatment of chronic progressive multiple sclerosis, analysis of the results revealed a favorable influence on the neurological performance scores in both the Kurtze extended disability status and the Scripps neurological rating scale and also on MRI findings in patients treated with 2CDA. In the first year, the most striking finding was that while clinical deterioration continued in the placebo-treated patients, the condition of patients who received 2CDA stabilized or even improved slightly (Beutler et al., 1996). In another study, a double-blind, placebo-controlled, randomized trial of 2CDA in relapsing-remitting MS was conducted in which 2CDA given by subcutaneous injection had a favorable effect on the frequency and severity of relapses and MRI findings. MRI-enhancing lesions were completely suppressed in the 2CDA patients by the sixth month of treatment (Romine et al., 1999). Thus, 2CDA has shown promise in the treatment of both progressive and relapsing-remitting MS.

In addition to MS, other nonmalignant disorders have responded to 2CDA treatment (Beutler et al., 1997a,b). Among them, autoimmune hemolytic anemia seems to respond to 2CDA in some patients but may be induced in others, usually in those with concomitant CLL after several courses of treatment (Robak et al., 1997), but has also been reported in some patients with Waldenstrom macroglobulinemia (Tetreault and Saven, 2000). The finding that human monocytes were as sensitive to 2CDA toxicity as lymphoid cells suggested that these analogs could offer a novel therapeutic strategy for
chronic inflammatory and autoimmune diseases characterized by inappropriate mono-
cyte deployment or function (Carrera et al., 1990). Rheumatoid arthritis (RA) and
inflammatory bowel disease are two inflammatory diseases for which 2CDA therapy
has shown some efficacy, although 2CDA is clearly not an anti-inflammatory agent.
Combined analysis of three studies involving 23 patients with longstanding RA who
were given either continuous or 4-hour infusions, or were given oral 2CDA, indicated
that more than half of the patients had >20% improvement in joint swelling during
2CDA treatment (Beutler et al., 1997a,b). The effect of the drug on pain and tenderness
was less dramatic. Collectively, these data support the hypothesis that 2CDA may
retard the pannus formation in patients with longstanding RA. Whether or not 2CDA
can decrease joint erosions in RA patients remains to be established.

Primary sclerosing cholangitis (PSC) is a chronic hepatic autoimmune disorder
of unknown etiology, thought to be mediated by biliary autoreactive cytotoxic lym-
phocytes. A study was conducted in four patients with stages I and II PSC who were
given 2CDA subcutaneously in an open-label pilot trial of 6-month duration and 2-year
followup (Duchini et al., 2000). Patients had a quantifiable decrease in the hepatic
inflammatory infiltrate on liver biopsy. No significant changes were found in symptom
scores, liver panel tests, or cholangiograms. However, the drug was well-tolerated, and
two of four patients reported remission of their inflammatory bowel disease symptoms.

Psoriasis and psoriatic arthritis are conditions for which an immune pathogen-
ess is indicated by the association of the disease with particular HLA types, by the
dense lymphocyte infiltrates seen in both the skin and affected joints, and by
beneficial effects of treatment with cyclosporine A and anti-CD4 antibodies (Bos
et al., 1989). A study of 2CDA given orally to six patients weekly for 12 weeks
followed by monthly maintenance therapy showed that after 6 months, four out of six
patients had improved joint disease, and five out of six had improved psoriasis
(Eibschutz et al., 1995).

Finally, because the only toxicity of 2CDA noted in the early studies was bone-
marrow suppression, it has been used in preparing patients with lymphomas and
leukemias for bone-marrow transplantation because destruction of the marrow in
such patients is a part of the therapeutic aim (Beutler et al., 1997a,b).

16.5 CONCLUSIONS AND FUTURE DIRECTIONS

The development of the purine nucleoside analog 2CDA as a novel anticancer and
immunosuppressive agent was accomplished by Carson and colleagues based on a
greater understanding of the biochemical basis of a rare human immunological
disease caused by an inborn error of metabolism. The result was the introduction
into the clinic of an agent that could function by virtue of its activity in both actively
dividing and resting lymphoid and myeloid cells. Thus, 2CDA and similar analogs
have found great utility in the indolent hematological malignancies. Purine analogues
such as 2CDA as single agents in the indolent lymphoproliferative diseases are likely
to be used for many years. The extended role will be in combinations in both
hematological malignancies and solid tumors. In addition, the immunosuppressive
effect of the purine analogues will continue to lead to increased use of these drugs in
preparative regimens for bone-marrow transplantation and perhaps use in enhancing cytotoxicity in association with preparative regimens in autologous and allogeneic bone-marrow transplants (Keating, 1997).

Newer, next-generation purine nucleoside analogues, such as the arabinofluoro derivative of 2CDA (clofarabine), may have improved pharmacokinetic properties and exhibit broader single-agent antitumor activity than 2CDA (Carson et al., 1992; Faderl et al., 2005). Thus, seemingly minor changes in the structure of nucleoside analogues can result in major differences in pharmacokinetics, bioavailability, and biological activity. Future studies of 2CDA and the next-generation purine nucleoside analogues will include combination strategies of rational biochemical modulation with other agents such as cytarabine (ara-C).

REFERENCES


17.1 INTRODUCTION

Apoptosis is caused by the activation of intracellular proteases, known as caspases. These cysteine proteases cleave their cellular targets at aspartic acid residues residing in the context of tetrapeptide motifs within polypeptide substrates (Boatright and Salvesen, 2003). The human genome encodes 11 or 12 caspases, depending on certain hereditary polymorphisms (Reed et al., 2004; Saleh et al., 2004). Numerous cellular targets of caspases have been identified, which when cleaved produce in aggregate the characteristic morphology we call “apoptosis”. Several pathways for triggering caspase activation exist, although two have been elucidated in great detail and have been the center of much attention in recent years. These two pathways for apoptosis are commonly referred to as the intrinsic and the extrinsic pathways (Fig. 17.1) (Salvesen, 2002).

The intrinsic pathway centers on mitochondria as initiators of cell death. Multiple signals converge on mitochondria, including DNA damage, hypoxia, and oxidative stress, causing these organelles to release cytochrome c (cyt c) and other apoptogenic proteins into cytosol. In the cytosol, cyt c binds caspase-activating

*Corresponding author: jreed@burnham.org

Drug Discovery Research: New Frontiers in the Post-Genomic Era, Edited by Ziwei Huang
Copyright © 2007 John Wiley & Sons, Inc.
protein Apaf1, triggering its oligomerization into a hepatmeric complex that binds pro-caspase-9, forming a multiprotein structure known as the “apoptosome” (Salvesen and Renatus, 2002). Physical binding of Apaf1 to pro-caspase-9 is mediated by their caspase recruitment domains (CARDs) through homotypic CARD-CARD binding. Activation of apoptosome-associated cell death protease caspase-9 then initiates a proteolytic cascade, where activated caspase-9 cleaves and activates downstream effector proteases, such as pro-caspase-3.

In contrast, the extrinsic apoptotic pathway relies on TNF (tumor necrosis factor)-family death receptors for triggering apoptosis. A subgroup of the TNF-family receptors contain a cytosolic death domain that enables their intracellular interaction with downstream adapter proteins, which link these receptors to specific caspases. Upon ligand binding, TNF-family receptors containing cytosolic death domains (e.g., Fas; TNFR; TRAIL-R1; TRAIL-R2) cluster in membranes, recruiting caspase-binding adapter proteins, including the bipartite adapter FADD that contains both a death domain (DD) and a death effector domain (DED) (Wallach et al., 1999). The DED of FADD binds DED-containing pro-caspases (e.g. caspase-8 and -10), forming a “death inducing signaling complex” (DISC) and resulting in

Figure 17.1 Caspase-activation pathways. See text for details.
caspase activation by an “induced proximity” mechanism (Boatright and Salvesen, 2003).

Aside from the intrinsic and extrinsic pathways, multiple additional routes to caspase activation are possible, although for some the pathophysiological relevance is less well established (Fig. 17.1). For example, a pathway activated by cytolytic T-cells (CTLs) and natural killer (NK) cells involves injection of the contents of cytotoxic granules into target cells, at least in part through perforin-mediated mechanisms. Among the contents of the cytotoxic granules is granzyme B, a serine protease with specificity for aspartyl residues that cleaves and activates multiple members of the caspase family [reviewed in Lieberman (2003)]. A caspase-activation pathway linked to stress in the Golgi and endoplasmic reticulum (ER) has also been proposed, although details are unclear at present—particularly for human cells that seem to differ from their rodent counterparts with respect to production of ER-associated caspases (Saleh et al., 2004; Breckenridge et al., 2003). A p53-inducible protein, PIDD, recently was demonstrated to activate caspase-2 via bipartite adapter protein RAIDD (Tinel and Tschopp, 2004). The RAIDD protein has a DD that binds a corresponding DD in PIDD, and a CARD that binds the CARD found within the N-terminal prodomain of pro-caspase-2. In addition, a variety of proteins are capable of activating pro-caspase-1 and possibly related members of the caspase family (e.g., caspase-4 and -5 in humans), but their relevance to apoptosis is unclear (Reed et al., 2004; Tinel and Schopp, 2004). This branch of the caspase family is involved principally in cytokine maturation rather than apoptosis, cleaving and activating pro-interleukin-1β (pro-IL-1β) among other cytokines (Thornberry and Lazebnik, 1998). However, under certain pathological circumstances, particularly in the setting of host–pathogen interactions and ischemia–reperfusion injury, excessive caspase-1 activity may trigger apoptosis (Reed, 1999a).

It is important to note that while the pathways described above can all lead to caspase activation and apoptosis, some of them also trigger parallel caspase-independent cell-death mechanisms, which nevertheless kill cells. For example, mitochondria not only release caspase-activating proteins such as cyt c, but also release endonuclease G (Endo G) and a chromatin-modifying protein apoptosis inducing factor (AIF) that promote genome digestion and cell death independent of caspases (Penninger and Kroemer, 2003). Similarly, ER stress typically promotes a caspase-independent cell death mechanism, which may involve disturbances in intracellular Ca$^{2+}$ homeostasis (Demaurex and Distelhorst, 2003). Also, granzyme B as well as other proteins found within the cytotoxic granules of CTLs and NK cells can invoke caspase-independent cell-death mechanisms [reviewed in (Lieberman (2003)). Thus, some types of cell-death stimuli can induce parallel paths to apoptotic (caspase-dependent) and non-apoptotic (caspase-independent) cell demise.

17.2 APOPTOSIS BLOCKERS

Given the critical importance of making the correct choices about cell life–death decisions in complex multicellular organisms, it is not surprising that the pathways governing caspase activation are under exquisite control by networks of proteins that
directly or indirectly communicate with these proteases. A delicate balance between pro-apoptotic and anti-apoptotic regulators of apoptosis pathways is at play on a continual basis, ensuring the survival of long-lived cells and the proper turnover of short-lived cells in a variety of tissues, including the bone marrow, thymus, and peripheral lymphoid tissues. However, imbalances in this delicate dance of pro- and anti-apoptotic proteins occur in disease scenarios, including cancer where the scales tip in favor of anti-apoptotic proteins and endow cells with a selective survival advantage that promotes neoplasia and malignancy.

The anti-apoptotic proteins responsible for creating roadblocks to apoptosis have been mapped to specific pathways, providing insights into the defective cell-death mechanisms that contribute to malignancy. Among these apoptosis blockers are members of the Bcl-2 family, a large group of proteins ($n = 27$ in humans) that control mitochondria-dependent steps in cell-death pathways, including dictating whether cyt $c$ is or is not released from these organelles (Fig. 17.1). These proteins control the intrinsic pathway (Kroemer and Reed, 2000), although effects on the ER pathway for cell death have also been documented (Demaurex and Distelhorst, 2003). Another anti-apoptotic protein known to create apoptosis roadblocks in cancer cells is FLIP, a DED-containing protein that competes with pro-caspase-8 and -10 for interactions with TNF/Fas-family death receptor complexes (Tschopp et al., 1998). FLIP thus can create blockades in the extrinsic pathway. Downstream at the convergence of the apoptosis pathways, IAP-family proteins thwart cell death by directly binding to and suppressing the activity of certain caspases. The specific caspases and mechanisms of inhibition vary, depending on the IAP ($n = 8$ in humans), but both effector (e.g., caspase-3 and -7) and initiator (e.g., caspase-9) proteases are among the known targets of IAP-mediated suppression (Deveraux and Reed, 1999; Salvesen and Duckett, 2002).

The explanations for overexpression of anti-apoptotic proteins in cancer and leukemia cells vary, but documented mechanisms include chromosomal translocations, gene amplification, gene hypomethylation, transcriptional upregulation, and perhaps altered protein stability [reviewed in (Kitada et al., 2002)]. Conversely, pro-apoptotic genes that oppose these cytoprotective proteins can become inactivated in malignant cells through gene deletion, somatic mutation, gene hypermethylation, transcriptional downregulation, and probably other mechanisms.

The apoptosis-blocking proteins that thwart signaling through specific apoptosis pathways provide targets for possible drug discovery as do agents that stimulate pro-apoptotic proteins within these cell-death pathways. Theoretically, all drug targets can be addressed in multiple ways, focusing on agents that modulate the activity of the target at the levels of the DNA (gene expression), mRNA, or protein. A summary of some of the more promising strategies for modulating the activity of apoptosis genes and proteins for cancer therapy is provided below, according to the apoptosis pathways they affect.

## 17.3 EXTRINSIC PATHWAY ACTIVATORS

The extrinsic pathway is activated *in vivo* by TNF-family ligands that engage death receptors such as TNF, Fas, or TRAIL, resulting in the activation of caspases and
inducing apoptosis. Interest has emerged in developing therapeutics that kill cancer cells via the extrinsic pathway, particularly since chemorefractory cells tend to have defects in their intrinsic pathway, given the predominant reliance of cytotoxic drugs and X-irradiation on the mitochondrial pathway for cell death (Green and Reed, 1998). In this regard, the TNF cytokine family consists of 18 members in humans, with 29 counter receptors [reviewed in Reed et al. (2004) and Locksley et al. (2001)]. Some of these TNF-family receptors transduce signals predominantly for cell survival, particularly those that bind intracellular TRAF-family adapter proteins that link to downstream protein kinase pathways. Blocking these receptors represents a potentially attractive strategy for a variety of lymphoid malignancies but will not be discussed here. Other members of the TNF family directly trigger apoptosis, particularly those that contain the death domain (DD) structure within their cytosolic tails \((n = 8\) in humans). Of these receptors and their corresponding ligands, four receptors (three ligands) have been studied in detail with respect to their potential exploitation as cancer therapeutics—namely, TNF (TNFR1), Fas (FasL), and TRAIL (DR4; DR5). Strategies for exploiting apoptosis-inducing TNF-family ligands and receptors for cancer therapy mostly focus on (a) production of recombinant ligands (biologics), expressing only the extracellular domain of these type II integral membrane proteins, or (b) on use of agonistic monoclonal antibodies that bind the receptors and trigger apoptosis. In addition, however, emerging knowledge about intracellular modulators of the apoptosis pathways utilized by TNF-family death receptors has suggested opportunities for engaging the extrinsic pathway at other levels, improving sensitivity of malignant cells to biologics that activate the extrinsic pathway.

17.3.1 TNF-Family Ligands and Receptor-Targeting Antibodies

Over three decades ago, TNF was reported to selectively kill tumor but not normal cells, raising hopes that this or similar cytokines might be exploited for improved treatment of cancer [reviewed in Bazzoni and Beutler (1996)]. Unfortunately, the pro-inflammatory actions of this cytokine precluded its systemic administration and squelched efforts to apply it clinically. Then, we knew little about how TNF transduces signals into cells through its specific receptors, thus making it difficult to propose alternative strategies. Today, in contrast, we have an in-depth knowledge of the pathways that TNF and related cytokines activate, and from this knowledge have come several potential targets for drug discovery. We also understand far better how TNF-family cytokines regulate pathways that control apoptosis, and we can begin to envision ways of exploiting that knowledge for sensitizing tumors to these cytokines that our immune cells use to do battle with cancer.

Binding of TNF to one of its principal cellular receptors expressed widely on cells, TNFR1, triggers two parallel signaling pathways. These pathways bifurcate at the adapter protein TRADD. One of these pathways results in activation of caspase-family cell-death proteases, triggering apoptosis. The other parallel pathway triggers activation of NFkB-family transcription factors. NFkB influences the expression of many target genes involved in host defenses and immune regulation, among which are
several genes that suppress apoptosis. As a result, this NFκB pathway nullifies the caspase pathway, negating apoptosis [reviewed in Karin and Lin (2002)] in addition to accounting for the untoward inflammatory actions of this cytokine.

The TNF-family cytokines Fas-ligand (FasL) and TRAIL (Apo2 ligand) trigger activation of the caspase pathway without concomitant induction of NFκB. These death ligands are expressed on CTLs, NK cells, and other types of immune-relevant cells (activated monocytes/macrophages and dendritic cells) and are used as weapons for eradication of virus-infected and transformed cells [reviewed in Locksley et al. (2001), Ashkenazi and Dixit (1998), and Nagata (1996)]. Various studies using mice with genetic alterations in genes encoding these death ligands or their receptors, as well as use of neutralizing antibodies and Fc-fusion proteins, have provided evidence that FasL and TRAIL play important roles in tumor suppression by cellular immune mechanisms. For example, FasL is important for CTL-mediated killing of some tumor targets, and TRAIL is critical for NK-mediated tumor suppression [see, for example, Rosen et al. (2000) and Johnsen et al. (1999)].

The absence of pro-inflammatory effects of FasL and TRAIL has raised hopes that they might be successfully applied for cancer therapy, where TNF failed due to toxicity. While agonistic antibodies that trigger the receptor, Fas (CD95), are unfortunately highly toxic to liver (Ogasawara et al., 1993), TRAIL and agonistic antibodies that bind TRAIL receptors appear to be well-tolerated in vivo. Indeed, Phase I clinical trials have recently been completed using an agonistic monoclonal antibody ETR1 (Human Genome Science, Inc.) directed against TRAIL-receptor-1 (TRAIL-R1) [also known as death receptor-4 (DR4)], revealing little toxicity (Le, 2004). In mouse xenograft models using selected tumor human cell lines, TRAIL and agonistic antibodies directed against TRAIL receptors have been demonstrated to possess potent antitumor activity (Ashkenazi et al., 1999), raising hopes of using these biological agents as a novel approach to cancer treatment and thereby mimicking some of the effector mechanisms normally employed by the immune system in its defense against transformed cells.

Preclinical animal models also suggest that TRAIL-based therapies may be synergistic with conventional cytotoxic anticancer drugs. The molecular basis for this apparent synergy can differ among tumor cell lines tested, including: (a) chemotherapy-induced increases in TRAIL receptors, particularly DR5, which is a direct transcriptional target of p53 (Wu et al., 1997); (b) p53-mediated increases in Bid, a pro-apoptotic Bcl-2 family protein that is cleaved and activated by the apical protease in the TRAIL-induced pathway (e.g., caspase-8), thus representing a point of cross-talk between the extrinsic and intrinsic pathways (Sax et al., 2002); (c) p53-independent increases in Bak (LeBlanc et al., 2002); and probably others [reviewed in Roth and Reed (2003)]. Thus, knowledge of the ways that conventional cytotoxic drugs modify pathways relevant to TRAIL receptor signaling may provide a rational basis for selecting combination therapies and identifying subgroups of patients for inclusion in clinical trials.

17.3.2 NFκB Inhibitors

Abnormal elevations in NFκB activity are found in many types of tumors, including many lymphoid malignancies and several types of solid tumors [reviewed in
Karin et al. (2002), Rayet and Gelinas (1999)]. In fact, the first member of the NFκB family identified, \(\text{c-REL}\), is the cellular counterpart of a viral transforming gene, \(\text{v-rel}\), discovered in the avian leukosis virus. The importance of NFκB-family transcription factors for suppression of TNF-induced apoptosis is well established (Baud and Karin, 2001). Consequently, interfering with the sequence of signal-transduction events that generates active NFκB has emerged as an attractive approach to sensitizing cancer cells to TNF and related cytokines. Several anti-apoptotic genes are among the direct transcriptional targets of Rel-family proteins, including the genes encoding \(\text{c-FLIP}, \text{cIAP2}, \text{Bcl-X}\_\text{L},\) and \(\text{Bfl-1}.\) Thus, suppression of NFκB activity is an attractive strategy for restoring apoptosis sensitivity to several apoptosis pathways simultaneously.

Efforts to elucidate the signaling events responsible for activation of NFκB by TNF have revealed several candidate targets that are potentially amenable to small-molecule drug discovery approaches, particularly protein kinases responsible for eradicating endogenous inhibitors of NFκB. For example, although alternative pathways exist (Fan and Maniatis, 1991), an evolutionarily conserved pathway for NFκB activation has been revealed involving various upstream kinases funneling signals into a pivotal kinase complex that consists of two homologous serine-kinases, IKK\(\alpha\) and IKK\(\beta\), and a scaffold protein, IKK\(\gamma\) (also known as NEMO) (DiDonato et al., 1997; Zandi et al., 1997, 1998; Rothwarf et al., 1998). This IKK complex is responsible for phosphorylation of endogenous inhibitors of NFκB, the IκB proteins, with phosphorylation targeting these NFκB inhibitors for ubiquitination and subsequent proteasome-dependent degradation. Degradation of IκB releases NFκB, allowing its translocation into the nucleus [reviewed in Karin and Ben-Neriah (2000)]. Thus, the protein kinases involved in this pathway have emerged as promising drug targets, as well as possibly the F-box protein that recognizes phosphorylated IκB and recruits it to E3 ubiquitin ligase complexes, and the proteasome itself (Karin and Ben-Neriah, 2000; Maniatis, 1999).

In this regard, small-molecule inhibitors of IKKs have been described, which display pro-apoptotic activity against some types of cultured malignant cells. These IKK inhibitors include BMS-345541 (Burke et al., 2003); Bay 11-7082 (Dai et al., 2004); SC-514 (Kishore et al., 2003); Sulfasalazine (Weber et al., 2000); Silibinin, a dietary flavonoid that seems to directly inhibit IKKs (Dhanalakshmi et al., 2003); pyrolidine dithiocarbamate (Sumitono et al., 1999); natural product Wedelolactone and semi-synthetic derivatives of B-carboline natural products (Li et al., 2003; Castro et al., 2003); and others. The specificity of some of these inhibitors for IKK is questionable, but toxicology studies will determine whether they might be clinically useful.

Proteasome inhibitors include bortezomib, recently approved for second-line treatment of myeloma. This compound is in clinical testing for several other types of cancer and leukemia (Richardson et al., 2003; Wright et al., 2003) (Millennium Pharmaceuticals, Inc.). Bortezomib (PS-341) (Velcade\textsuperscript{TM}) inhibits the chymotrypsin-like protease activity of the proteasome (Wright et al., 2003; Paramore and Frantz, 2003). Pharmacological inhibition of proteasome activity not only suppresses IκB degradation, but also has effects on the stability of a large repertoire of cellular proteins, and thus the molecular basis for the antitumor activity of proteasome inhibitors remains unclear and may only partly be related to NFκB activity.
17.3.3 FLIP

Many tumors exhibit intrinsic resistance to TRAIL, anti-Fas antibodies, or both of these death stimuli, despite expressing the necessary cell-surface receptors (Snell et al., 1997). Multiple antagonists of the extrinsic pathway have been identified, including several DED-containing proteins that compete for binding to the adapter proteins or pro-caspases that participate in TNF-family death receptor signaling, including c-FLIP, BAR, and possibly Bap31 [reviewed in Krammer (2000) and Griffith et al. (1998)]. Among these, FLIP has received the most attention for its role in producing Fas- and TRAIL-resistant states in tumor cells (Tschopp et al., 1998; Irmter et al., 1997).

The c-FLIP protein is highly similar in overall sequence to pro-caspase-8 and -10, containing tandem copies of the DED, followed by a pseudo-caspase domain that lacks enzymatic activity. The c-FLIP gene resides in a tandem gene cluster on human chromosome 2, which contains the genes encoding pro-caspase-8 and -10, suggestive of gene-duplication events. Two isoforms of c-FLIP are produced from a single gene, including the long form described above and a shorter isoform that consists of only the DED domains. The shorter version of FLIP thus resembles analogous proteins encoded in the genomes of some mammalian viruses (Tschopp et al., 1998b). FLIP-S is exclusively anti-apoptotic, while FLIP-L can be either pro- or anti-apoptotic, depending on its levels of expression relative to pro-caspase-8 and -10 (Chang et al., 2002). FLIP proteins form complexes with pro-caspase-8 and -10, preventing their effective activation, as well as competing for binding to adapter proteins required for caspase recruitment to death receptor complexes [reviewed in Tschopp et al., (1998a, 1999)]. Overexpression of FLIP occurs commonly in cancers, implying in vivo selection, probably as a result of confrontations of neoplastic cells with the immune system.

Although no drugs have been described to date that directly bind FLIP proteins, experimental agents have been found that reduce FLIP protein expression. Among these are synthetic triterpenoids 2-cyano-3,12-dioxoolean-1,9-bien-28-oic acid (CDDO), CDDO-methyl ester (CDDO-me), and CDDO-imidazole (CDDO-Im), which induce ubiquitination and proteasome-dependent destruction of FLIP proteins in cultured cancer cells, sensitizing solid tumor lines to TRAIL apoptosis induction by TNF, Fas, and TRAIL (Kim et al., 2002). CDDO-Im was shown to synergize with recombinant TRAIL in a tumor xenograft model study using mice, revealing promising antitumor activity without toxicity to normal tissues (Hyer et al., 2005). Moreover, these and other CDDO analogs exhibit single agent activity in terms of inducing apoptosis in vitro of some types of cancer and leukemia cells (Pedersen et al., 2002b; Ikeda et al., 2004). The mechanism of CDDO and related compounds is still under investigation, but several studies have found activation of caspase-8 as a proximal event, indicative of stimulation of the extrinsic pathway (Pedersen et al., 2002b; Ito et al., 2000), although other mechanisms have also been proposed (Konopleva et al., 2002; Lapillone et al., 2003; Hail et al., 2004). Given its ability to induce activation of components of the extrinsic pathway, CDDO and its analogues may provide an option for bypassing the blocks to apoptosis that typically arise within the mitochondrial apoptosis pathway in chemorefractory tumors. CDDO and CDDO-me are undergoing
final preclinical development steps in preparation for human clinical trials, in collaboration with NCI DTP.

Additional possible routes to suppression of FLIP expression include NFκB pathway inhibitors (see above) and HDAC inhibitors (Aron et al., 2002). Also, two chemicals (4-(4-chloro-2-methylphenoxy)-N-hydroxybutanamide and 1-(3,5-di-tert-butyl-4-hydroxyphenyl)-2-(3-ethyl-2-imino-2,3-dihydro-benzoimidazol-1-yl)ethanone) were recently identified by cell-based high-throughput screening that reduce FLIP expression and sensitize tumor cells to TRAIL and FAS at concentrations of 5–50 μM (Schimmer et al., 2005). The cellular target(s) of these compounds are unknown, but siRNA experiments support a FLIP-dependent mechanism of action.

### 17.3.4 Additional Strategies for Overcoming Extrinsic Pathway Resistance

Other routes to restoring integrity of the extrinsic pathway have been revealed that may apply to leukemia and cancer. Protein kinase C (PKC) modulators such as byrostatin induce TNF production and apoptosis of myeloid leukemia cell lines via a caspase-8-dependent mechanism (Cartee et al., 2003). Also, all trans-retinoid acid (ATRA) induces TRAIL production and triggers autocrine induction of apoptosis in acute promyelomonocytic leukemia (APML) cells that harbor the RARα-PML fusion protein resulting from t(15;17) chromosomal translocations (Altucci et al., 2001).

### 17.4 INTRINSIC PATHWAY ACTIVATORS

Bcl-2-family proteins are the most prominent of the intrinsic pathway targets for cancer drug discovery. Though the human genome encodes 27 Bcl-2-family proteins, only six of these are anti-apoptotic and thus represent logical targets for therapy. The six anti-apoptotic members of the family are Bcl-2, Bcl-XL, Mcl-1, Bcl-W, Bfl-1, and Bcl-B. Overexpression of several of these anti-apoptotic Bcl-2-family proteins has been documented in various malignancies, although data are most extensive for the founding member of the family Bcl-2. Attempts to overcome the cytoprotective effects of Bcl-2 or related anti-apoptotic proteins in cancer and leukemia include three strategies: (1) shutting off gene transcription, (2) inducing mRNA degradation with antisense oligonucleotides, and (3) directly attacking the proteins with small-molecule drugs. A fourth strategy attempts to invoke endogenous antagonists of anti-apoptotic Bcl-2-family proteins to nullify their cytoprotective functions.

### 17.4.1 Drugs Regulating BCL-2 Gene Expression

Some synthetic retinoids are known to reduce levels of BCL-2 or BCL-XL mRNA in leukemia cells, suggesting a potential explanation for the pro-apoptotic effect of these agents that are already approved and in clinical use (Reed, 1999b). Estrogen receptor antagonists also reduce BCL-2 expression in breast cancer cell lines (Pratt, 1998), providing at least a partial explanation for the pro-apoptotic activity of these compounds in ER-positive tumors of mammary origin. Steroid hormone, 1α,25(OH)₂
vitamin D also suppresses expression of anti-apoptotic genes $BCL-2$, $BCL-X$, and $MCL-1$ in prostate cancer cell lines that contain the vitamin D receptor (VDR) (Guzey et al., 2002) and may underlie in part the antitumor activity observed recently in clinical trials of high dose $1\alpha,25(OH)_2$ vitamin D (calcitriol) for treatment of hormone refractory prostate cancer (Johnson et al., 2002). Compounds that inhibit histone deacetylases (HDACs), chromatin-modifying enzymes that interact with many transcription factors, also reduce the expression of $BCL-2$, $BCL-X$, or $MCL-1$ at a transcriptional level (Khan et al., 2004; Mori et al., 2004; Rosato et al., 2003). Clinical trials of HDAC inhibitors are progressing, with hints of activity documented for lymphoma and some solid tumors (Kelly et al., 2003). PPAR$\gamma$-modulating drugs also can reduce expression of $BCL-2$ or other anti-apoptotic $BCL-2$ family genes (83).

### 17.4.2 Drugs Attacking Bcl-2 mRNA

Antisense oligodeoxynucleotides (ODNs) targeting the $BCL-2$ mRNA are undergoing clinical evaluation, with Phase III trials underway or recently completed for relapsed or refractory CLL, AML, and myeloma and also for hormone refractory prostate cancer (Buchele, 2003). The molecules in clinical trials are composed of nuclease-resistant phosphorothioates, which hybridize via Watson–Crick base-pairing to the first 18 nucleotides within the open reading frame of Bcl-2 mRNAs and induce RNaseH-mediated degradation (Reed et al., 1990). This DNA-based drug (G3139) (oblimersen sodium) (Genasense$^\text{TM}$) is being developed in combination with cytotoxic chemotherapy, in an attempt to exploit the chemosensitizing properties of Bcl-2-based therapies. The first reported Phase I trial of $BCL-2$ antisense therapy showed promising activity against chemorefractory non-Hodgkin lymphoma (NHL) (Webb et al., 1997). Uncontrolled Phase II data also suggest that this $BCL-2$ antisense molecule (Oblimersen sodium) (Genasense$^\text{TM}$) has promising bioactivity against B-cell malignancies and adult AML, while producing little toxicity at doses $<4$ mg/kg. A recently completed trial of oblimersen sodium at 7 mg/kg for advanced melanoma (in combination with alkylating agent dacarbazine) demonstrated improved response rates and time to progression (TTP) but failed to meet its primary endpoint of prolonging patient survival. Further follow-up of the melanoma patient cohort, however, suggests a survival benefit for patients with lower pretreatment tumor burden, as defined by serum levels of lactate dehydrogenase (LDH). A recently completed Phase III trial of oblimersen sodium at 3 mg/kg in patients with refractory CLL met its primary endpoint objective of demonstrating improved responses to cytotoxic drug therapy (cyclophosphamide plus fludarabine). These clinical studies further validate $BCL-2$ as a promising target for cancer therapy.

Demonstration of a correlation between clinical responses and antisense-mediated knockdown of Bcl-2 mRNA or protein levels in patients undergoing treatment with oblimersen sodium has been documented in some but not all studies, raising questions about the mechanism of this novel targeted therapy (Rudin et al., 2004; Marcucci et al., 2003). However, it has been argued that cells with successful antisense-mediated reductions in Bcl-2 mRNA are difficult to recover from patients because of their rapid in vivo clearance induced by apoptosis. The Bcl-2 protein also exhibits a very slow spontaneous degradation rate ($\sim$12–36 hr), suggesting that
prolonged suppression of mRNA accumulation is necessary to effect a decline in Bcl-2 protein [reviewed in Reed (1996a)]. A pro-inflammatory effect of oblimersen sodium, seen particularly at higher doses, suggests an additional non-antisense bioactivity, which may or may not be relevant to tumor responses. In this regard, CpG motifs in synthetic DNA molecules can trigger inflammatory responses by engaging toll-like receptor-9 (TLR9), tapping into a mechanism of innate immunity intended to sense the presence of unmethylated bacterial DNA (Kimbrell and Beutler, 2001). Interestingly, mismatched ODNs containing CpG motifs were shown to reduce BCL-2 expression and promote apoptosis of prostate cancer cell lines through an antisense-independent mechanism (Lai et al., 2003). Nevertheless, tumor xenograft studies of oblimersen sodium in immunocompromised mice provide strong support for an antisense-based mechanism, whereby sequence-specific reductions in BCL-2 expression either directly trigger tumor apoptosis or sensitize malignant cells to traditional cytotoxic anticancer drugs (Cotter et al., 1994; Klasa et al., 2000; Guinness et al., 2000; Jansen et al., 1998; Miyake et al., 1999).

The genetic lesions complementing BCL-2 expression in tumors probably dictate whether inhibiting BCL-2 is sufficient to induce apoptosis (single agent activity) versus merely sensitizing cells to chemotherapy. For example, B-cell malignancies with simultaneous activation of C-MYC and BCL-2 have been described, arising from transformed follicular lymphomas where a t(14;18) activating BCL-2 is combined with a t(8;14) activating C-MYC (Gauwerky et al., 1988). Activated C-MYC drives both cell proliferation and cell death, the latter therefore requiring complementation by BCL-2 (Evan and Vousden, 2001). Thus, depriving cells with activated C-MYC of their Bcl-2 protein is tantamount to a death sentence, suggesting a possible strategy for patient selection based on simple biomarker profiling that demonstrates overexpression of both BCL-2 and C-MYC.

Aside from oblimersen sodium, no nucleic acid-based inhibitors of Bcl-2 or its anti-apoptotic relatives are presently in clinical testing, but multiple preclinical programs have been reported. For example, antisense DNA molecules that target Bcl-XL mRNA and induce its RNaseH-mediated destruction (ISIS16009) have been generated and shown to reduce resistance of tumor and leukemia cell lines to cytotoxic anticancer drugs in culture and in xenograft models in mice (Guensberg et al., 2002; Fennell et al., 2001). Mcl-1-targeting antisense ODNs (Derenne et al., 2002) also reportedly overcome the survival-promoting activity of follicular dendritic cells on cultured CLL B-cells, which induce Mcl-1 expression when co-cultured with these leukemic B-cells (Pedersen et al., 2002a). Mcl-1 antisense also induces apoptosis of multiple myeloma cells more efficiently than antisense targeting either Bcl-2 or Bcl-XL (Derenne, 2002), and displays activity in preclinical animal models of myeloma (Thallinger et al., 2003). A clever strategy using antisense chemistry that fails to support RNaseH-based mRNA destruction has also been described, wherein a 2’-methoxyethyl-based antisense ODN hybridized to a splice acceptor site in the Bcl-XL pre-mRNA forced splicing to a downstream exon to produce a shorter Bcl-X protein with pro-apoptotic rather than anti-apoptotic activity (Taylor et al., 1999). A dual inhibitor of both Bcl-2 and Bcl-XL has also been reported that takes advance of nucleotide sequence conservation, targeting both of these genes simultaneously (Zangemeister-Wittke et al., 2000).
17.4.3 Drugs Attacking the Bcl-2 Protein

Small-molecule inhibitors that directly bind Bcl-2 or related anti-apoptotic proteins have also entered clinical trials for cancer. The most advanced of these is a natural product, gossypol, whose ability to bind and inhibit Bcl-2 was unknown at the time initial clinical testing began (Stein et al., 1992; Kitada et al., 2003) (Table 17.1). Gossypol is a compound found in cottonseeds originally used as an herbal medicine in China. Gossypol promotes apoptosis of tumor cells in vitro and shows antitumor activity in tumor xenograft studies in mice [reviewed in Pellecchia (2004)]. Gossypol binds a hydrophobic pocket found on the surface of anti-apoptotic Bcl-2-family proteins. This binding pocket represents a regulatory site where endogenous antagonists dock onto Bcl-2 and related anti-apoptotic proteins, negating their cytoprotective activity (Fesik, 2000). The endogenous antagonists bind via a conserved 16 amino-acid motif called a Bcl-2 homology-3 (BH3) domain (Fig. 17.2). Proof-of-concept experiments using BH3 peptides have suggested that compounds docking at this regulatory site on Bcl-2 and its related anti-apoptotic proteins could provide a route to effective suppression of these proteins and thereby promote apoptosis of malignant cells (Holinger et al., 1999).

Gossypol interacts with the BH3-binding pockets of five of the six anti-apoptotic Bcl-2-family proteins, displacing BH3 peptides with an IC$_{50}$ of $0.5–2\mu M$ from Bcl-2, Bcl-X$_L$, Bcl-B, Bcl-W, and Mcl-1 in vitro (Zhai et al., 2005), thus making it a broad-spectrum inhibitor. An enantiomer of gossypol is touted to have superior activity compared to the naturally occurring racemer and is in clinical development by ASCENTA, Inc., in collaboration with the Developmental Therapeutic Branch (DTB) of the National Cancer Institute (NCI-USA) (Qiu et al., 2002). However, gossypol is a highly reactive compound, containing two aldehydes that may explain some of the toxicities originally seen in Phase I trials of this natural product, in addition to producing unfavorable pharmacological properties. For this reason, attempts to produce semi-synthetic analogues that retain activity against Bcl-2 have begun. The most advanced of these is the compound called apogossypol, in which the two aldehydes were eliminated (Becattini et al., 2004). This compound is also under evaluation by the DTB of NCI in the hope of achieving stable formulations that can be advanced into human clinical trials.

GX-15-070 is a synthetic broad-spectrum antagonist of anti-apoptotic Bcl-2-family proteins currently being evaluated in Phase I trials by GEMINEX, Inc. This inhibitor is a cycloprodigiosin derivative touted as a broad-spectrum inhibitor of anti-apoptotic Bcl-2-family proteins (unpublished information presented by sponsoring company, GeminX, Inc.), which is said to have roughly 1$\mu M$ affinity for Bcl-2 and several anti-apoptotic Bcl-2 family proteins. Little information about GX-15-070, however, is available in peer-reviewed literature.

Several other small-molecule antagonists of Bcl-2, Bcl-X$_L$, and Mcl-1 have been reported, all of which are still in preclinical evaluation at this time. Among these are: (a) HA14-1, a chromene derivative identified by a computational modeling approach (Wang et al., 2000); (b) BH3I-1 and -2, a thiazolidin derivative and benzene sulfonyl derivative, respectfully, identified by screening using a BH3 peptide displacement assay (Degterev et al., 2001); (c) antimycin analogues, identified by
TABLE 17.1 Structures of Bcl-2 Antagonists

- Chelerythrine
- ABT-737
- EGCG
- GX-15
- Apogossypol
- YC137
- Gossypol
- BH3-1
- Antimycin A3
- HA14-1
computational modeling (Tzung et al., 2001); (d) certain theaflavins, a class of natural products abundant in black tea, some epigallocatechins, natural products abundant in green tea (Leone et al., 2003), and the natural product chelerythrine (Chan et al., 2003); (e) ABT737, a synthetic small-molecule inhibitor produced by NMR-guided, structure-based drug design (Abbott Laboratories, Inc.) (Olfersdorf et al., 2005) (Table 17.1). Side-by-side comparisons of these small-molecule inhibitors of anti-apoptotic Bcl-2-family proteins have recently been reported, suggesting a rank-order potency with respect to affinity for the BH3 pocket of Bcl-2 or Bcl-XL of ABT-737 > EGCG > theaflavins > gossypol/apogossypol > chelerythrine > HA14-1 and antimycin (Zhai 2006).

Several issues will be encountered as these small-molecule antagonists of Bcl-2 move forward with their preclinical and clinical development, including: (a) compound stability and formulation; (b) pharmacokinetics (PK) and metabolism; (c) toxicity; and (d) off-target mechanisms. For example, some of the natural products that target Bcl-2 are inherently unstable molecules prone to rapid inactivation by oxidation in vitro, in addition to rapid metabolism to inactive products in vivo. With respect to toxicity, since some of these small molecules attack multiple anti-apoptotic members of the Bcl-2 family, toxicity to normal tissues is more likely than with antisense-based drugs that selectively target only one family member. The potential advantage of broad-spectrum inhibitors of anti-apoptotic Bcl-2-family proteins, however, is that they may be more effective due to redundancy within the Bcl-2 family and the tendency of tumors to express more than one anti-apoptotic member of the family. Finally, most of the natural products that bind the BH3 pocket of Bcl-2 have additional cellular targets, and these off-target activities may or may not be important to the antitumor activity of the compounds. Provided these off-target activities do not add untoward toxicity, this may be an advantage rather than a disadvantage for anticancer drug development. Nevertheless, multitarget activity potentially complicates structure-activity relation (SAR) studies in which one might envision producing semi-synthetic derivatives that retain activity against Bcl-2 since chemical modifications that enhance Bcl-2 binding might simultaneously reduce interaction with other cancer-relevant targets.

Figure 17.2  Structure of Bcl-2 and associated BH3 peptide. The 3D structure of Bcl-XL is depicted (blue) with Bak BH3 peptide bound (red). See color plates.
17.4.4 Activating Endogenous Antagonists of Bcl-2

In addition to Bcl-2-targeting strategies that attack at the level of gene expression, mRNA, or proteins, a fourth possible route to dealing with the problem of pathological overexpression of Bcl-2 in malignant cells has recently been revealed. This strategy seeks to induce expression or activation of endogenous proteins that bind and negate the cytoprotective action of Bcl-2. For example, an orphan member of the retinoid/steroid family of nuclear receptors, TR3 (Nur77) has been shown to translocate from the nucleus to the cytosol in response to certain cell-death stimuli in some types of cells (Li et al., 2000). In the cytosol, TR3 binds a regulatory domain within the Bcl-2 protein, thus accounting for accumulation of TR3 on the surface of mitochondria. TR3 induces a profound conformational change of Bcl-2, causing exposure of this BH3 domain and converting Bcl-2 from a protector to a killer (Lin et al., 2004). The BH3 domain of Bcl-2 binds both pro-apoptotic Bcl-2-family members Bax and Bak as well as anti-apoptotic Bcl-2-family proteins, probably activating the pro-apoptotic while inactivating the anti-apoptotic.

Recently, a class of compounds representing analogs of the retinoid APHN has been identified that induces TR3 expression and translocation into the cytosol. The prototype compound, 3Cl-AHPC, lacks activity against retinoid receptors although having evolved from attempts to synthesize highly selective retinoid ligands (Dawson et al., 2004). 3Cl-AHPC has demonstrated activity against cultured leukemia cells and in preclinical animal models of AML (Zhang et al., 2002; Dawson et al., 2001) and is active even against retinoid refractory leukemia cell lines. 3Cl-AHPC appears to induce phosphorylation of TR3, possibly via a Jun N-terminal kinase (JNK) pathway, correlating with its exodus from the nucleus (Kolluri et al., 2003). This compound may therefore provide a mechanism to invoke the TR3 pathway, converting Bcl-2 from a protector to a killer and thus making Bcl-2 a liability rather than an advantage for malignant cells.

Opposing this TR3-mediated pathway for apoptosis is Akt/PKB, which appears to nullify pro-apoptotic actions of TR3 by phosphorylation at serine 530 (Pekarsky et al., 2001; Masuyama et al., 2001). Thus, TR3 joins a long list of apoptosis-relevant substrates of Akt, which includes Bcl-2 antagonist BID, human caspase-9, pro-apoptotic kinase Ask1, p53 antagonist Mdm2, and others [reviewed in Testa and Bellacosa (2001)]. Small wonder then that Akt has emerged as a prominent target for cancer drug discovery [reviewed in Mitsiades et al., (2004)]. In this regard, relatively selective but fairly weak inhibitors of Akt such as 1-L-hydroxymethyl-chiro-inositol-2-((R)-2-methyl-3-O-octadecyl/carbonate) have been described, which has preclinical activity against multiple myeloma cells (Mitsiades et al., 2002). Interestingly, 3Cl-AHPC may also reduce Akt activity in tumor cells (Farhana et al., 2004), further promoting the TR3 pathway for cell death. It remains to be determined what effect the Tcl-1 oncoprotein has on nuclear export of TR3, but it may be of interest that this protein overexpressed in CLLs promotes nuclear targeting of Akt and regulates the transcriptional activity of TR3 (Pekarsky et al., 2001).

The revelation that the Bcl-2 protein can have two opposing phenotypes, depending upon its interactions with other proteins such as TR3, suggests a possible
explanation for the paradoxical association of higher Bcl-2 with favorable clinical outcome in some types of cancer (Reed, 1996b). Interestingly, TR3 is among a small set of genes identified by DNA microarray analysis demonstrating a strong association with longer survival among patients with diffuse large B-cell lymphoma (Shipp et al., 2002). Examination of Bcl-2 and TR3 expression in clinical specimens may help select patients most likely to benefit from therapeutic strategies designed to stimulate the TR3 pathway for Bcl-2 conversion. Clearly, one would not want to combine TR3-based therapies with drugs affecting BCL-2 expression, given that antisense-mediated suppression of BCL-2 expression protects tumor cell lines in culture from TR3-induced apoptosis (Lin et al., 2004).

Binding of TR3 to Bcl-2 requires the presence of an ~50 amino-acid flexible loop region located between the first two α-helices in the Bcl-2 protein (Lin et al., 2004). Prior studies have shown that this domain in Bcl-2 and Bcl-XL negatively regulates the anti-apoptotic functions of these proteins (Chang et al., 1997). It is currently unknown which of the six anti-apoptotic Bcl-2 family proteins can bind TR3, but all of them have predicted "loop" domains. Interestingly, at least one anti-apoptotic Bcl-2-family member, mouse A1 (whose human counterpart is BH1) reportedly resides in the nucleus (Somogyi et al., 2001), raising suspicions that it may target nuclear proteins such as TR3. The loop of Bcl-2 is a documented site of phosphorylation, correlating with inactivation of the anti-apoptotic function of Bcl-2 in some scenarios (Haldar et al., 1998). It remains to be explored whether phosphorylation of Bcl-2 affects TR3 binding, but microtubule-modulating drugs such as the taxanes are the most potent inducers of Bcl-2 loop phosphorylation identified thus far. Of note, somatic mutations have been documented in rearranged BCL-2 genes in lymphomas, corresponding to amino-acid substitutions in the loop of the Bcl-2 protein (Tanaka et al., 1992). This observation implies that Bcl-2 may acquire mutations that abolish TR3 sensitivity, a hypothesis that awaits experimental confirmation.

To some extent, the concept of invoking endogenous antagonists of Bcl-2 is not novel. Anti-apoptotic Bcl-2-family proteins are normally held in check by endogenous proteins that contain a BH3 domain, especially the so-called BH3-only proteins (BOPs). Several of these BOPs are linked to pathways that help explain how the currently available cytotoxic anticancer drugs trigger cell death, including: (a) the genes encoding Puma, Noxa, and Bid, whose expression is directly induced by p53; (b) isoforms of the Bim proteins, which are normally sequestered on microtubules; and (c) Bax, which is held in an inactive conformation in the cytosol by the regulatory subunit (Ku70) of the DNA-dependent kinase, becoming released during responses to DNA damage (Huang and Strasser, 2000; Sawada et al., 2003). Unfortunately, defects in these specific pathways involving BOPs or pathological overexpression of Bcl-2 that overwhelms the activity of these endogenous antagonists all too frequently create roadblocks to apoptosis that must be overcome by alternative means. The observation that commonly used cytotoxic anticancer drugs are capable of modulating endogenous pathways impinging on Bcl-2 and related anti-apoptotic proteins underscores why Bcl-2-targeted therapies are often synergistic with conventional cytotoxic agents in terms of inducing apoptosis of malignant cells. Knowledge of the molecular details of the
mechanisms by which traditional cytotoxic anticancer drugs affect Bcl-2 family proteins thus may help in selecting optimal combination therapies for clinical trials of Bcl-2 antagonists, where the Bcl-2 antagonists are used to sensitize cancer and leukemia cells to apoptosis induction by conventional cytotoxic drugs or radiation therapy. Underscoring the importance of this knowledge base are recent studies suggesting that alkylating agents are capable of inducing cell death via a necrotic rather than apoptotic mechanism, which is independent of Bcl-2/Bax-family proteins (Zong et al., 2004). It is therefore unfortunate that many of the early clinical trials of Bcl-2 antagonists (e.g., oblimersen sodium [Genasense™]) have involved combination with alkylating agents since this newfound knowledge suggests that this would not be a synergistic combination.

As we continue to learn more about apoptosis mechanisms involving Bcl-2-family proteins, novel strategies for therapeutic intervention emerge. Another recent example is the discovery that a cytoplasmic pool of p53 associates with mitochondria, directly inducing activation of the Bax protein through a transcription-independent mechanism (Chipuk et al., 2004). Importantly, even mutant p53 is capable of activating this cell-death pathway, raising hopes of finding pharmacological interventions that would entice mutant p53 to attack mitochondria and trigger apoptosis of cancer cells in which this important tumor suppressor gene product has suffered somatic mutations that inactivate its nuclear functions.

### 17.5 CONVERGENCE PATHWAY

The intrinsic and extrinsic pathways for apoptosis converge on downstream effector caspases. Certain effector caspases are targets of suppression by an endogenous family of anti-apoptotic proteins called IAPs (inhibitor of apoptosis proteins). The IAPs represent an evolutionarily conserved family of proteins, originally identified in insect viruses (Miller, 1999). IAPs contain one or more copies of a domain called the baculoviral IAP repeat (BIR). BIR domains are sometimes accompanied by other domains, including RING domains and ubiquitin-conjugating enzyme domains (E2s), which endow these proteins with additional properties such as an ability to target themselves and associated proteins for ubiquitination and proteasome-dependent degradation. The human genome encodes eight IAP-family members: XIAP, cIAP1, cIAP2, Naip, Survivin, ML-IAP (Livin; K-IAP), ILP2 (Ts-IAP), and Apollon (Bruce) (Reed et al., 2004).

Pathological overexpression of IAPs has been documented in cancer and leukemia [reviewed in Deveraux and Reed (1999) and Prefontaine et al., (1998)]. For example, the IAP-family member Survivin is overexpressed in most cancers (Ambrosini et al., 1997) and has become a topic of considerable attention for its dual role as a regulator of cell division (chromosome segregation and cytokinesis) and apoptosis [reviewed in Altieri et al., (1999), Reed and Reed (1999), and Reed and Bischoff (2000)]. Similarly, the IAP-family member ML-IAP is rarely expressed in normal tissues but is found at elevated levels in melanomas and some renal cancers (Vucic et al., 2000; Lin et al., 2000; Kasof and Gomes, 2001). Chromosomal translocations activating the cIAP2 gene are found in some mucosal-associated
lymphomas (Motegi et al., 2000). The XIAP gene is overexpressed in a substantial proportion of AMLs (Tamm et al., 2000), CLLs (Schimmer et al., 2003), hepatocellular carcinomas (Shiraki et al., 2003), ovarian cancers (Li et al., 2005), and prostate cancers (Krajewska et al., 2003).

The functional importance of overexpressed IAPs for apoptosis suppression in cancers has been supported by antisense experiments in which knocking down expression of SURVIVIN, XIAP, or other IAPs family genes has been shown to induce apoptosis of tumor cell lines in culture or to sensitize tumor cell lines to apoptosis induced by anticancer drugs [Chen, 2000; Gordon, 2002; Li, 2001; Sasaki, 2000; Holcik, 2000]. In contrast, gene knockout studies in mice imply that most normal cells are potentially less dependent on IAPs compared to tumor cells, inasmuch as targeted disruption of the genes for xiap, ciap1, and ciap2 produces little phenotype (Harlin et al., 2001).

Routes to small-molecule drug discovery have been revealed through knowledge of the structural and molecular details of how IAPs inhibit caspases, enlightened by insights gained from studies of endogenous inhibitors of the IAPs. For example, specific segments within IAP-family proteins have been identified that bind specific caspases, providing structural information useful for envisioning generations of compounds that disrupt these protein interactions, freeing caspases to induce apoptosis (Deveraux and Reed, 1999; Salvesen and Duckett, 2002; Fesik and Shi, 2001).

XIAP, for example, which contains three tandem BIRs, utilizes different domains to suppress both downstream effector caspases that operate at points of convergence of apoptosis pathways (e.g., caspase-3 and -7), and caspase-9, the apical protease in the mitochondrial pathway for apoptosis. The BIR2 domain and a flanking upstream linking segment of XIAP bind active caspase-3 and -7, while the BIR3 domain is responsible for binding active caspase-9. In contrast, ILP2, which contains a single BIR domain, is a potent suppressor only of caspase-9 but not caspase-3 and -7.

### 17.5.1 Drugs Targeting IAP-Family Proteins

Using an enzyme derepression assay in which XIAP-mediated suppression of caspase-3 is overcome by chemical compounds, at least two classes of XIAP antagonists have been described that target the vicinity of the BIR2 domain and that free caspase-3, including diphenyl and triphenylureas and benzenesulfonamide derivatives (Schimmer et al., 2004; Wu et al., 2003) (Table 17.2). The phenylureas exhibit broad-spectrum apoptosis-inducing activity against cultured human tumor cell lines and primary leukemias and suppress growth of tumor xenografts in mice without apparent toxicity to normal tissues (Schimmer et al., 2004). Importantly, these compounds induce apoptosis through a Bcl-2/Bcl-XL-independent mechanism, which provides encouragement that they may prove useful for chemorefractory cancers and leukemias.

Attempts to mimic the actions of endogenous antagonists of IAPs define another strategy for achieving specific inhibitors. Endogenous proteins such as SMAC and OMI (HtrA2) bind and suppress IAPs, releasing caspases to kill cells (Du et al., 2000; Verhagen et al., 2000; Suzuki et al., 2000). A 4’mer peptide corresponding to the N-terminus of the mature SMAC protein is reported to be sufficient to bind IAPs and block their association with caspases (Fesik and Shi, 2001;
Chai et al., 2000), which is analogous to similar mechanisms previously defined in *Drosophila* where IAPs are kept in check by endogenous antagonists Reaper, Hid, Grim, and Sickle [reviewed in Goyal (2001)]. By fusing membrane-penetrating peptides onto SMAC, OMI, Hid, or Reaper peptides, it is possible to induce apoptosis of human cancer cell lines in culture as well as to suppress tumor growth in xenograft models in mice (Deng et al., 2002; Guo et al., 2002; Carson et al., 2002; Chauhan et al., 2001). These data thus provide proof-of-concept evidence that small molecules that mimic the effects of these IAP-binding peptides could potentially be exploited as drugs for cancer treatment.

The three-dimensional structure of the BIR3 domain of XIAP complexed with SMAC reveals that N-terminal 4-amino acids of the mature SMAC protein bind in the same crevice normally occupied by the N-terminus of the small subunit of processed caspase-9, competing for binding (Fesik and Shi, 2001) (Fig. 17.3). Consequently, small-molecule compounds that mimic the SMAC 4’mer peptide dislodge active caspase-9 from BIR3, promoting apoptosis. Analogous SMAC-peptide binding crevices are anticipated on the BIRs of many other IAP-family proteins, based on molecular modeling, suggesting SMAC peptide-mimicking compounds would be rather broad-spectrum in their activity. Several groups have initiated preclinical studies of SMAC-mimics. Among the reported antagonists that target the SMAC binding site are embelin, a natural product (Nikolovska-Coleska et al., 2004), as well as several synthetic peptidyl and nonpeptidyl SMAC mimics (Li et al., 2004; Sun et al., 2005), with the most potent binding with low nanomolar affinity (Oost et al., 2004).

In addition to SMAC and OMI, additional endogenous antagonists of IAPs have been described, including ARTS, XAF1, and NRAGE (Gottfried et al., 2004; Liston...
**TABLE 17.2 Structures of IAP Antagonists**

<table>
<thead>
<tr>
<th>SMAC</th>
<th>Non-SMAC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ABT-Thiazol derivative 20 g</strong></td>
<td><strong>1396-11</strong>&lt;br&gt;N=-(N-(2S)-1-[2-(1-adamantyl)ethyl]pyrrolidine-2-ylmethyl)-N'-[(1R)-1-[[anilinocarbonyl]amino]methyl]-5-[<a href="methyl">anilinocarbonyl</a>amino]penty]-N'-phenylurea;</td>
</tr>
<tr>
<td>2-Amino-N-[4-(4-bromo-phenyl)-5-phenyl-thiazol-2-yl]-propionamide</td>
<td><strong>1396-12</strong>&lt;br&gt;R = 2-Cyclohexyl-ethyl (N-{[(2S)-1-(4-cyclohexylbutyl)pyrrolidine-2-yl][methyl]amino}hexyl]-N-methyl-N'-phenylurea)</td>
</tr>
<tr>
<td><strong>Capped Tripeptides</strong></td>
<td></td>
</tr>
<tr>
<td>R&lt;sub&gt;1&lt;/sub&gt; = H or phenyl; R&lt;sub&gt;2&lt;/sub&gt; = t-butyl or Cyclohexyl; R&lt;sub&gt;3&lt;/sub&gt; = H or methyl</td>
<td></td>
</tr>
<tr>
<td><strong>Embelin</strong></td>
<td><strong>TWX006</strong>&lt;br&gt;X=−CO−&lt;br&gt;(4-[4-[3-{1-(Biphenyl-4-carbonyl)-piperidin-4-yl]propyl}-piperidin-1-yl]-N-ethyl-1-N-(3-ethylamino-propyl)-2-fluorobenzene sulfonamide</td>
</tr>
<tr>
<td>2,5-Dihydroxy-3-undecyl-[1,4]benzoquinone</td>
<td><strong>TWX024</strong>&lt;br&gt;X=−CONH−&lt;br&gt;(4-[3-{1-[4-{Ethyl-(3-ethylamino-propyl)-sulfamoyl]-3-fluorophenyl}-piperidin-4-yl]propyl]-piperidine-1-carboxylic acid biphenyl-4-ylamide)</td>
</tr>
<tr>
<td><strong>Dimer of N-{2-But-2-ynyloxy-1-[2-(5-phenylsulfanyl-tetrazol-1-ylmethyl)]pyrroldine-1-carbonyl}-2-methylamino-propionamide.</strong></td>
<td></td>
</tr>
</tbody>
</table>
et al., 2001; Jordan et al., 2001). The structural details of how these proteins interact with their IAP targets is thus far undefined but, when known, may provide additional strategies for generation of small-molecule antagonists that work through non-SMAC mechanisms. Given that the aforementioned phenylurea-base antagonists of XIAP have been reported to target a non-SMAC site on XIAP (Schimmer et al., 2004), it will be interesting to determine whether those compounds in fact mimic ARTS, NRAGE, or XAF1 in terms of their docking site on XIAP.

Finally, a variety of indirect strategies for negating the activity of IAPs are beginning to emerge. For example, studies have shown that phosphorylation of Survivin on threonine 34 is required for execution of cell division and suppression of apoptosis (O’Connor et al., 2000). The predominant kinases responsible for phosphorylation of Survivin appear to be Cdc2 or related cyclin-dependent kinases. Compounds inhibiting Cdc2-family kinases include flavopiridol, currently in clinical trials.

### 17.5.2 Drugs Inhibiting IAP-Family Gene Expression

In addition to targeting IAP-family proteins directly with small-molecule drugs, efforts to reduce expression using antisense-based drugs are also well underway. Antisense ODNs targeting Survivin mRNA (Chen et al., 2000) have begun clinical development using second-generation antisense chemistry (ISIS Pharmaceuticals/Lilly Inc). An antisense-based approach to Survivin is particularly attractive because the mechanism by which this protein suppresses caspases differs from most other IAPs (Marusawa et al., 2003), and a clear path forward for small-molecule drug discovery remains uncertain based on available protein structure data (Fesik, 2000). In addition, a phosphorothioate-based antisense ODN targeting XIAP (AEG35156/GEM640) has been designated for human clinical trials (Agera Pharmaceuticals, Inc.). The classical dilemma of specificity versus redundancy is highly germane to the question of whether small-molecule compounds versus nucleic-acid-based drugs are preferable. Most strategies for small-molecule drugs are likely to target multiple members of the IAP family and thus could afford advantages in terms of addressing more drug targets but also might suffer disadvantages with respect to possible toxicities. In this regard, the observation that some cancers overexpress four or more IAP members simultaneously argues in favor of a small-molecule strategy involving broad-spectrum inhibitors (Krajewska et al., 2003). Yet, experiments using antisense or siRNA to knock-down expression of individual IAPs have suggested that highly selective inhibitors could also be successful. Future work will empirically define the best paths forward for cancer drug discovery efforts that target single or multiple members of the IAP family.

### 17.6 CONCLUSIONS

Emerging knowledge about the molecular mechanism of apoptosis dysregulation in cancer and leukemia has revealed a plethora of potential drug discovery targets. Detailed structural analysis of apoptosis proteins and studies of their biochemical mechanisms have suggested strategies for lead generation, resulting in numerous
novel chemical entities (NCEs) with mechanism-based activity. Although providing encouraging proof-of-principle data that help validate several apoptosis targets, much work lies ahead in terms of optimizing the spectrum of activity of compounds that interact with multiple members of apoptosis protein families, improving their stability and pharmacological properties, establishing optimal formulations for stability and delivery, and elucidating attendant rate-limiting toxicities. Many of the most logical targets for promoting apoptosis of cancer and leukemia cells are technically challenging, often involving either disrupting protein interactions or altering gene expression, as opposed to traditional pharmaceuticals that typically attack the active sites of enzymes. Modern techniques of structure-based drug optimization render this task feasible, but still challenging. Such targets require long-term commitments, often outstripping the usual drug discovery and development cycle time incorporated into the practices of pharmaceutical companies. For those with the stamina, the rewards are likely to be great, creating a new era in cancer therapy where the intrinsic or acquired resistance of malignant cells to apoptosis can be pharmacologically reversed, reinstating natural pathways for cell suicide.

There is good reason to suspect that malignant cells will be preferentially susceptible to restoration of apoptosis sensitivity, compared to normal cells. Cancer cells exhibit a wide variety of abnormal behaviors and molecular processes that normally trigger an apoptosis response, including cell-cycle checkpoint dysregulation, oncogene activation, chromosome segregation defects, cell detachment from substratum, and outgrowth of blood supply (hypoxia). These defects render tumor cells more dependent on apoptosis-suppressing genes and proteins, and thus withdrawing this crutch from malignant cells may promote self-destruction of transformed cells while sparing normal cells. The full validity of this hypothesis awaits experimental verification in human clinical trials, but the present insights from animal studies and the current forays into the clinic provide encouragement. Altogether, therefore, promises are high that apoptosis-based strategies for cancer drug discovery will yield effective weapons for use in the war on cancer, and thus these efforts should be supported and encouraged.

Acknowledgment

We thank M. Hanaii for manuscript preparation.

References


REFERENCES


REFERENCES


MECHANISM-BASED DEVELOPMENT OF MEMANTINE AS A THERAPEUTIC AGENT IN TREATING ALZHEIMER’S DISEASE AND OTHER NEUROLOGIC DISORDERS: LOW-AFFINITY, UNCOMPETITIVE ANTAGONISM WITH FAST OFF-RATE

Huei-Sheng Vincent Chen* and Stuart A. Lipton*

18.1 INTRODUCTION
18.2 EXCITOTOXICITY
18.3 RATIONAL DRUG DESIGN OF CLINICALLY TOLERATED NMDA RECEPTOR ANTAGONISTS
18.4 MEMANTINE
18.5 NITROMEMANTINES
18.6 CONCLUSIONS

ACKNOWLEDGMENTS
REFERENCES

*Corresponding author: hsv_chen@burnham.org; slipton@burnham.org

Drug Discovery Research: New Frontiers in the Post-Genomic Era, Edited by Ziwei Huang
Copyright © 2007 John Wiley & Sons, Inc.
18.1 INTRODUCTION

Neurologic diseases are among the leading causes of death, disability, and economic expense in the world. For example, Alzheimer’s disease (AD) ranks fourth as a cause of mortality in the United States. In fact, it has been estimated that as the population continues to age, treatment of patients with dementia will consume our entire gross national product by the latter decades of this century. Excitotoxic cell death, also termed excitotoxicity, is thought to contribute to neuronal cell injury and death in this and other neurodegenerative disorders. Excitotoxicity is defined as excessive exposure to the neurotransmitter glutamate or overstimulation of its membrane receptors, leading to neuronal injury or death. Glutamate, however, is the major excitatory neurotransmitter in the brain and mediates critical synaptic transmission for the normal functioning of the nervous system. Based on the pharmacology of agonist sites, there are three classes of glutamate-gated ion (or ionotropic) channels, known as \( \alpha \)-amino-3-hydroxy-5-methyl-4-isoxazolepropionate (AMPA), kainate, and \( N \)-methyl-D-aspartate (NMDA) receptors. Among these, the ion channels coupled to classical NMDA receptors (NMDARs) are generally the most permeable to calcium (\( Ca^{2+} \)). Excitotoxic neuronal cell death is mediated in part by overactivation of NMDARs, which results in excessive \( Ca^{2+} \) influx through the receptor’s associated ion channel. As a consequence of \( Ca^{2+} \) accumulation, excessive activation of the NMDARs leads to production of damaging free radicals and other enzymatic processes contributing to cell death (Fig.18.1) (Lipton and Rosenberg, 1994; Lipton and Nicotera, 1998). Many neurodegenerative diseases, including AD, Parkinson’s disease, Huntington’s disease, HIV-associated dementia, multiple sclerosis, amyotrophic lateral sclerosis (ALS), and glaucoma, are caused by different mechanisms but may share a final common pathway to neuronal injury due to the overstimulation of glutamate receptors, especially of the NMDA subtype (Lipton and Rosenberg, 1994). Acute disorders, such as stroke, CNS trauma, and epilepsy, also manifest a component of excitotoxicity. Hence, NMDAR antagonists could potentially be of therapeutic benefit in a number of acute and chronic neurological disorders manifesting excessive NMDA receptor activities (Kemp and McKernan, 2002).

NMDAR is made up of different subunits: NR1 (whose presence is mandatory), NR2A-D, and, in some cases, NR3A or B subunits. The receptor is probably composed of a tetramer of these subunits. The subunit composition determines the pharmacology and other parameters of the receptor-ion channel complex (Wollmuth and Sobolevsky, 2004). Alternative splicing of some subunits, such as NR1, further contributes to the diversity of pharmacological properties of the receptor (Zukin and Bennett, 1995). The subunits are differentially expressed both regionally in the brain and temporally during development. Physiological NMDAR activity is therefore essential for normal neuronal function (Cull-Candy et al., 2001). Potential neuroprotective agents that block virtually all NMDA receptor activity will therefore very likely have unacceptable clinical side effects. For this reason, many previous NMDAR antagonists have disappointingly failed advanced clinical trials for a number of neurodegenerative disorders. In contrast, studies in our laboratory have shown that the adamantane derivative memantine preferentially blocks excessive NMDA receptor activity without disrupting normal activity. Memantine does this
through its action as a low-affinity, uncompetitive open-channel blocker with a relatively rapid off-rate from the channel. Here we review the molecular mechanism of memantine’s clinically tolerated action and also the basis for the drug’s development in treating several neurological disorders.

Of note, other calcium-permeable channels and routes of calcium entry, such as transient receptor potential (TRP) channels and calcium-permeable AMPA receptors, are also known to contribute to excitotoxicity. In addition, there are alternative approaches to providing protection against NMDAR-mediated excitotoxicity, such as methods uncoupling NMDARs from their downstream effectors (e.g., nitric oxide synthase; also see Kemp and McKernan, 2002). It is not our intention here to cover all possible strategies against calcium-mediated excitotoxicity. The purpose of the present review is to provide a brief and perhaps somewhat surprising primer on

Figure 18.1  Schematic illustration of the apoptotic-like cell death pathways triggered by excessive NMDA receptor activity. The cascade of steps leading to neuronal cell death include: (1) NMDA receptor (NMDA-R) hyperactivation; (2) activation of the p38 mitogen activated kinase (MAPK)–MEF2C (transcription factor) pathway. MEF2 is subsequently cleaved by caspases to form an endogenous dominant-interfering form that contributes to neuronal cell death (Okamoto et al., 2002); (3) toxic effects of free radicals such as nitric oxide (NO) and reactive oxygen species (ROS); and (4) activation of apoptosis-inducing enzymes including caspases and apoptosis-inducing factor. nNOS: nitric oxide synthase; Cyt C: cytochrome C.

(Adapted from the Lipton Web site at www.burnham.org)
excitotoxicity as a promising target of neuroprotective strategies and to present a scientific and clinical overview of the excitotoxicity blocker memantine. Some preliminary information on second-generation memantine derivatives, termed NitroMemantines, is also provided.

18.2 EXCITOTOXICITY

18.2.1 Definition and Clinical Relevance

The ability of the nervous system to rapidly convey sensory information and complex motor commands from one part of the body to another, and to form thoughts and memories, is largely dependent on a single powerful excitatory neurotransmitter, glutamate. There are other excitatory neurotransmitters in the brain, but glutamate is the most common and widely distributed. Most neurons (and also glia) contain high concentrations of glutamate (~10 mM) (Lipton and Rosenberg, 1994); after sequestration inside synaptic vesicles, glutamate is released for very brief amounts of time (milliseconds) to communicate with other neurons via synaptic endings. Because glutamate is so powerful, however, its presence in excessive amounts or for prolonged periods of time can literally excite cells to death. This phenomenon was first documented when Lucas and Newhouse (1957) observed that subcutaneously injected glutamate selectively damaged the inner layer of the retina (representing primarily the retinal ganglion cells). John Olney, in seminal work, later coined the term “excitotoxicity” to describe this phenomenon (Olney, 1969; Olney and Ho, 1970).

A large variety of insults can lead to the excessive release of glutamate within the nervous system and, thus, excitotoxicity. When the nervous system suffers a severe mechanical insult, as in head or spinal cord injury, large amounts of glutamate are released from injured cells. These high levels of glutamate reach thousands of nearby cells that had survived the original trauma, causing them to depolarize, swell, lyse, and die by necrosis. The lysed cells release more glutamate, leading to a cascade of autodestructive events and progressive cell death that can continue for hours or even days after the original injury. A similar phenomenon occurs in stroke; the ischemic event deprives many neurons of the energy they need to maintain ionic homeostasis, causing them to depolarize and propagate the same type of autodestructive events seen in traumatic injury (Lipton, 1993a; Lipton and Rosenberg, 1994). This acute form of cell death occurs by a necrotic-like mechanism, although a slower component leading to an apoptotic-like death can also be present, as well as a continuum of events somewhere between the two (see below).

A slower, subtler form of excitotoxicity is implicated in a variety of chronic and slowly progressing neurodegenerative disorders as well as in the penumbra of stroke damage. In disorders such as AD, Huntington’s disease, Parkinson’s disease, multiple sclerosis, HIV-associated dementia, ALS, and glaucoma, it is hypothesized that chronic exposure to moderately elevated glutamate concentrations or glutamate receptor hyperactivity for longer periods of time than occur during normal neurotransmission trigger cellular processes in neurons that eventually lead to
apoptotic-like cell death, a form of cell death related to the programmed cell death that occurs during normal development (Fig. 18.1) (Lipton and Nicotera, 1998; Bonfoco et al., 1995; Quigley et al., 1995; Vorwerk et al., 1996; Naskar et al., 1999). More subtle or incipient insults can lead to synaptic and dendritic damage in the early stages of disease. This process may be reversible and so is of considerable therapeutic interest.

Importantly, elevations in extracellular glutamate are not necessary to invoke an excitotoxic mechanism. Excitotoxicity can come into play even with normal levels of glutamate if NMDAR activity is increased—for example, when neurons are injured and thus become depolarized (more positively charged); this condition relieves the normal block of the ion channel by magnesium (Mg²⁺) and thus abnormally increases NMDAR activity (Zeevalk and Nicklas, 1992). In addition, increased activity of the enzyme nitric oxide synthase (NOS) is associated with excitotoxic cell death. The neuronal isoform of the enzyme is physically tethered to the NMDA receptor and activated by Ca²⁺ influx via the receptor-associated ion channel, and increased levels of nitric oxide (NO) have been detected in animal models of stroke and several neurodegenerative diseases (Mullins et al., 1996).

18.2.2 Links Between Vascular Dementia and Excitotoxic Damage

As stated earlier, the glutamate content of most cells in the brain is approximately 10 mM. Because of the activity of glutamate transporters, most of this glutamate is intracellular. The extracellular glutamate concentration in the brain has been estimated to be approximately 0.6 μM. The sensitivity to excitotoxicity of cultured cortical neurons isolated away from astrocytes or of hippocampal neurons in intact tissue is approximately 2–5 μM glutamate (Lipton and Rosenberg, 1994). Therefore, the ambient concentrations of glutamate are close to those that can cause neuronal death, and it is important that extracellular glutamate concentration and compartmentalization be exquisitely controlled to prevent excitotoxicity. On the other hand, with 10 mM glutamate present in cells, the potential for disaster is obviously great.

Extracellular glutamate levels have been shown to rise in the face of hypoxic–ischemic insults. There will probably prove to be several mechanisms for the excess accumulation of glutamate, even in a single disorder such as ischemia (Lipton and Rosenberg, 1994). Energy failure might cause abnormal accumulations of glutamate either by impairment of uptake (into neurons and especially astrocytes) mediated via glutamate transporters or by reversal of the direction of transport. This series of events would be followed by injury to some neurons and abnormal potentiation of glutamate release from others. With glutamate release from injured neurons and excess physiological release from otherwise intact neighboring neurons, the process might then develop a self-propagating, vicious cycle that extends the area of neuronal damage.

Deprivation of oxygen and glucose—for example, during ischemia—cause a decrement in the production of high-energy phosphate compounds and “energy failure.” However, short-term energy failure per se is not particularly toxic to neurons. What does make energy failure highly neurotoxic is the activation of glutamate receptor-dependent mechanisms. If suitable glutamate antagonists block
these mechanisms, then neurons can survive a period of oxygen and metabolic substrate deprivation (reviewed by Lipton and Rosenberg, 1994).

18.2.3 Possible Links Between Excitotoxic Damage and Alzheimer’s Disease

There are several potential links between excitotoxic damage and the primary insults of Alzheimer’s disease, which, based on rare familial forms of the disease, are believed to involve toxicity from misfolded mutant proteins (reviewed by Rogawski and Wenk, 2003). These proteins include soluble oligomers of β-amyloid peptide (Aβ) and hyperphosphorylated tau proteins (Selkoe, 2001). For example, oxidative stress and increased intracellular Ca^{2+} generated by Aβ have been reported to enhance glutamate-mediated neurotoxicity in vitro. Additional experiments suggest that Aβ can increase NMDA responses and thus excitotoxicity (Koh et al., 1990; Mattson et al., 1992; Wu et al., 1995). Another potential link comes from recent evidence that glutamate transporters are downregulated in Alzheimer’s disease and that Aβ can inhibit glutamate reuptake or even enhance its release (Topper et al., 1995; Harkany et al., 2000). Finally, excessive NMDAR activity has been reported to increase the hyperphosphorylation of tau, which contributes to neurofibrillary tangles and is involved in NMDA-mediated neurotoxicity (Couratier et al., 1996). The NMDAR antagonist memantine has been found to offer protection from these aforementioned neurotoxic processes (see below).

18.2.4 Pathophysiology of Excitotoxicity: Role of the NMDA Receptor

NMDA receptors are implicated in neuronal survival and maturation (Simon et al., 1984; Balazs et al., 1989), neuronal migration (Komuro and Rakic, 1993), induction of long-term potentiation (LTP, a cellular/electrophysiological correlate of learning and memory formation; Bliss and Collingridge, 1993), formation of sensory maps (Cline et al., 1987; Simon et al., 1992), and neurodegeneration (Meldrum and Garthwaite, 1990; Lipton, 1993b; Lipton and Rosenberg 1994, Cull-Candy et al., 2001). In contrast to AMPA- and kainate-type glutamate receptors, all functional NMDARs are heteromultimers (Wollmuth and Sobolevsky, 2004). Conventional NMDARs composed of NR1 and NR2A-D subunits require dual agonists, glutamate and glycine, for activation (Fig. 18.2A). The activity of the NMDAR-associated channel is modulated by a voltage-dependent block of Mg^{2+} (Nowak et al., 1984; Mayer et al., 1984), and the channel manifests high permeability to Ca^{2+} (Dingledine et al., 1999). NMDA is generally not thought to be an endogenous substance in the body; it is an experimental tool that is highly selective for this subtype of glutamate receptor, hence its name. Under normal conditions of synaptic transmission, the NMDAR-channel is gated by extracellular Mg^{2+} sitting in the channel and is only activated for brief periods of time. This brief opening of the NMDARs allows Ca^{2+} (and other cations) to move into the cell for the subsequent physiological functions. Under pathological conditions, however, overactivation of the receptor relieves the Mg^{2+} block and causes an excessive amount of Ca^{2+} influx into the nerve cell, which
then triggers a variety of processes that can lead to necrosis, apoptosis, or dendritic/synaptic damage. These detrimental processes include (a) Ca\(^{2+}\) overload of mitochondria, resulting in oxygen free radical formation, activation of caspases, and release of apoptosis-inducing factor; (b) Ca\(^{2+}\)-dependent activation of neuronal NOS, leading to increased NO production and the formation of toxic peroxynitrite (\(\text{ONOO}^{-}\)) and S-nitrosylated glyceraldehyde-3-phosphate dehydrogenase (GAPDH); and (c) stimulation of mitogen-activated protein kinase p38 (MAPK p38), which
activates transcription factors that can go into the nucleus to influence neuronal injury and apoptosis (Fig. 18.1; Dawson et al., 1991, 1993; Lipton et al., 1993; Bonfoco et al., 1995; Yun et al., 1998; Budd et al., 2000; Okamoto et al., 2002; Hara et al., 2005).

In NMDARs, the following membrane topology has been proposed (Fig. 18.2B, reviewed by Wollmuth and Sobolevsky, 2004): (1) The N-terminal domain contains the first ~380 amino acids that are related to the bacterial periplasmic binding protein sequence designated leucine/isoleucine/valine-binding protein (LIVBP), the Zn$^{2+}$ binding site, the proton site, and other modulatory sites; (2) four transmembrane domains (M1-M4) are present, and the selectivity filter of the channel pore is formed by M2 (a P loop region); (3) the ligand-binding domains are formed by the pre-M1 (S1) and M3-M4 linker region (S2); (4) a cytoplasmic C-terminal domain interacts with intracellular proteins (Dingledine et al., 1999); and (5) the pre-M1 segment, the C-terminal portion of the M3 segment, and the N-terminal region of the M4 segment form the extracellular channel vestibule (Beck et al., 1999).

In general, NMDAR antagonists can be categorized pharmacologically into four major groups according to site of action on the receptor-channel complex (Wong and Kemp, 1991): Drugs acting at the (1) NMDA (agonist) recognition site, (2) glycine (co-agonist) site, (3) channel pore, and (4) modulatory sites, such as the redox modulatory site, the proton sensitive site, the high-affinity Zn$^{2+}$ site, and the polyamine site. The degree of NMDAR activation and consequent influx of Ca$^{2+}$ and Na$^+$ into the cell can be altered by higher levels of agonists and by substances binding to one of the modulatory sites on the receptor. The two modulatory sites most relevant to this review are open-channel blocker sites within the ion channel pore and S-nitrosylation site(s) located toward the N-terminus (and hence extracellular region) of the receptor. Note that S-nitrosylation reactions represent transfer of NO to a thiol or sulfhydryl group (—SH) of a critical cysteine residue (Lipton et al., 2002). This reaction modulates protein function, in this case decreasing channel activity associated with stimulation of the NMDA receptor. Each of these sites can be considered potential targets for therapeutic intervention to block excitotoxicity, as explained below. Moreover, other modulatory sites also exist on the NMDA receptor and may in the future prove to be of therapeutic value. These include binding sites for Zn$^{2+}$, polyamines, the drug ifenprodil, and a pH (i.e., proton)-sensitive site (Kemp and McKernan, 2002). Additionally, three pairs of cysteine residues at extracellular domains contribute to the redox sites and can modulate NMDAR function by virtue of their redox sensitivity (Lipton et al., 2002). These redox-sensitive cysteine residues may constitute a unique “NO-reactive molecular oxygen sensor” in the brain, enhancing the degree of downregulation of NMDA receptor function by S-nitrosylation in the presence of low pO$_2$ levels and thus dictating the pathological effects of hypoxia that are mediated via the receptor (Takahashi et al., 2007).

18.3 RATIONAL DRUG DESIGN OF CLINICALLY TOLERATED NMDA RECEPTOR ANTAGONISTS

Excitotoxicity is a particularly attractive target for neuroprotective efforts because it is implicated in the pathophysiology of a wide variety of acute and chronic...
neurodegenerative disorders (Lipton and Rosenberg, 1994). The challenge facing those trying to devise strategies for combating excitotoxicity is that the same processes which, in excess, lead to excitotoxic cell death are, at lower levels, absolutely critical for normal neuronal function. To be clinically acceptable, an anti-excitotoxic therapy must block excessive activation of the NMDAR while leaving normal function relatively intact in order to avoid side effects. Drugs that simply compete with glutamate or glycine at the agonist binding sites block normal function and therefore do not meet this requirement, and thus they have failed in clinical trials to date because of side effects (drowsiness, hallucinations, and even coma) (Hickenbottom and Grotta, 1998; Lutsep and Clark, 1999; Palmer, 2000). In fact, competitive antagonists compete one for one with the agonist (glutamate or glycine) and therefore will block healthy areas of the brain (where lower, more physiological levels of these agonists exist) before they can affect pathological areas (where higher levels of agonists accumulate). Thus, such drugs would preferentially block normal activity and would most likely be displaced from the receptor by the high concentrations of glutamate for prolonged periods that can exist under excitotoxic conditions.

### 18.3.1 Advantages of Uncompetitive Antagonism and Open-Channel Block

The term “open-channel blocker of the NMDAR” means that the drug enters the receptor-associated ion channel only when it is open. Importantly, this type of drug will be most effective in the face of excessive (pathological) activity because statistically more channels are open and available to be blocked. This mechanism of inhibition, whose action is contingent upon prior activation of the receptor by the agonist, is defined as “uncompetitive” antagonism. Open-channel block is a most appealing strategy for therapeutic intervention during excessive NMDAR activation since the action of blockade requires prior activation of the receptors. This property, in theory, leads to a higher degree of channel blockade in the presence of excessive levels of glutamate and little blockade at relatively lower levels—for example, during physiological neurotransmission (see Section 18.4 for details; Rang, 1981; Chen, 1992; Chen et al., 1992; 1998; Lipton, 1993b; Lipton and Rosenberg, 1994; Chen and Lipton, 1997).

In fact, an uncompetitive open-channel blocker would prevent more severe excitotoxic processes better than mild disease. One would predict, for example, that moderate-to-severe dementia, involving excessive NMDAR activity leading to neuronal injury or death, would be treated more effectively than mild dementia or other excitotoxic disorders, involving only somewhat increased physiological firing. (Of course, these drugs will not reverse very severe disease because the neurons will already be lost.) Although preferential neuroprotection from moderate-to-severe excitotoxic processes seems counterintuitive, the drug’s uncompetitive mechanism of action readily explains this uncanny phenomenon.

### 18.3.2 The Importance of Off-Rate from Channel Block

Importantly, we have shown that among open-channel blockers of the NMDAR, the details of the kinetics of blockade are important in order to avoid side effects...
(Chen, 1992; Chen et al., 1992; Chen and Lipton, 1997). A relatively fast off-rate (and hence short dwell time in the channel) would prevent the drug from accumulating in open channels. This avoids progressive blockade of normal synaptic transmission. In contrast, a drug with a slow off-rate would build up in the ion channels that underlie synaptic events and consequently interfere with normal neurological function. The apparent affinity of a channel-blocking drug is related to its off-rate divided by its on-rate. At a given membrane potential, the on-rate is not only a property of diffusion and channel open probability, but also the drug’s concentration. In contrast, the off-rate is an intrinsic property of the drug–receptor complex, unaffected by drug concentration. A relatively fast off-rate is a major contributor to a drug’s low affinity for the channel pore. Thus, we propose that a clinically tolerated neuroprotective drug would consist of a low-affinity, open-channel blocker with a relatively fast off-rate. Hence, the drug would not substantially interfere with normal synaptic neurotransmission in an accumulative fashion. As a result, the drug would be both effective and well-tolerated.

As a useful analogy, the NMDA receptor can be thought of like a television set. The agonist sites are similar to the “on/off” switch of the television. Drugs that block here cut off all normal NMDAR function. What we need to find is the equivalent of the “volume” control (or in biophysical terms, the gain) of the receptor. Then, excessive Ca\(^{2+}\) influx through the NMDAR-associated ion channel would be prevented by simply turning down the “volume” of the Ca\(^{2+}\) flux toward normal values. A blocker that binds at a site within the channel, similar to the action of physiological levels of Mg\(^{2+}\), could act as a sensor and provide an “automatic” volume control. Importantly, the automatic volume control needs to reach an optimal level. In the case of Mg\(^{2+}\) itself, the block is too ephemeral, a so-called “flickery block,” and the cell continues to depolarize (become positively charged because of Ca\(^{2+}\) and Na\(^{+}\) entry) until Mg\(^{2+}\) is repelled, and the block is totally relieved. Hence, in most cases Mg\(^{2+}\) does not effectively block excessive Ca\(^{2+}\) influx to the degree needed to prevent neurotoxicity. If, on the other hand, a channel blocker binds with too high an affinity, it will accumulate in the channels, block normal activation, and thus prove clinically unacceptable. Following the television set analogy, turning the volume all the way down is as bad as turning off the “on/off” switch in terms of normal functioning of the television. This is the case with MK-801: It is a very good blocker of excitotoxicity, but because its dwell time in the ion channel is so long (reflecting its slow off-rate and high affinity), it progressively blocks critical normal functions. MK-801 can thus produce coma. Drugs with slightly shorter but still excessive dwell times (slow off-rates) make patients hallucinate (e.g., phencyclidine, also known as Angel Dust) or feel drowsy so that they classify as anesthetics (e.g., ketamine).

A clinically tolerated NMDAR antagonist would not make a patient feel drowsy, hallucinate, or become comatose, and in fact it should spare normal neurotransmission while blocking the ravages of excessive NMDA receptor activation. An uncompetitive, open-channel mechanism of blockade coupled with a longer dwell time in the channel (and consequently a slower off-rate) than Mg\(^{2+}\) but a substantially shorter dwell time (faster off-rate) than MK-801 would yield a drug that blocks NMDAR-operated channels only when they are excessively open, while relatively sparing normal neurotransmission.
18.3.3 Targeting Therapeutic Agents to the Pathologically Active Area

An NMDAR open-channel blocker binds increasingly well to neurons manifesting excessive channel activity; these are by definition potentially vulnerable neurons. Hence, we can attach other anti-apoptotic or pro-survival moieties to an open-channel blocker to form a new adduct that is targeted to vulnerable neurons, giving the second moiety specificity of action to the sick neuron. Using this concept, we propose that a series of second-generation drugs can be formulated that will have even greater neuroprotective properties than the original. These second-generation drugs can take advantage of the fact that the NMDAR has other modulatory sites (or “volume” controls) in addition to its ion channel that offer safe but effective clinical intervention.

One example of an additional modulatory site on the NMDAR that we can take therapeutic advantage of involves the action of nitric oxide (NO). We have shown that transfer of NO to thiol (−SH) groups on critical cysteine residues of the NMDAR (a reaction termed S-nitrosylation) decreases excessive receptor activity (Lei et al., 1992; Lipton et al., 1993; Choi et al., 2000; Takahashi et al., 2007). However, if administered systemically, NO can cause serious side effects, including severe hypotension (low blood pressure) by virtue of its ability to produce vasodilatation, and may even be toxic by nitrosylating other targets such as GAPDH or reacting with superoxide anion (O₂⁻) to form peroxynitrite (ONOO⁻). To avoid this problem, we can tether the NO group to an appropriate open-channel blocker in order to specifically target NO to the NMDAR nitrosylation sites. To date, such combinatorial drugs, linking the principles of open-channel block and S-nitrosylation of the NMDAR to provide two “volume controls,” show great clinical potential. We propose that this newly recognized mode of action for drugs be designated “pathologically activated therapeutics” (PAT, meaning a gentle tap). By virtue of their relatively gentle binding, PAT drugs work best under pathological conditions, while exerting minimal effects on normal brain activity. We believe that these simple concepts embody the future of clinically tolerated neuroprotective drug design. Interestingly, we were the first to show that memantine fulfills these mechanistic criteria (Chen, 1992; Chen et al., 1992; Chen and Lipton, 1997).

18.4 MEMANTINE

18.4.1 Uncompetitive Open-Channel Block of NMDARs

Memantine (MEM, 1-amino-3, 5-dimethyl-adamantane, Fig. 18.3) was first synthesized by Eli Lilly and Company and patented in 1968, as documented in the Merck Index. It is a derivative of amantadine, an anti-influenza agent (Tominack and Hayden, 1987). Memantine has a three-ring (adamantane) structure with a bridgehead amine (−NH₂) that under physiological conditions carries a positive charge (−NH₃⁺). Memantine has been used clinically with an excellent safety record for over 20 years in Europe to treat Parkinson’s disease, spasticity, convulsions, vascular dementia (Ditzler, 1991), and Alzheimer’s disease (Fleischhacker et al., 1986). The
reported efficacy of amantadine and memantine in Parkinson’s disease, which was discovered by serendipity in a patient taking amantadine for influenza, led scientists to believe that these compounds were dopaminergic or possibly anticholinergic drugs. Work at a small German company named Merz first suggested that the drug might be an NMDA receptor inhibitor (Bormann, 1989; Kornhuber et al., 1989). We are the first group to characterize that memantine acts as an open-channel blocker of the NMDAR-coupled channel pore. We found that the drug blocks NMDA-evoked responses via uncompetitive antagonism with a 50% inhibition constant (IC50) of ~1 μM at -60 mV. Kinetic analysis with single exponential fitting revealed that the on-time (time until peak blockade) of 12 μM memantine was approximately 1 sec, while the off-time constant (recovery time constant, τ) from the effect was ~5 sec. (C) Dose–response curve for MEM constructed using I_{MEM}/I_{control} (%) versus MEM concentration.

Figure 18.3  Structure, kinetic, and dose–response analysis of memantine (MEM). (A) Left: Chemical structure of amantadine. Right: Chemical structure of memantine, which has methyl group (–CH₃) side chains (unlike amantadine). (B) Blockade of 200 μM NMDA-activated currents by 1 and 12 μM MEM recorded from a solitary neuron at a holding potential of -60 mV. Kinetic analysis with single exponential fitting revealed that the on-time (time until peak blockade) of 12 μM memantine was approximately 1 sec, while the off-time constant (recovery time constant, τ) from the effect was ~5 sec. (C) Dose–response curve for MEM constructed using I_{MEM}/I_{control} (%) versus MEM concentration.

reported efficacy of amantadine and memantine in Parkinson’s disease, which was discovered by serendipity in a patient taking amantadine for influenza, led scientists to believe that these compounds were dopaminergic or possibly anticholinergic drugs. Work at a small German company named Merz first suggested that the drug might be an NMDA receptor inhibitor (Bormann, 1989; Kornhuber et al., 1989). We are the first group to characterize that memantine acts as an open-channel blocker of the NMDAR-coupled channel pore. We found that the drug blocks NMDA-evoked responses via uncompetitive antagonism with a 50% inhibition constant (IC50) of ~1 μM at -60 mV. Kinetic analysis with single exponential fitting revealed that the on-time (time until peak blockade) of 12 μM memantine was approximately 1 sec, while the off-time constant (recovery time constant, τ) from the effect was ~5 sec. (C) Dose–response curve for MEM constructed using I_{MEM}/I_{control} (%) versus MEM concentration.
within the effective range of its NMDA-antagonistic action \( (K_i \approx 1 \mu M) \) but is below the effective level of memantine at any other known receptor or ligand-gated channel (Chen et al. 1992, but see below). The antagonistic action of memantine on NMDARs is therefore thought to be the principal mechanism in the therapy of Parkinson’s disease and possibly cerebral ischemia, dementia, and epilepsy (Lipton, 1993b; Lipton and Rosenberg, 1994; Rogawski, 2000).

Most importantly, we also showed why memantine could be clinically tolerated as an NMDA receptor antagonist; namely, it was an uncompetitive open-channel blocker with a dwell time/off-rate from the channel that limited pathological activity of the NMDA receptor while sparing normal synaptic activity (Chen et al., 1992; 1998; Chen and Lipton, 1997). We illustrate this most astonishing property of memantine in Fig. 18.4 and Fig. 18.5 (Chen et al., 1992). In this experiment, the concentration of memantine was held constant (at a clinically achievable low micromolar level) while the concentration of NMDA was increased over a wide range. It was found that the degree to which this fixed concentration of memantine blocked NMDA receptor activity actually increased as the NMDA concentration was increased to pathological levels. This is classical “uncompetitive” antagonist behavior. In fact, the component of the excitatory postsynaptic current due to physiological activation of NMDA receptors is inhibited by only 10 to 30% (Chen et al., 1998). During prolonged activation of the receptor, however, as occurs under excitotoxic conditions, memantine becomes a very effective blocker. In essence, memantine only acts under pathological conditions without much affecting normal function, thus relatively sparing synaptic transmission, preserving long-term potentiation, and maintaining physiological function on behavioral tests such as the Morris water maze (Chen et al., 1998). This notion is supported by the safety and efficacy profiles of memantine in two recent clinical trials for the treatment of Alzheimer’s disease (Reisberg et al., 2003; Tariot et al., 2004). This strategy of selecting NMDAR antagonists of low-affinity/fast off-rate is in contrast to most drug discovery by big

![Figure 18.4](image-url) Paradoxically, a fixed dose of memantine (i.e., 6 \( \mu M \)) blocks the effect of increasing concentrations of NMDA to a greater degree than lower concentrations of NMDA. This finding is characteristic of an uncompetitive antagonist.
Pharma, which uses high-affinity screens of the target. Neuroprotective agents that work by high-affinity binding to the NMDAR result in blockade of virtually all receptor activity; thus, these drugs manifest unacceptable clinical side effects. Instead, several years ago, we had proposed to protect the brain with drugs that do not bind very well under physiological conditions but are nevertheless selective under pathological conditions for a particular target, such as the NMDAR (Chen et al., 1992). Importantly, one should not confuse affinity with selectivity; as long as a drug acts selectively and specifically on the target of interest and the effective concentration can be achieved, a high affinity per se is not the key issue.

18.4.2 Memantine Interacts with the Intracellular Mg\textsuperscript{2+} Blocking Site in the NMDAR Channel Pore

Our earlier studies indicated that memantine exerts its effect on NMDA receptor activity by binding at or near the Mg\textsuperscript{2+} site within the ion channel (Chen, 1992; Chen et al., 1992; Chen and Lipton, 1997). Because of its interaction with external Mg\textsuperscript{2+} and based on mutational analysis of the NMDAR by others (Kashiwagi et al., 2002), the specific site of memantine action was assumed to be near the external Mg\textsuperscript{2+} blocking site at the selectivity filter region of the NMDAR-associated channel (Danysz and Parsons, 2003). This region is formed by asparagine (N) residues at the “N site” of NR1 and “N + 1 site” of NR2 subunits (reviewed by Dingledine et al., 1999). Compared to physiological block by external Mg\textsuperscript{2+}, a common explanation for the safety and effectiveness of memantine has been that memantine represented a “better magnesium,” manifesting a somewhat slower unblocking rate, moderate voltage dependence, and slightly higher affinity (Danysz and Parsons, 2003). However, when applied from the intracellular versus extracellular surface, Mg\textsuperscript{2+} interacts differently on the N-site residues of NR1 and NR2 subunits (Wollmuth et al., 1998a,b). The N-site asparagine of the NR1 subunit represents the dominant blocking site for intracellular Mg\textsuperscript{2+}, whereas the N and N + 1 site asparagines of the NR2A subunit form the critical blocking site for extracellular Mg\textsuperscript{2+}. We recently performed a series of experiments using point mutations and the substituted cysteine accessibility method (SCAM) to show that the N-site asparagine of the NR1 subunit, located at the selectivity filter of the NMDAR-associated channel, is the specific and predominant blocking site for memantine (Chen and Lipton, 2005). The N and N + 1 sites of NR2A subunits provide the major electrostatic interaction with memantine upon binding to this deep, specific site (Fig. 18.6). The differential contribution to memantine block by the N- and N + 1 site asparagines in NR1 and NR2 subunits is reminiscent of their effects on intracellular Mg\textsuperscript{2+} blockade (Wollmuth et al., 1998b). The distinct patterns of interaction of memantine with the channel selectivity filter may confer upon memantine unique kinetic features leading to the drug’s excellent clinical tolerability. In line with these results, memantine, in the absence of extracellular Mg\textsuperscript{2+}, displays minimal differences in blocking NMDARs containing various NR2 subunits (Bresink et al., 1996). We also found that none of the NR1 splice variants differentially affected memantine antagonism in the absence of external Mg\textsuperscript{2+} (Chen and Lipton, 2005).
Figure 18.5 A trapping, uncompetitive scheme for memantine action. (A) Response to 200 μM NMDA was first blocked by 2 μM MEM at -60 mV, and then both agonists and antagonists were washed out by rapidly exposing the cell to a solution containing only the control solution for 11.3 sec to demonstrate MEM trapping (left). After washout with control solution, a second application of NMDA alone displayed a fast-rising phase of channel activation followed by slow relaxation, representing recovery from MEM blockade. This slow relaxation was similar to the regular recovery phase from MEM block that occurred without trapping (right). (B) Difference in predicted degrees of blockade between noncompetitive and uncompetitive antagonist action of memantine (MEM). Top left: Scheme for noncompetitive antagonism with C representing the closed channel; O, the open channel; C*-MEM, the blocked and closed channel; O-MEM, the open but blocked channel; a, the microscopic on-rate; b, the microscopic off-rate; [MEM], the concentration of the blocker; K, the equilibrium constant for opening from the closed state; and P₀, the open probability of the channel. The affinity of the blocker for the closed and open channel is the same. The open probability of the unblocked channel is the same as that of the blocked channel. Top right: Scheme for uncompetitive antagonism with K/θ (θ = 0.146) is the equilibrium constant for opening from the C*-MEM state; P₀, the open probability of the blocked channel, and the rest of the symbols have the same meaning as above. The blocker does not bind to the closed channel in this paradigm. (C) Computer-simulated degree of blockade for a noncompetitive antagonist (dotted line) and uncompetitive antagonist (solid curve) with the models and parameters indicated in A and B. The degree of blockade of NMDA-elicited current in Fig. 18.4 is replotted here. The inhibition equilibrium constant (Kᵯ) for the blocker is 1.2 μM, and [MEM] is 6 μM. The empirical data points were very close to those predicted theoretically for pure uncompetitive antagonism (Chen and Lipton, 1997).
18.4.3 A Second Binding Site for Memantine in NMDA-Gated Channels

Several studies have also reported a second binding site for memantine in NMDA-gated channels (Antonov and Johnson, 1996; Blanpied et al., 1997; Bresink et al., 1996; Sobolevsky et al., 1998). This second site was reported to have a much lower affinity, minimal voltage dependence, and a noncompetitive mechanism of block. There were, however, several differences among these studies, including the estimated IC$_{50}$, the precise degree of voltage dependence, and the location of this blocking site in the NMDA-gated channel. Using point mutations and SCAM, we found that the second (superficial) memantine-blocking site, located at the extracellular vestibule of the channel, appears to be nonspecific and overlaps the site occupied by the nonspecific pore blocker hexamethonium. Residues in the post-M3 segment of the NR1 subunit are not directly involved in memantine binding (Chen and Lipton, 2005).

There is an important therapeutic implication of our recent description of the location of the second memantine binding site that concerns the uncompetitive mechanism of antagonism displayed by memantine. Uncompetitive, unlike competitive or noncompetitive, antagonists can block excessive activation of NMDARs while sparing normal neurotransmission; this is the property most likely responsible for the clinically tolerated mechanism of action of memantine, as explained above. Memantine blocks NMDARs via an uncompetitive mechanism at low micromolar concentrations, yet possesses a noncompetitive component (sometimes called “partial trapping” in the channel) at higher concentrations (Blanpied et al., 1997; Chen and Lipton, 1997). Lipophilic leak of memantine from its blocking site cannot explain this noncompetitive component (Mealing et al., 2001; Bolshakov et al., 2003). Instead, the noncompetitive behavior may be explained by the second site of memantine binding that we identified, which has very low affinity and is located at the channel vestibule. Occupancy by memantine of this shallow site may allow dissociation of the drug in either the open or closed conformation, resulting in a form of noncompetitive antagonism. This superficial site of memantine action, however, represents nonspecific binding and may also explain the noncompetitive component of many other very low-affinity, open-channel blockers.

Most importantly, we showed that a large difference in the affinity between these two sites of memantine binding is crucial for maintaining the selectivity of this type of “low-affinity” NMDAR open-channel blocker (Chen and Lipton, 2005). We believe that the distinct patterns of interaction and the relative degree of affinity of memantine for these two binding sites contribute to the drug’s excellent pharmacological profile of clinical tolerability.

18.4.4 Voltage Dependence, Partial Trapping, and Lipophilic Leak of Memantine

Many possible factors have also been suggested for memantine’s clinical tolerability, including moderate-to-low affinity, moderate voltage dependence, fast blocking and unblocking kinetics, and partial trapping in the NMDAR-associated channel.
(Rogawski and Wenk, 2003). All these macroscopic explanations are based on the assumption that memantine and other open-channel blockers bind at the same site as extracellular \(\text{Mg}^{2+}\) in the channel selectivity filter. We were recently the first to show that this assumption is incorrect. Instead, memantine interacts with the intracellular \(\text{Mg}^{2+}\) blocking site, which is located slightly deeper than the extracellular \(\text{Mg}^{2+}\) blocking site (Chen and Lipton, 2005). We also showed that the second, superficial site of memantine action is nonspecific and may explain the noncompetitive (or nontrapping) component of memantine at near millimolar concentrations. For so-called “low-affinity” NMDA open-channel blockers, the apparent affinity, unblocking rate, and voltage dependence as well as the lipophilic leak (or closed-channel egress) (Blanpied et al., 1997) and degree of trapping of each open-channel blocker will represent “mixed” properties of both sites if the relative affinities are not too far apart. Therefore, despite prior reports, none of these properties at one site alone can explain the variable clinical tolerability of low-affinity NMDAR antagonists. In the case of memantine, however, the affinities of the two sites are sufficiently distinct so that the pharmacological properties of the specific site may account for its lack of side effects (Chen and Lipton, 2005). In our hands, memantine at therapeutic concentrations displays minimal closed-channel block or egress (minimal lipophilic leak) and therefore behaves like a perfect uncompetitive blocker [Fig. 18.6; see Chen and Lipton, (1997, 2005) for details].

**Figure 18.6** Atomic model showing two memantine binding sites in the channel permeation pathway of the NMDAR. Locations of memantine binding sites in the channel permeation pathway are shown at the level of the channel selectivity filter (the specific site) and at the L651 residue of the NR1 subunit (the nonspecific site). Internal permeant monovalent cation (\(\text{K}^+\) or \(\text{Cs}^+\)) can compete with the externally applied MEM for binding, and MEM binding interacts with the intracellular \(\text{Mg}^{2+}\) binding site (Chen and Lipton, 1997, 2005).
18.4.5 Other Possible Effects of Memantine

At various concentrations, many different mechanisms of action of memantine have been reported: (1) antiviral action by inhibition of viral coat protein function (Tominack and Hayden, 1987), (2) inhibition of muscle-type nicotinic acetylcholine (ACh) receptors at the frog neuromuscular junction (Masuo et al., 1986), (3) antagonistic effects on NMDARs, (4) potentiation of strychnine-sensitive, glycine-activated currents (Lampe and Bigalke, 1991) (however, this result has not been substantiated by us or others), and (5) at concentrations above 100 μM, memantine may block voltage-sensitive sodium channels (Netzer and Bigalke, 1990) and has nonspecific effects on cell membranes (Osborne et al., 1982; Wesemann et al., 1983). In rat brain preparations, memantine does not interact directly with dopamine, opioid, GABA, or α1- and α2-adrenergic receptors and has no effect on the uptake of the neurotransmitters noradrenaline or serotonin [reviewed in Osborne et al. (1982)].

Memantine has been found to block serotonin (5-HT₃) receptor channels at concentrations approaching those that block NMDA receptor channels (Rammes et al., 2001; Reiser et al., 1988). Memantine’s effect on 5-HT₃ receptors may possibly further enhance cognitive performance. Memantine has also been reported to inhibit α7 nicotinic acetylcholine (ACh) receptors in the Xenopus oocyte expression system and in rat cultured hippocampal neurons with an affinity similar to that reported for native NMDA receptors (Maskell et al., 2003; Aracava et al., 2005). These results, however, remain controversial, in part because desensitizing currents were evaluated rather than steady-state currents, and their significance thus needs to be confirmed (Banerjee et al., 2005).

Concerning its use in Alzheimer’s disease, in the rat, memantine at therapeutic concentrations reduced the loss of cholinergic neurons in the nucleus basalis caused by NMDA-mediated toxicity or mitochondrial toxins [reviewed in Rogawski and Wenk (2003)]. Memantine has been reported to offer protection from neurotoxicity engendered by intrahippocampal injection of Aβ (Miguel-Hidalgo et al., 2002) and to enhance the processing of non-amyloidogenic β-amyloid precursor protein (Rogawski and Wenk, 2003). Memantine also improved performance on behavioral tests (T-maze and Morris water maze) in a transgenic mouse model of familial AD consisting of a mutant form of amyloid precursor protein and presenilin 1 (Tanila et al., 2003). Additionally, memantine was recently found to reduce tau hyperphosphorylation in culture (Iqbal et al., 2003). Chronic infusion of memantine also attenuated neuronal loss, improved short-term memory impairment, and reduced learning deficits and neurotoxicity caused by quinolinic acid-induced entorhinal cortex lesions in the rat [reviewed by Parsons et al. (1999)]. However, the exact mechanism(s) of protection by memantine from these animal or culture models of AD remains to be elucidated.

18.4.6 Neuroprotective Efficacy

The neuroprotective properties of memantine have been studied in a large number of in vitro and in vivo animal models by several laboratories [reviewed in Parsons et al. (1999)]. Among neurons protected in this manner both in culture and in vivo are cerebrocortical neurons, cerebellar neurons, and retinal neurons (Chen et al., 1992, 1998; Lipton, 1992; Osborne, 1999). Additionally, in a rat model of stroke,
memantine, given as long as 2 hours after the ischemic event, reduces the amount of brain damage by approximately 50% (Chen et al., 1992, 1998).

A series of human clinical trials that investigated the efficacy of memantine for the treatment of AD, vascular dementia, HIV-associated dementia, diabetic neuropathic pain, depression, and glaucoma have recently been completed. Some of these studies have only recently been completed and remain unpublished at this time except in abstract form. One recent high-profile publication reported the result of a U.S. phase 3 (final) clinical study showing that memantine (20 mg/day) is efficacious for moderate-to-severe AD (Reisberg et al., 2003). Another study reported that, in combination with Aricept, memantine treatment offers some improvement of memory and function in moderate-to-severe Alzheimer’s patients (Tariot et al., 2004). Both studies revealed excellent clinical tolerance and minimal side effects from therapeutic doses of memantine. These positive results of clinical studies for treating AD convinced the European Union and the U.S. FDA to approve memantine for the treatment of this form of dementia. Concerning other forms of dementia, one European multicenter, randomized, controlled trial reported that memantine was beneficial in severely demented patients, probably representing both AD and vascular dementia (Winblad and Poritis, 1999). Another recent publication of a randomized, placebo-controlled clinical trial also described significant benefit from memantine therapy (20 mg/day) in mild-to-moderate vascular dementia (Orgogozo et al., 2002). These clinical trials are summarized in Table 18.1. Most trials have reported minimal adverse effects from memantine. The only memantine-induced side effects encountered were rare dizziness and occasional restlessness/agitation at higher doses (40 mg/day), but these effects were mild and dose-related. Memantine is also under investigation as a potential treatment for other neurodegenerative disorders, including HIV-associated dementia, neuropathic pain, and glaucoma as well as depression and movement disorders.

Importantly, as we discovered and outlined above, the uncompetitive mode of memantine action would predict that, at a fixed dose, memantine would work better for a severe condition (e.g., excessive glutamate receptor activity to the point of causing cell death) than for more mild conditions manifested by slightly elevated synaptic transmission. Bearing this out, recent studies support that memantine may have a larger effect in moderate-to-severe dementia than in mild dementia. Another case in point is neuropathic pain, which is thought to be mediated at least in part by excessive NMDAR activity. Given the uncompetitive antagonism of memantine, more severe pain (e.g., the nocturnal pain of diabetic neuropathy) might be expected to benefit from memantine to a greater extent than more mild forms of neuropathic pain. In fact, a phase 2B clinical trial suggested that this is indeed the case, whereas preliminary reports indicate that more

<table>
<thead>
<tr>
<th>TABLE 18.1 Clinical Trials with Memantine</th>
</tr>
</thead>
<tbody>
<tr>
<td>• German/Merz phase 3 trial for vascular dementia and Alzheimer’s disease +</td>
</tr>
<tr>
<td>• Karolinska/Italian phase 3 trial for vascular dementia +</td>
</tr>
<tr>
<td>• Two USA multicenter phase 3 trials for Alzheimer’s disease +</td>
</tr>
<tr>
<td>• United Kingdom phase 3 trial for vascular dementia +</td>
</tr>
<tr>
<td>• French phase 3 trial for Vascular dementia +</td>
</tr>
<tr>
<td>• USA phase 2 trials for neuropathic pain and HIV-associated dementia +/-</td>
</tr>
</tbody>
</table>

SOFTbank E-Book Center Tehran, Phone: 66403879,66493070 For Educational Use.
mild pain conditions were not statistically benefited by memantine in phase 2/3 clinical trials. Along these same lines, one would predict that a higher concentration of memantine would be needed to combat pain than to prevent neuronal cell death because greater NMDA receptor activity is associated with cell death (a greater proportion of channels will be blocked by memantine in the face of increasing NMDAR activity). Again, clinical trials have suggested that this is indeed the case because 40 mg/day of memantine have been needed in successful pain studies but only 20 mg/day in severe dementia studies. However, further clinical trials will be necessary to prove the efficacy of memantine for severe neuropathic pain.

As promising as the results with memantine are, we are continuing to pursue ways to use additional modulatory sites (the “volume controls”) on the NMDA receptor to block excitotoxicity even more effectively and safely than memantine alone. New approaches in this regard are explored below.

18.5 NITROMEMANTINES

NitroMemantines are second-generation memantine derivatives designed to have enhanced neuroprotective efficacy without sacrificing safety. As mentioned earlier, a nitrosylation site(s) is located on the N-terminus or extracellular domain of the NMDA receptor, and S-nitrosylation of this site (NO reaction with the sulfhydryl group of the cysteine residue) downregulates (but does not completely shut off) receptor activity (Fig. 18.2). The drug nitroglycerin, which generates NO-related species, can act at this site to limit excessive NMDA receptor activity. In fact, in rodent models, nitroglycerin can limit ischemic damage (Lipton and Wang, 1996), and there is some evidence that patients taking nitroglycerin for other medical reasons may be resistant to glaucomatous visual field loss as well (Zurakowski et al., 1998).

Consequently, we carefully characterized S-nitrosylation sites on the NMDA receptor in order to determine if we could design a nitroglycerin-like drug that could be more specifically targeted to the receptor. In brief, we found that five different cysteine residues on the NMDA receptor could interact with NO. One of these, located at cysteine residue #399 (C399) on the NR2A subunit of the NMDA receptor, mediates approximately 90% of the effect of NO under our experimental conditions (Choi et al., 2000). From crystal structure models and electrophysiological experiments, we further found that NO binding to the NMDAR at C399 apparently induces a conformational change in the receptor protein that makes glutamate and Zn$^{2+}$ bind more tightly to the receptor. The enhanced binding of glutamate and Zn$^{2+}$ in turn causes the receptor to desensitize and, consequently, causes the ion channel to close (Lipton et al., 2002). Electrophysiological studies have demonstrated this effect of NO on the NMDA channel (Lei et al., 1992; Lipton et al., 1993; Choi et al., 2000).

Unfortunately, nitroglycerin is not very attractive as a neuroprotective agent. The same cardiovascular vasodilator effect that makes it useful in the treatment of angina could cause dangerously large drops in blood pressure in patients with dementia, stroke, traumatic injury, or glaucoma. However, the open-channel block mechanism of memantine not only leads to a higher degree of channel blockade in the presence of excessive levels of glutamate, but also can be used as a homing signal for targeting drugs (e.g., the
NO group) to hyperactivated, open NMDA-gated channels. We have been developing combinatorial drugs (NitroMemantines) that theoretically should be able to use memantine to target NO to the nitrosylation sites of the NMDAR in order to avoid the systemic side effects of NO. Two sites of modulation are analogous to having two volume controls on your television set for fine-tuning the audio signal.

Preliminary studies have shown NitroMemantines to be highly neuroprotective in both *in vitro* and *in vivo* animal models. In fact, it appears to be more effective than memantine. Moreover, because of the targeting effect of memantine moiety, NitroMemantines appear to lack the blood pressure lowering effect typical of nitroglycerin. More research still needs to be performed on NitroMemantine drugs, but by combining two clinically tolerated drugs (memantine and nitroglycerin) we created a new, improved class of PAT drugs that are both clinically tolerated and neuroprotective.

**18.6 CONCLUSIONS**

Necrosis- and apoptosis-mediated excitotoxic cell death is implicated in the pathophysiology of many neurological diseases. This type of excitotoxicity is caused, at least in part, by excessive activation of NMDA-type glutamate receptors. Intense insults, such as that occurring in the ischemic core after a stroke, trigger massive stimulation of NMDA receptors, leading to neuronal cell swelling and lysis (necrosis). In contrast, more moderate NMDAR hyperactivity, such as that occurring in the ischemic penumbra of a stroke and in many slow-onset neurodegenerative diseases, results in a moderately excessive influx of calcium ions into nerve cells, which, in turn, triggers free radical formation and multiple pathways leading to the initiation of synaptic damage and apoptotic-like neuronal cell loss (Lipton and Nicotera, 1998). However, NMDA receptor activity is also required for normal neural function. Until recently, all drugs that showed promise as inhibitors of excitotoxicity also blocked normal neuronal function and consequently had severe and unacceptable side effects, so clinical trials for stroke, traumatic brain injury, and Huntington’s disease all failed (Lees et al., 2000; Sacco et al., 2001). In the past decade, we have shown that memantine represents a class of drugs that have relatively low affinity and act as an uncompetitive, open-channel blocker. Due to its uncompetitive antagonism and relatively fast off-rate, memantine blocks excessive NMDAR activation but spares low (physiological) levels of NMDAR activity seen during normal neurotransmission. Importantly, memantine binds at the “intracellular” Mg$^{2+}$ site in the channel pore and displays differential affinity for specific and nonspecific binding sites on the NMDAR. These molecular interactions confer upon memantine favorable kinetic properties that contribute to the drug’s clinical tolerability as well as its neuroprotective profile (Chen and Lipton, 2005). The discovery that memantine, a low-affinity but still highly selective agent with a mechanism of uncompetitive antagonism, is neuroprotective yet clinically tolerated has triggered a paradigm shift in the history of drug development by the pharmaceutical industry. Clinical studies have borne out our hypothesis that low-affinity/fast off-rate memantine is a safe NMDA receptor antagonist in humans and beneficial in the treatment of neurological disorders mediated, at least in part, by excitotoxicity.
The NitroMemantines are second-generation NMDA receptor antagonists that may work even better than memantine. They use the memantine moiety as a homing signal for the targeted delivery of NO to a second modulatory site on the NMDAR. Work is progressing rapidly in this area of investigation.

Further clinical studies of the efficacy of memantine in the treatment of AD, vascular dementia, HIV-associated dementia, glaucoma, and severe neuropathic pain are currently underway, and there is every reason to expect the results to be positive, although this is, of course, not yet proven except in the case of Alzheimer’s disease and possibly vascular dementia (Table 18.1). The efficacy of memantine in neurodegenerative diseases and its ability to protect neurons in animal models of both acute and chronic neurological disorders suggest that memantine and drugs acting in a similar manner could become very important new weapons in the fight against neuronal damage.

ACKNOWLEDGMENTS

We would like to thank our colleagues for their contributions to this work, which is updated here with an emphasis on Alzheimer’s disease. We are especially grateful to Drs. Joachim Bormann, Yun-Beom Choi, and Jonathan S. Stamler for their discussions or collaborations. This work was supported in part by an AHA SDG and NIH grants P01 HD29587, R01 EY50477, and R01 EY09024. Dr. Lipton was a Senior Scholar in Aging of the Ellison Medical Foundation.

REFERENCES


19.1 INTRODUCTION

While human diseases may have various origins and forms—including viral or bacterial infectious diseases; inherited disorders such as muscular dystrophy; metabolic diseases such as obesity; diabetes; aging-associated diseases such as Alzheimer’s; and cancers—many of these diseases can be modeled in animals by means of infectious agents, genetic manipulation, diet or drug-induced regimens, and surgery or xenograft explants. These animal models have offered important insights into the molecular mechanism of human diseases, served as proof-of-concept studies to validate gene targets for drug discovery, and, ultimately, provided indispensable platforms for evaluating the efficacy and safety of drugs in development. Decisions on whether to further or halt the development of a candidate drug often hinges on

*Corresponding author: chen_ling@gibh.ac.cn
comprehensive analysis of its efficacy in proven animal models that faithfully predict
the therapeutic value in treating the corresponding human disease, its pharmacody-
namic/pharmacokinetic characteristics, and, importantly, its short- and long-term
safety profiles.

The genomic era has brought forth tremendous knowledge on the gene sequences
and genomic organizations of humans, rodents, and various other animals. An increasing
amount of information is being obtained on the tissue, developmental, and disease-
related expression patterns of candidate genes. Based on these data, one can easily
identify the corresponding orthologue and homologue of a particular target gene of
interest from primates to rodents. Combined with advances in techniques such as
conventional and conditional transgenics or knockouts, it is relatively straightforward
to create rodent models of human monogenic disorders through manipulating the
expression levels of those equivalent genes in rodents, while creating models of human
diseases of a multifactorial nature still remains a challenge. Nonetheless, these genetic
models can be combined with diet-/drug-induced regimens and/or surgery/xenograft
explants to further mimic the complex interaction of gene function and the environment,
create animal disease models that more closely parallel that of complex human diseases,
and allow more predictive evaluation on the effectiveness of newly developed drugs.

Besides demonstrating efficacies in animal disease models, effective drugs are
also required to have desirable characteristics in absorption, distribution, metabolism,
excretion, and pharmacokinetic (ADME/PK) profiles. Advances in techniques such as
genechip/microarray for studying gene expression or proteomics for protein–
protein interactions have presented novel testing paradigms in understanding the
mechanisms of diseases as well as evaluating the ADME/PK profiles of drugs in both
in vitro and in vivo rodent models. These data from ADME/PK studies of candidate
drugs are then utilized for better drug design through further medicinal chemistry
work and additional efficacy studies, forming an essential feed-forward loop in the
drug discovery process.

Although rodent models of human diseases have provided a broader spectrum
of human disease models to choose from for drug testing, the differences in primate
and rodent physiology—in particular, the regulation of metabolic pathways—have
generated some uncertainty in relying on only rodent models for drug development
decisions (Ktorza et al., 1997). Nonhuman primates are closer to humans in terms of
physiology; thus, an increasing number of nonhuman primate models are gaining
popularity among pharmaceutical companies for testing developmental track drugs
for their efficacy and safety profiles. In the following sections, we will attempt to
highlight the impact of recent advances in genomics, proteomics, and pharmacoge-
nomics on generating highly predictive animal models in both rodents and nonhuman
primates, with emphasis on how differences in physiology affect the outcomes of
studying metabolic diseases and decisions on drug discovery.

19.2 RODENT MODELS OF HUMAN DISEASES

Years of traditional breeding, crossing, and selecting desirable phenotypes from
different strains have yielded a number of mouse lines that show phenotypic
similarity to human diseases. Mapping and crossing of these lines reveal the genetic markers or loci that are linked to the phenotypes. For instance, the obese and insulin-resistant strains of \textit{ob/ob} and \textit{db/db} mice and \textit{fa/fa} mutations in Kolestsky, ZDF, and Zucker rats are shown to be defective in leptin and its receptor (Chen et al., 1996; Chua et al., 1996; Takaya et al., 1996; Tartaglia et al., 1995; Wu-Peng et al., 1997; Zhang et al., 1994). Leptin is an important adipocyte-derived cytokine or adipokine that regulates appetite and energy metabolism. Defects in this system—in either the adipokine itself or its receptor—leads to obese and insulin-resistant phenotypes. Although genetic defects in the leptin system are rare in humans, rodents provide a platform in which antiobesity or antidiabetic drugs can be tested for their effects on body weight, blood glucose, and insulin levels as well as sensitivity to glucose or insulin challenges. Other spontaneous mutants used as obesity and diabetes models include tubby mice, yellow \textit{A^{v/y}} mice with ectopic expression of the agouti gene, and growth-hormone-deficient dwarf rats (\textit{dw/dw}) (Kleyn et al., 1996; Clark et al., 1996; Yen et al., 1994). This type of traditional cloning and mapping of mutant genes associated with a particular phenotype, although useful, requires large-scale breeding and crossing to identify phenotype-associated loci and eventually the associated gene. Yet other, more direct testing of candidate target genes of human disease by reverse genetics is carried out through genetic manipulations in mice to either knock in or knock out those genes to study the associated disease phenotypes.

19.2.1 Genetic Deletion Models

Even though the technique of generating mouse embryonic stem cells had been established before the genomic era, the advances in sequencing and mapping the human and mouse genomes have greatly assisted the process of targeted mutagenesis—in particular, conditional deletion of target genes by the Cre-loxP recombinase system (Branda and Dymecki, 2004; Yu and Bradley, 2001). While conventional target deletions have been fruitful in yielding useful and relevant information on the function of genes (Table 19.1), due to the essential function of some of these genes in early development, many of these targeted deletions have resulted in embryonic lethal phenotypes, preventing the evaluation of their functions in adults and in a tissue-specific fashion. The utilization of the Cre-loxP recombinase system, combined with the knowledge gained from studying promoter and enhancer elements that govern the tissue and developmental stage-specific expression of genes, allows the design of targeting vectors for conditional expression of Cre recombinase that would be expressed in desirable tissues during specific developmental stages. Therefore, targeted homologous recombination can be tailored to excise the targeted gene of interest in a controlled manner, generating deletion mutants that yield more informative phenotypes.

An example of utilizing conditional knockout to clarify the function of a gene in adult animals due to embryonic homozygous lethality involves a well-documented drug discovery target in diabetes mellitus, peroxisome proliferator-activated receptor gamma (\textit{PPAR\gamma}) (Berger et al., 2005). \textit{PPAR\gamma} is a nuclear hormone receptor, a class of transcription factor in which activity is modulated by small molecular ligands (Lazar, 2005). Nuclear hormone receptors also include other well-established drug
discovery targets such as peroxisome proliferator-activated receptor alpha (PPARα) for dyslipidemia, estrogen receptor (ER) for osteoporosis, and glucocorticoid receptor (GR) for anti-inflammation (Staels, 2005). The activity of PPARγ is modulated by endogenous ligands such as prostaglandin J2 (PGJ2) and synthetic drugs such as rosiglitazone that belong to the thiazolindinedione (TZD) class of molecules for the treatment of type II diabetes (Forman et al., 1995). Since a point mutation in PPARγ leads to diabetes and hypertension in humans and TZD stimulates the activity of PPARγ and alleviates the symptoms of diabetes, one would simply expect that a deletion of the PPARγ gene would prime the animal for the development of diabetes (Deeb et al., 1998). A homozygous deletion of PPARγ results in embryonic lethality; however, a heterozygous deletion of PPARγ actually partially protects the animal from developing diabetes (Kubota et al., 1999; Miles et al., 2000).

PPARγ is expressed in multiple tissues, more abundantly in adipose than in muscle, the primary organ responsible for glucose disposal upon insulin stimulation. The paradoxical observation from heterozygous mice has led to the hypothesis that differential regulation of PPARγ activity in these insulin-responsive tissues is required to confer protection from developing diabetes. By selecting appropriate targeting vectors that would result in tissue-specific homozygous deletion of PPARγ, different strains of adipose- or muscle-specific knockouts have revealed the complex relationship of PPARγ and glucose metabolism (He et al., 2003; Hevener et al., 2003). Adipose-specific PPARγ knockout causes adipose hypertrophy, elevated levels of

| Table 19.1 Selected Examples of Gene Deletions Used in Obesity and Diabetes Research |
|-------------------------------|----------------|----------------|
| Gene                          | Phenotypes          | References                      |
| Pro-opiomelanocortin (Pomc1)  | Obese, IR<sup>a</sup> | Yaswen et al. (1999)             |
| Melanocortin-3 receptor (Mc3r) | Obese, IR          | Chen et al. (2000)               |
| Melanocortin-4 receptor (Mc4r) | Obese, IR          | Huszar et al. (1997)             |
| NPY1 receptor (Npy1r)          | Obese<sup>b</sup>   | Kushi et al. (1998)              |
| NPY2 receptor (Npy2r)          | Obese<sup>b</sup>   | Naveilhan et al. (1999)          |
| NPY5 receptor (Npy5r)          | Obese<sup>b</sup>   | Marsh et al. (1998)              |
| 5-Hydroxytryptamine 2C receptor (Htr2c) | Obese, IR | Nonogaki et al. (1998) |
| Bombesin receptor subtype-3 (Brs3) | Obese            | Okhi-Hamazaki et al. (1997)     |
| Peroxisome proliferator-activated receptor α (Ppara) | Obese          | Costet et al. (1998)             |
| Steroid receptor coactivator 1 (Src1) | Obese<sup>c</sup> | Picard et al. (2002)            |
| Glucocorticoid receptor interacting protein 1 (Grip1) | Lean<sup>d</sup> | Picard et al. (2002)            |
| Receptor interacting protein 140 (Rip140) | Lean<sup>d</sup> | Leonardsson et al. (2004)       |
| Estrogen receptor α (Esr1)     | Obese              | Heine et al. (2000)              |
| Estrogen-related receptor α (Esrra) | Lean<sup>d</sup> | Luo et al. (2003)               |

<sup>a</sup>IR, insulin resistance.  
<sup>b</sup>Late-onset obesity.  
<sup>c</sup>Obesity due to high-fat diet.  
<sup>d</sup>Resistant to high-fat-diet-induced obesity.
plasma-free fatty acids and triglyceride, and decreased levels of adipokines leptin and adiponectin, showing some of the characteristics of metabolic syndrome X. Despite causing insulin resistance in adipose and the liver, the blood glucose level, glucose and insulin tolerance, and insulin-stimulated muscle glucose uptake remain normal. On the other hand, muscle-specific PPARγ knockout causes serve muscle glucose intolerance and insulin resistance with an 80% reduction in insulin-stimulated glucose uptake. Importantly, the animals are not responsive to TZD treatment, strongly suggesting the primary action of TZD is through regulating the activity of PPARγ in the muscle, paving ways for medicinal modification of TZD derivatives for improved efficacies (Argmann et al., 2005).

19.2.2 Transgenic and Viral-Mediated Overexpression Models

While knockout studies have been useful in identifying the necessary functions of genes, transgenic models and viral-mediated overexpression are of equal value in building predictive animal models for drug development. By choosing the appropriate conventional or conditional transgenic vector, it is possible to link the elevated expression level of candidate genes associated with human diseases to the cause of the corresponding diseases in rodents (Table 19.2). Moreover, through viral-mediated overexpression in combination with conventional knockout, one can ameliorate the disease condition caused by gene knockout, providing additional proof-of-concept evidence.

Chronic and subacute inflammation is associated with obesity, insulin resistance, and diabetes (Bullo et al., 2003; Haffner, 2003; Shoelson et al., 2003). In diabetic patients, elevations of pro-inflammatory markers, particularly those under the control of inflammatory transcription activator NF-κB, are found in adipose and liver tissues. The activity of NF-κB is activated by the inhibitor IκB, which is in turn inhibited by its negative regulating kinases, IKKα and IKKβ. Cai et al. (2004)

**TABLE 19.2 Selected Examples of Transgenic and Adenovirus-Mediated Overexpression Used in Obesity and Diabetes Research**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Method</th>
<th>Phenotypes</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peroxisome proliferator-activated receptor δ (<em>Ppard</em>)</td>
<td>Muscle transgenic</td>
<td>ISa, Leanb</td>
<td>Wang et al. (2004)</td>
</tr>
<tr>
<td>Angiopoietin-related growth factor (<em>Angptl6</em>)</td>
<td>Adenovirus</td>
<td>IS, Leanb</td>
<td>Oike et al. (2005)</td>
</tr>
<tr>
<td>Retinol binding protein-4 (<em>Rbp4</em>)</td>
<td>Muscle transgenic</td>
<td>IRc</td>
<td>Yang et al. (2005)</td>
</tr>
<tr>
<td>Forkhead transcription factor C2 (<em>Foxc2</em>)</td>
<td>Adipose transgenic</td>
<td>Leanb</td>
<td>Kim et al. (2005)</td>
</tr>
<tr>
<td>Glucose transporter 4 (<em>Glut4</em>)</td>
<td>Adipose transgenic</td>
<td>IS</td>
<td>Shepherd et al. (1993)</td>
</tr>
<tr>
<td>Agouti-related transcript (<em>Agrt</em>)</td>
<td>Transgenic</td>
<td>Obese</td>
<td>Graham et al. (1997)</td>
</tr>
</tbody>
</table>

aIS, insulin-sensitive.
bResistant to high-fat-diet-induced obesity.
cIR, insulin-resistant.
generated a hepatic-specific transgenic model in which constitutively active IKKβ is overexpressed in the liver to induce a state of subacute chronic inflammation. These mice develop hyperglycemia with severe hepatic insulin resistance and moderate systemic insulin resistance, including effects in the muscle. Using a similar strategy, the same group of researchers tested the effect of muscle-specific expression of IKKβ and only observed profound muscle wasting resembling clinical cachexia. These tissue-specific transgenic studies provided strong evidence that chronic inflammation in the liver can cause insulin resistance and that anti-inflammatory drugs may be of therapeutic potential.

An alternative to generating transgenic lines is through adenovirus-mediated studies, although these studies show a limited range of overexpression. Adenovirus carrying the desirable gene of interest can be easily engineered by commercially available vectors and helper cell lines. Tail-vein injection of recombinant adenovirus typically leads to overexpression of genes in the liver. Previous work on conventional knockout of insulin signal negative regulator protein tyrosine phosphatase 1B (PTP1B) showed that these mice were hypersensitive to insulin (Elchebly et al., 1999). However, controversies arise as to the primary tissue of PTP1B negative action. By using adenovirus to re-express PTP1B in the liver, Haj and co-workers observed marked attenuation of insulin sensitivity in these animals, thus demonstrating that the liver is the primary site of action for the peripheral effect of PTP1B and suggesting that a liver-specific PTP1B inhibitor may be of clinical value in restoring insulin sensitivity in diabetic patients (Haj et al., 2005; Harley and Levens, 2003).

Adenovirus-mediated in vivo gene delivery is even more commonly used to investigate the functions of secreted proteins and constitutes an important aspect of the target validation process. Oike et al. (2005) demonstrated the potential therapeutic value of angiopoietin-related growth factor (ANGPTL6) as an antiobesity/antidiabetic agent through a combination of gene knockout, transgenic overexpression, and adenovirus-mediated gene delivery. Specifically, they found that ANGPTL6 knockout mice displayed an obese phenotype with insulin resistance. In contrast, ANGPTL6 transgenic mice were lean and resistant to diet-induced obesity. Importantly, when mice that had already gained weight on a high-fat diet were treated with an adenovirus driving the expression of ANGPTL6, they lost some of the weight gained and displayed improved insulin sensitivity, strongly suggesting that ANGPTL6 may be a novel therapeutic target in obesity/diabetes.

### 19.2.3 Diet- or Drug-Induced Models

Besides genetic manipulation by deletion or transgenic, an even broader range of animal models for testing the effectiveness of therapeutic compounds can be built through administration of special diet or drug regimens. For example, rodents can be induced to develop non-insulin-dependent diabetes mellitus (NIDDM) or type I diabetes through injection of streptozotocin (STZ), which works primarily through impairing the glucose-dependent insulin secretion from pancreatic beta islet cells (Tsujii et al., 1988). This model can be used to study the changes of a gene expression profile by microarray in beta islet cells or other insulin-responsive tissues such as the liver, adipose, and muscle for mechanistic studies or target discovery purposes as the
animal begins to lose its ability to secrete a sufficient amount of insulin to maintain euglycemia (Fan et al., 2003; Kume et al., 2005).

In addition to drug-induced models, more physiologically relevant studies through altering animal diet composition can also help gain important insights into mechanistic studies. Obesity is an important risk factor in humans for developing type II diabetes. Mice become overweight when put on a high-fat and/or high-carbohydrate diet and develop symptoms of metabolic diseases such as hyperinsulinemia and hyperlipidemia (Corsetti et al., 2000; Luo et al., 1998; Wang et al., 2001). This strategy is particularly informative when combined with genetic manipulation and drug-induced models.

Glucagon-like peptide-1 (GLP-1) is a gastric peptide hormone that functions to enhance the glucose-dependent insulin secretion from pancreatic beta islet cells (Stoffers, 2004). The half-life of GLP-1 is relatively short due to its cleavage by the dipeptidyl peptidase IV (DPP-IV) (Chen et al., 2003). Conarello et al. (2003) studied the effects of DPP-IV knockout on a high-fat diet. They found that DPP-IV knockout mice were refractory to weight gain and hyperinsulinemia compared to wild-type mice. With an elevated level of GLP-1, these mice were also protected against STZ-induced loss of beta cell mass and hyperglycemia, highlighting the potential of a GLP-1 analogue or DPP-IV inhibitor to be effective in enhancing glucose-dependent insulin secretion for managing diabetes (Nielsen, 2005).

Another intriguing example of combining a diet-induced regimen with that of genetic manipulation involves conditional knockout of IKK$\beta$ in either hepatocyte or myeloid cells to address the effect of modulating inflammatory pathway in different tissue would result in distinctive outcomes. Arkan et al. (2005) found out that mice with liver-specific knockout of IKK$\beta$ developed insulin resistance in muscle and fat in response to a high-fat diet despite retaining liver insulin responsiveness. On the other hand, myeloid-specific knockout mice remained insulin-sensitive and protected against insulin resistance on a high-fat diet. Together with the IKK$\beta$ transgenics mentioned above, these lines of evidence suggest that specific anti-inflammatory drugs may be of therapeutic value in treating metabolic diseases.

19.3 ANIMAL MODELS FOR PHARMACOLOGICAL AND TOXICOLOGICAL STUDIES

While these animal models can be utilized to establish the proof-of-concept studies and examine the efficacies of drugs under development, other important parameters such as ADME/PK and safety profiles that constitute the overall therapeutic value of drugs are being evaluated in pharmacological and toxicological studies (Theil et al., 2003). The advances in genomics, proteomics, and pharmacogenomics have brought forth improved in vitro assays and in silico predictions to learn about the absorption, distribution, metabolism, excretion, and pharmacokinetic profiles even before drugs are administered in animals for in vivo studies. These techniques if appropriately applied can help eliminate drug candidates that may pose problems early in the drug development process. However, to truly analyze the ADME/PK profiles, it is essential to use indicative animal models. Primarily, these pharmacodynamic/pharmacogenetic
studies are conducted using rodent models with preferred studies in larger animals, including canine and nonhuman primates.

19.3.1 Humanized Mice for Drug Metabolism Studies

Rodents such as mice and rats have been utilized as standards for learning the ADME/PK profiles of leading drug candidates for years. Nonetheless, due to the species difference between primates and rodents in xenobiotic sensors responsible for regulating the expression of drug-metabolizing enzymes, it is noteworthy to conduct at least some ADME/PK studies in nonhuman primates. Before engaging in these kinds of extensive and expensive studies, one can utilize a more humanlike mouse in terms of drug metabolism for such studies.

The majority of drug metabolism takes place in the liver where transcription factors including pregnane X receptor (PXR) and constitutive androstane receptor (CAR), both nuclear hormone receptors, regulate the expression of phase I cytochrome P450 oxygenases (CYPs), phase II UDP-glucuronosyltransferases (UGTs), and phase III transporters such as multidrug resistance-associated proteins (MRPs) and organic anion transporting polypeptide 2 (OATP2) for elimination and detoxification in response to xenobiotics or drugs (Bock and Kohle, 2004; Sonoda et al., 2003). The activities of these nuclear receptors are regulated by binding to structurally distinct xenobiotics through their respective ligand-binding domains.

Due to the differences in the primary sequence of the ligand-binding domain of PXR between humans and mice, drugs recognized by human PXR, such as rifampicin, are not necessarily recognized by mouse PXR; therefore, the predictive value of using only mice as a model system to learn about drug metabolism in humans is less than ideal (Jones et al., 2000). In addition, the enhancer elements that are recognized by PXR in controlling Cyp3A gene expression are only partially conserved between humans and mice, leading to differential action of Cyp3A genes in response to xenobiotics. In order to create a mouse model that recapitulates the specificity and regulatory pattern to that of a human, a line of PXR homozygous knockout was first created and then followed by replacing the deleted mouse gene with that of human PXR (Xie et al., 2000). These humanized mice can then recognize the plethora of drug molecules that would lead to undesirable activation of the Cyp3A gene. Thus, it is now possible to learn about the potential regulation of CYP genes that may be the basis for adverse drug–drug interactions.

19.3.2 Pharmacodynamic/Pharmacokinetic Studies

Besides understanding the regulation of metabolizing enzymes and transporters by potential drug candidates by humanized mice, it can also be used to learn about the pharmacodynamic and pharmacokinetic profiles. Depending on the optimal treatment regimen, a desirable drug may be long- or short-lived once taken. It is therefore important to determine the kinetics and distribution of drugs in circulation and at target tissues, analyze the metabolite derivatives and excretion rate of drugs, obtain the maximal tolerance dose for short-term safety evaluation, and establish the optimal dosing regimens to define the therapeutic window (Li, 2004).
Therefore, by first selecting the appropriate animal model for testing efficacy, pharmacodynamic parameters measured like the concentration that produces the half-maximal effect, the duration and intensity of responses can then be conducted.

Another key parameter for a desirable drug is bioavailability. Bioavailability is the amount of exposure measured as free concentration in plasma—the amount of drugs not bound to plasma protein. Bioavailability is often determined by the rates of absorption, extraction, elimination, and variability due to genetic and environmental factors. Bioavailability data including the maximum concentration \( (C_{\text{max}}) \), time to maximum concentration \( (t_{\text{max}}) \) in single or multiple-dose studies, area under curve (AUC), peak-trough fluctuation (PTF), and time required to reach steady state can sometimes be extended to calculate the potential safety risk where plasma free concentration is considered to be relevant to compound safety. In general, there is a 30-fold safety multiple between therapeutic activity and concentration causing QT prolongation, a predictive risk of ventricular tachycardia.

**19.4 NONHUMAN PRIMATE MODELS FOR METABOLIC DISEASES AND DRUG DISCOVERY**

Despite the progress made in creating more humanized rodents to produce more predictive disease models, the differences in physiology and metabolism in particular have raised concern about the relevancy of data obtained from some rodent models. Noticeably, the metabolic regulatory pathways governing the biosynthesis of cholesterol and bile acids in the liver are not entirely conserved between primates and rodents. In humans, one of the major risk factors for the development of atherosclerosis is an elevated circulating level of cholesterol that correlates with high dietary cholesterol and fat intake (Goldstein and Brown, 1990). However, mice and rats on a high-cholesterol diet are resistant to hypercholesterolemia (Horton et al., 1995).

Liver conversion of cholesterol into bile acids is through the rate-limiting step enzyme cytochrome P450 (CYP) 7A1 in which expression is governed by a network of nuclear hormone receptors in response to their respective ligands (Goodwin et al., 2000; Lu et al., 2000; Russell, 1999). CYP7A1 expression is feedback-controlled by the pathway endproduct bile acids through binding to a farnesoid X receptor (FXR). Together with heterodimer partner 9-cis retinoic acid receptor RXR, FXR activates its target gene small heterodimer partner (SHP), which in turn acts as a repressor of positive factor liver receptor homolog 1 (LRH1) on the expression of the CYP7A1 gene. In mice and rats, CYP7A1 is induced by oxysterol metabolites of cholesterol, which bind to and activate liver X receptor (LXR). LXR binds to the CYP7A1 gene through a response element that is not conserved in humans. Instead, activation of LXR represses human CYP7A1 expression through inducing the expression of SHP (Goodwin et al., 2002). Thus, there is no substitute for testing the efficacy and safety of metabolic disease drugs in nonhuman primate models, especially those that are designed for treating hypercholesterolemia.
Nonhuman primate models for metabolic diseases have been utilized successfully to provide more informative mechanism studies and predictive efficacy studies for cholesterol-lowering and antidiabetic drugs. Barbara Hansen et al. have established a colony of obese rhesus monkeys at the Obesity and Diabetes Research Center. These monkeys first develop fasting hyperinsulinemia, hyperglycemia, and insulin resistance and then eventually become diabetic with retinopathy and nephropathy upon aging, displaying many characteristics similar to that of the human condition (Ortmeyer et al., 2000). When these monkeys were treated with the antidiabetic drug TZD, the compromised glycogen synthase activity in diabetic muscle was restored close to the normal level, contributing to improved insulin sensitivity and indicating that these animals respond to antidiabetic drugs in a similar way to humans.

In addition to testing antidiabetic drugs, this model has also been shown to be valuable in demonstrating the effectiveness of a novel peroxisome proliferator-activated receptor δ (PPARδ) ligand. Oliver et al. first observed that the synthetic PPARδ ligand can stimulate the expression of LXR and some LXR-regulated genes in vitro and speculated that this PPARδ ligand may regulate cholesterol biosynthesis at least partly through LXR (Oliver et al., 2001). They tested this idea in the obese monkey model and observed a dose-dependent decrease in small-dense low density lipoprotein, fasting triglycerides, and insulin and an increase in high-density lipoprotein, dramatically improving the serum profile of these monkeys to a lower risk category for cardiovascular disease. These cases strongly indicate the predictive value of such a nonhuman primate model in assessing drug efficacy.

19.5 CONCLUSION AND FUTURE PROSPECTS

With the ingenious use of various genetic manipulations such as conditional knock-ins or knockouts with an appropriate genetic strain of rodents in combination with diet- or drug-induced regimens, there is a wide range of possibilities for testing the functions of genes. Once the therapeutic potential of a target has been established by these animal models, lead candidate compounds can also be tested in these models to study their on-target and off-target effects. A compound designed to modulate the activity of a particular target should lose its efficacy in the background of the appropriate target knockout. Residual activity may then suggest off-target effects, possibly through acting on similar family members of the intended target or through unanticipated mechanisms. To further evaluate if these types of compounds still possess therapeutic value, it is then desirable to test the compound in more sophisticated models. The future directions may thus depend on the generation of even more desirable conditional compound knockins and knockouts of multiple genes to tailor the timing and duration of genetic manipulations. One possibility is to combine the current techniques with the siRNA technology that would allow modification of a combination of genes as opposed to a single genetic knockout. Once these animal models have been established to reflect the pathological condition of human diseases, they become highly valued in the drug development process for learning the efficacy of developing compounds, potential off-target effects, and ADME/toxicity profiles.
REFERENCES


regulation of rat and human *CYP7A1* by the nuclear oxysterol receptor liver X receptor-α. *Mol


phosphatase 1B (PTP1B) re-expression alters glucose homeostasis of PTP1B-/-mice. *J Biol Chem* **280**(15), 15038–15046.


He, W., Barak, Y., Hevener, A., Olson, P., Liao, D., Le, J., Nelson, M., Ong, E., Olefsky, J. M., and


Hevener, A. L., He, W., Barak, Y., Le, J., Bandyopadhyay, G., Olson, P., Wilkes, J., Evans, R. M., and


Huszar, D., Lynch, C. A., Fairchild-Huntress, V., Dunmore, J. H., Fang, Q., Berkemeier, L. R., Gu, W.,

receptor: A promiscuous xenobiotic receptor that has diverged during evolution. *Mol Endocrinol* **14**(1),
27–39.

Liu, Z. X., Dong, J., Cline, G. W., Enerback, S., and Shulman, G. I. (2005). Adipocyte-specific over-
expression of FOXC2 prevents diet-induced increases in intramuscular fatty acyl CoA and insulin

Holmgren, L., Charlat, O., Woolf, E. A., Tayber, O., Brody, T., Shu, P., Hawkins, F., Kennedy, B., Baldini,
L., Ebeling, C., Alperin, G. D., Deeds, J., Lakey, N. D., Culpepper, J., Chen, H., Glucksman-Kuis,

animal models of diabetes relevant to the study of the genetics of non-insulin-dependent diabetes in

Kubota, N., Terauchi, Y., Miki, H., Tamemoto, H., Yamauchi, T., Komeda, K., Satoh, S., Nakano, R., Ishii,
C., Sugiyama, T., Eto, K., Tsubamoto, Y., Okuno, A., Murakami, K., Sekihara, H., Hasegawa, G., Naito,
M., Toyoshima, Y., Tanaka, S., Shiota, K., Kitamura, T., Fujita, T., Ezaki, O., Aizawa, S., Kadowaki, T.,


20.1 INTRODUCTION

While the presence of a pluripotent or multipotent “stem” cell in embryonic and adult animals is conceptually intuitive, a concerted study of these cells as a discipline has acquired prominence only since the early 1990s (Gage et al., 1995; Snyder and Macklis, 1995). It is now widely accepted that stem cells of varying potentials exist in adult organisms ranging from plants to invertebrate and vertebrate animals and are present to an even greater extent in developing and young organisms. Stem cells reside in several tissues in the adult and embryonic stages of mammalian development and have the potential to give rise to several mature lineages (Blau et al., 2001). This complex field is rapidly emerging largely due to the promise it holds for cellular therapies in humans. In this chapter we will attempt to review some salient aspects of the field in general and demonstrate research and therapeutic uses of stem cells with some specific examples.

*Corresponding author: esnyder@burnham.org

Drug Discovery Research: New Frontiers in the Post-Genomic Era, Edited by Ziwei Huang
Copyright © 2007 John Wiley & Sons, Inc.
Embryonic stem cells (ESCs) are isolated from the inner cell mass of a developing embryo, are pluripotent, and have the unbridled capacity to proliferate and differentiate into almost all the cell types that populate the adult organism. Conversely, somatic stem cells are present to varying degrees in young and adult tissues and are thought to have a more restricted potential in terms of the types of fates into which they can differentiate. The best-described of the somatic stem cells include neural stem cells (NSCs), which may be isolated from both adult and embryonic brains; hematopoietic stem cells (HSCs), which are isolated from adult bone marrow; and mesenchymal stem cells (MSCs), which are also isolated from adult bone marrow. Stem cells have also been described from adult adipose tissues (adipose-derived stem cells, ADSCs), umbilical cord blood, adult muscle, and liver, among other tissues. While somatic stem cells are more restricted in their differentiation potential compared to their embryonic counterparts, the plasticity of embryonic stem cells presents as a double-edged sword because they do not form tumors and/or differentiate into unexpected lineages as readily when used in transplantation therapies. Adult stem cells also have fewer political encumbrances in terms of limits on the generation and practical use of human stem cell lines.

While it is possible to harvest HSCs, MSCs, and ADSCs with relative ease from adult humans, ESCs and NSCs present more of a challenge due to size and anatomical considerations. Thus, cell lines have been created from mouse and human NSCs that have been created either by immortalization with genes such as myc or by propagation with selected growth factors. NSCs have been selected from fetal human tissue by a battery of surface antibodies, including CD133 and 5E12 (Uchida et al., 2000), and NSC lines have been derived by immortalization with myc (Villa et al., 2000) and selected by culture techniques and growth factors (Snyder, unpublished data). Similarly, several lines of ESCs have been created and characterized to varying extents (Hoffman and Carpenter, 2005). A registry of hESCs currently eligible for federally funded research is available on http://stemcells.nih.gov/research/registry/. There are currently 21 hESCs on this list, which have been derived in universities and institutes all over the globe. A concerted effort is currently being made to characterize these ES cells and establish culture conditions, including selection and identification criteria that may serve as a benchmark in global stem cell efforts. The results of the characterization of some of these cell lines are presented in the Web site http://stemcells.nih.gov/research/nihresearch/, where selected cell lines have been characterized and the results made public. Similar efforts are in progress in other labs where techniques such as several variations of microarrays (several variations), massively parallel signature sequencing (MPSS), and proteomics are being used in an effort to sift out the various cell populations that compose the complex hESC cultures. Some of these will be discussed in greater detail below. Other groups have also independently created hESC lines, including the 17 lines described in Cowan et al. (2004).

The obvious in vivo application for stem cells is transplantation of stem or progenitor cells that lead to “cell replacement therapies.” Ameliorative effects could potentially be the result of one or all of these events: (1) appropriate differentiation of the stem cells to mature cells that are functionally and/or phenotypically similar to the cell compromised or lost; (2) trophic/paracrine effects provided by the transplanted cells to the injured tissue; and (3) fusion of transplanted cells to the injured cells.
Although hESCs hold the most therapeutic value because of their pluripotency, our lack of knowledge regarding control of their differentiation and their propensity to promote tumors after transplantation is a major issue that needs to be addressed before they can be serious contenders in human therapeutics. In addition to transplantation, the homing properties of stem cells may prove advantageous in the use of these cells for the delivery of therapeutic agents to compromised tissues in the body, especially tumors (Aboody et al., 2000; Lee et al., 2003; Nakamizo et al., 2005). The latter use may be more tangible as the cells could be engineered in such a manner that they are dispensed with once the therapeutic is delivered.

While the most obvious translation of the potential of stem cells into realizable and immediate benefit to humans is in cell replacement therapies with the endpoint of short- and long-term amelioration of disease phenotypes, there are other equally important and interesting applications of these remarkable cells. Important in vitro uses of stem cells include the use of differentiating and mature differentiated phenotypes at several stages of the drug discovery process. In addition to these two applications, the dynamic multipotentiality of stem cells can also be used to establish in vitro paradigms that represent diseased states (particularly of developmental disorders) so that the mechanisms of onset and reversal of disease may be studied at the molecular and cellular levels. The insights thus obtained may be used to design therapies for these illnesses. In addition, at a purely basic level of scientific curiosity, the mechanisms of developmentally related events such as fate choice, cell proliferation, and cell motility may be studied in these dynamic cultures. An understanding of the mechanisms of development will drive the manipulation and exploitation of stem cells for novel cell-replacement therapies.

### 20.2 DEFINITION OF STEM CELLS

The enormous potential of stem cells arises from their plasticity. The use of stem cells in all pertinent applications will be greatly aided by a precise definition of their state in the continuum of “multipotent cell” to “committed precursor.” Cell surface markers, in particular, will greatly aid in the isolation of particular populations of stem cells when used in conjunction with magnetic bead technology and fluorescence-activated cell sorting (FACS); these populations of stem cells may be used in both in vivo and in vitro applications. Cytoplasmic markers could largely be used for the identification of cell types. At the present time there is a dearth of markers that precisely define stem cells and their progeny across the field. Although work in the field is progressing by the absence or presence of groups of markers thought to identify particular groups of cells, the lack of comprehensive sets of markers (or a single marker) defining specific stem cells, precursors, and mature phenotypes leads to a lack of precision, which is a serious limitation in the field.

Several techniques have been used in attempts to define unique markers for embryonic and somatic stem cells markers and the mature phenotypes that they yield. Some of the earliest publications attempting to define “stemness” genes—that is, genes that are unique to the stem cell state irrespective of the source of the stem cells—met with muted success (Ivanova et al., 2002; Ramalho-Santos et al., 2002).
These and other similar efforts suggest that the study of differential gene expression alone is not sufficient to define markers of particular stem cells or restricted lineages, because there doesn’t appear to be a clear demarcation of specific markers for different cell types at the level of gene expression. Other markers related to stem cells including Oct4, nanog, nestin, and musashi have been isolated using techniques such as invertebrate genetics, expression cloning, generation of antibodies to whole cells and cell lysates, differential display, and activation of proteins during certain cellular events to define markers of particular cell types or processes (Mitsui et al., 2003; Chambers et al., 2003; Lendahl et al., 1990; Sakakibara et al., 2002). Using subtractive cDNA technology, Brandenberger et al. (2004b) have compared differences in gene expression in a population of undifferentiated hESCs with three populations of cells differentiated from hESCs and have attempted a characterization of differences in levels of expression (not activation) of signaling proteins and related transcription factors. Such “transcriptome” analyses will possibly lead to the definition of cohorts of markers that define particular fates or may be useful for the characterization of populations of cells. The HSC and MSC fields have relied largely on the “CD” surface markers gleaned over time by defining markers of cells as they differentiate into various lineages in the blood (Shizuru et al., 2005; Pittenger and Martin, 2004). Markers such as CD133, which has been used to characterize HSCs as well as NSCs (Uchida et al., 2000; Shizuru et al., 2005), may be a general marker of stem cells. Stem cells have also been isolated based on their position in a fluorescence activated cell sorting (FACS) profile, where a small group of cells exclude dyes with greater efficiency than the rest of the population of cells within that particular tissue. These cells, called “side population” cells, have now been characterized from various tissues including blood, bone marrow, breast, skin, and eye and appear to coincide with the appearance of membrane channels of the ATP binding cassette (ABC) family that includes the MDR and BCRP1 receptors. The stemlike properties of this population are still undergoing characterization (Jonker et al., 2005; Camargo et al., 2005; Uchida et al., 2004). The ongoing characterization of ESCs and the various somatic stem cells will continue to broaden our definition of each kind of cell and perhaps improve our methods of assessing the use of these cells in clinical and research projects.

Other novel identifying events, perhaps more dynamic, are also being explored as markers of particular states of stem cells. Using gene-trapping technology, Scheel et al. (2005) studied differentiating neural stem cells and identified known and novel markers of differentiation. Such data will yield clues as to the dynamic process of differentiation in stem cells and perhaps yield markers, which define several stages of differentiation before the mature phenotype is reached. Since the final arbiter of fate choice is transcription of specific genes that dictate fate choice, of particular interest is the transcriptional state of the cell, the ability of the DNA to bind to proteins such as histones and transcription factors, and the ease with which the DNA can be transcribed—that is, the epigenetic state of the genome. Several clever variations of array technology are being developed where the methylation states of the DNA and histones and the availability of DNA to bind to transcription factors can be detected [reviewed in van Steensel (2005)]. The use of these techniques could possibly lead to novel insights into the “transcriptome” of proliferating and differentiating stem cells.
This includes the ChIP-chip (chromosome immunoprecipitation) that can analyze the sites in the genome that are bound by particular proteins. Alternatively, the presence of specific protein complexes in the cell could be considered a marker, for instance, if it relates specifically to the activation of a transcription factor required for a particular fate. In the example presented in Fig. 20.3, the complex of FRAP-Stat3 would serve as a marker of glial differentiation in similar paradigms, since Stat3 has been shown to be required for glial differentiation by at least two independent groups (Rajan et al., 2003; Bonni et al., 1997). However, Stat3 activation alone cannot serve as a marker for glial differentiation since the neuroepithelial cytokines activate STAT proteins in several tissues and under diverse conditions (Aaronson and Horvath, 2002). Such complexes could be detected by colocalization of signal by immunofluorescence or by fluorescence energy transfer (FRET) in combination with bead technology or FACS, depending on the application in question. Stable RNA molecules such as miRNAs have recently been shown to regulate cell cycling in stem cells (Hatfield et al., 2005) and regulate the differentiation of cardiomyocytes (Zhao et al., 2005).

More global efforts at signaling pathway analysis have also been recently attempted. In an example of trying to model the global signaling state of a hippocampal neuron, Ma’ayan et al. (2005) have used a combination of published literature along with computer modeling to propose networks of signals arising from the ligand–receptor interaction at the cell surface to transcription factor activation in the nucleus and have described regulatory motifs along the cellular network. While this study has taken only the chemical reactions into consideration without spatial and temporal criteria, it provides an initial look at a complex cellular network and could develop into more mature definitions of networks of plastic versus mature cells. Others have used microarray approaches to globally detect genes that are inactive and active in hESCs with respect to particular transcription factors and postulate possible feed-forward and feedback loops at the gene expression level (Boyer et al., 2005). It is very possible that a precise description of a stem cell as it traverses its path to differentiation may be defined by combinations of two or more kinds of markers, “static” and “dynamic,” some of which are described above. Given the relatively nebulous nature of the markers presently being used to define stem cells and their derivatives, the only stringent definition of a stem cell and progenitor remains a functional one. While it is unrealistic, or sometimes impossible, to perform “clonal assays” to prove that the cell in question is a stem/progenitor cell because it is able to give rise to all the requisite fates, there is no equivalent alternative at the present time. The evolution of markers precise enough to identify the position of the cell in the continuum of “stemness” through “mature fate” will greatly enhance the efficiency of the experimental and therapeutic avenues being pursued (Table 20.1).

While several approaches are being used to define markers as described above, existing knowledge is being used to characterize cultures currently being used in laboratories. This is a particular issue for hESCs, which, due to their enormous plasticity, are notoriously heterogeneous due to variations introduced by passaging of cultures, minor variations in culture protocols that exist in different laboratories. Various techniques characterizing the expression profiles of different stem cells have been used. Expression microarrays have been used to determine the similarities and differences between hESCs (Bhattacharya et al., 2004; Yang et al., 2005). Description
of the gene expression profiles of the ESC lines HES3 and HES4 has been done by serial analysis of gene expression (SAGE) to compare between these cell lines and also to compare them with mouse ESC lines and cancer cells (Richards et al., 2004). Massively parallel signature sequencing (MPSS) has also been used to confirm some of the data generated by other techniques (Brandenberger et al., 2004a).

MSCs and HSCs have been characterized largely using FACS to determine the surface markers present on the cells. NSCs have also been characterized using FACS, again by the description of their surface markers (Uchida et al., 2000).

A conundrum in the stem cell field is the fact that due to the limited numbers of stem cells present in an embryo or adult animal, stem cells have to be expanded in culture to generate enough numbers of cells for experimental and clinical manipulations. However, due to the inherent plasticity of the cells, the stem cell cultures that result are entirely a product of the culture conditions. This must always be considered while making interpretations about the plasticity, or lack thereof, of stem cells.

### TABLE 20.1  Current Markers Used to Identify ESCs, NSCs, and MSCs

<table>
<thead>
<tr>
<th>ESCs</th>
<th>NSCs</th>
<th>MSCs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nanog</td>
<td>Nestin</td>
<td>CD 44</td>
</tr>
<tr>
<td>Sox 2</td>
<td>Sox 1</td>
<td>CD 105</td>
</tr>
<tr>
<td>Oct 4</td>
<td>Musashi</td>
<td>CD 71</td>
</tr>
<tr>
<td>Rex 1</td>
<td>CD 133</td>
<td>CD 29</td>
</tr>
<tr>
<td>CD 9</td>
<td>PSA-NCAM</td>
<td>CD 45 negative</td>
</tr>
<tr>
<td>SSEA3</td>
<td>5E12</td>
<td>CD 73</td>
</tr>
<tr>
<td>SSEA 4</td>
<td></td>
<td>CD 90</td>
</tr>
<tr>
<td>Telomerase</td>
<td></td>
<td>p75</td>
</tr>
<tr>
<td>Alkaline phosphatase</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*a* This is a parsimonious list of markers for some human stem cells. While more extensive lists are in use, the most commonly used markers for these three varieties of human stem cells are listed. Because no single marker is unambiguously specific (i.e., there may be overlap in immunoreactivity between different stages of differentiation or even between cells from different organ systems), designation of a given stem cell type should rest only the use of panels of multiple markers (not simply one) complemented by functional assays.
particularly relating to somatic stem cells. An excellent example of this is the culture of mesenchymal stem cells. When cultured in medium containing 10% serum as is common in the field, these cells differentiate into adipocytes, chondrocytes, and osteoblasts (Pittenger et al., 1999). However, when cultured under alternate culture conditions, including low serum and regulated density, these cells can also be coaxed into ectodermal fates, such as neurons, and endodermal fates, such as endothelium (Jiang et al., 2002; Gregory et al., 2005; Rajan and Mackay, unpublished data). Thus, the differentiation potential of stem cells is performed on stem cells that have undergone proliferation in specific media, and it is possible that the conditions used for, and thus the signals received during, proliferation of the stem cells dictate the restrictions of the differentiation potential in the resulting culture. Similarly, neural stem cells cultured in serum and/or as neurospheres have a greater propensity to glial fates when compared to neural stem cells maintained as monolayers and in defined culture media—that is, without serum (Fig. 20.1). A corollary to this is that

**Figure 20.1** Sixty-seven percent of NSC genes are lost in neurospheres. This figure, taken from Parker et al. (2005), shows that there are substantial qualitative differences in NSCs that are grown in different conditions. Here comparisons were made between cells grown as neurospheres and monolayers, and differences are seen in genes involved in cell signaling, fate choice, and cell migration.

**Key**
P = present expression  
M = marginal expression  
A = absent expression  
E = expressed dChip call  
- = absent dChip call
the culture of stem cells isolated from various tissues such as bone marrow, adipose tissue, and umbilical cord under the same culture conditions, such as that used for the culture of MSCs, will possibly result in the isolation and culture of a similar stem cell from these seemingly unrelated tissues (Strem et al., 2005; Li et al., 2005a). Such problems are only magnified when dealing with embryonic stem cells, which are even more plastic and less defined. The creation of disparate ESC lines from the same starting cell line due to variations in culture conditions extant in independent laboratories is an issue that is being addressed in a concerted manner by the International Stem Cell Initiative (Andrews et al., 2005). Another caveat in the field is that most of the experiments defining plasticity of particular stem cells are performed \textit{in vitro} with cultured stem cells. Although the differentiation potential of some stem cells such as HSCs and ESCs have been confirmed \textit{in vivo}, the limits of multipotentiality of stem cells \textit{in vivo}, and especially \textit{in situ}, remain to be defined. All these observations serve to reiterate the continuing need for monitoring the populations of embryonic and somatic stem cells that are in culture, and possibly \textit{in vivo}, by markers that need to be more comprehensive and stringent in defining the cell types being used or derived.

The culture and expansion of stem cells is evolving into an independent discipline, both at the level of small-scale cultures for research and on a large scale for potential clinical applications. The challenge of stem cell culture is to achieve the dual goals of maintaining the cells in an undifferentiated state while maximizing for yield. More often than not, dense cultures tend to differentiate spontaneously. Three practical considerations impinge on the design of a culture method. (1) A recent comprehensive review by Hoffman and Carpenter (2005) gives an idea of the intricacies of hESC culture. Traditionally, hESCs have been maintained on mouse fibroblasts that serve as “feeder” layers and promote maintenance of the undifferentiated state. MSCs are routinely cultured in serum. Due to the implications of nonhuman pathogens associated with cells to be used in transplantation, the culture conditions for all stem cells destined to be used for human therapies will preferably be “animal origin free”—that is, free of all undefined reagents of animal origin, including serum, usually fetal bovine serum (FBS), and feeder cells. (2) Although FBS is a rich source of growth, survival, and cell adhesion factors, it is essentially an undefined cocktail of proteins containing large amounts of albumin. Again, culture of stem cells in defined medium will permit better control of the culture and a more reproducible maintenance of the undifferentiated and differentiated states. (3) However, the creation of appropriate culture conditions by the addition of the requisite individual growth factors that promote proliferation and survival of stem, differentiating, and differentiated cultures, along with the appropriate adhesion molecules, can be prohibitively expensive. This is particularly true when large-scale production of cells needs to be achieved for patient and drug screening applications. Another practical consideration is that suspension cultures are preferred over adherent cultures for the large-scale production of cells, because they are more feasible in terms of expenditure of reagents and manpower. Several stem cell cultures, including MSCs, NSCs, and ESCs, are best maintained in adherent cultures. Creative methods are being devised in order to adapt these cells to more suspension-like conditions.
culture methods, including the use of beads as “suspended substrata” on which cells may grow (Schroeder et al., 2005; Sen et al., 2004).

20.3 USES OF STEM CELLS IN THERAPEUTIC APPLICATIONS

The most obvious use of stem cells is in replacement therapies where compromised tissues are replaced with functional equivalents. This had been envisioned as the goal in several disease processes, most commonly in diseases where a specific known cell type is compromised, such as diabetes, spinal cord injury, myocardial infarctions, Parkinson’s disease (PD), and some blood cancers reviewed in (Daley et al., 2003; Shizuru et al., 2005). While initial impetus for the field of therapeutic transplantation was given by bone marrow transplantations for blood-related malignancies (Shizuru et al., 2005), it is now being contemplated for several of the somatic stem cells—including NSCs and MSCs and, in the longer term, ESCs as well—when the processes of uncontrolled differentiation and tumorigenesis in these latter cells is understood better. Amelioration of disease phenotypes in animal models have been attempted with various stem cells including hESCs, NSCs, HSCs, and MSCs (Kim et al., 2002; Yang et al., 2002; Eglitis and Mezey, 1997; Zhao et al., 2002). We will discuss the use of NSCs in transplantation with specific reference to PD.

20.3.1 NSCs: Parkinson’s Disease

NSCs have been used with varying degrees of success in transplantation therapies in the central nervous system (CNS) (Snyder et al., 2004; Picard-Riera et al., 2004). The impact of NSCs in directly rescuing endangered host neurons was first evinced in a series of experiments in aged rodents in which the nigrostriatal system was impaired (Fig. 20.2). PD is a degenerative disorder characterized by a loss of midbrain dopamine (DA) neurons with a subsequent reduction in striatal DA (Rosenthal, 1998). The disease, in addition to incapacitating many thousands of patients, has also long served as a model for testing neural cell replacement strategies. Transplantation therapy for this CNS disorder has a long history [for review see (Dunnett (1999)]. It was the neural disease that was first treated clinically by neural transplantation, using primary tissue from human fetal ventral mesencephalon to replace DA-expressing cells (Lindvall et al., 1990ab). Indeed, it was in this disease that the limitations of fetal tissue grafts in not only rodent and primate models of PD (Mehta et al., 1998), but also in clinical trials (Lindvall et al., 1990a,b), was first recognized. These limitations include (a) short graft survival and limited integration of the grafts and (b) the possibility of unregulated DA production in improper regions, leading adversely to dyskinesias. Given PD’s storied role in the development of cellular therapies, it is appropriate that a model of this disease should have also played a pivotal role in revealing a little-suspected but powerful therapeutic action that NSCs may play in preserving degenerating host cells by some heretofore-unheralded mechanisms that are nevertheless inherent to stem cell biology.
Figure 20.2 NSCs possess an inherent mechanism for rescuing dysfunctional neurons. Shown here is evidence from the effects of NSCs in the restoration of mesencephalic dopaminergic function. [Modified from Ourednik et al. (2002)].

I. TH expression in mesencephalon and striatum of aged mice following MPTP lesioning and unilateral NSC engraftment into the substantia nigra/ventral tegmental area (SN/VTA). A model that emulates the slow dysfunction of aging dopaminergic neurons in substantia nigra (SN) was generated by giving aged mice repeated high doses of MPTP. Schematic on top indicates the levels of the analyzed transverse sections along the rostrocaudal axis of the mouse brain. Representative coronal sections through the striatum are presented in the left column (A, C, E, G) and through the SN/VTA area in the right column (B, D, F, H). A and B. Immunodetection of TH (black cells) shows the normal distribution of DA-producing TH⁺ neurons in coronal sections in the intact SN/VTA (B) and their projections to the striatum (A). C and D. Within 1 week, MPTP treatment caused extensive and permanent bilateral loss of TH immunoreactivity in both the mesostriatal nuclei (C) and the striatum (D), which lasted for a lifetime. Shown in this example, and matching the time point in g and h, is the situation in a mock-grafted animal 4 weeks after MPTP treatment. E and F. Unilateral (right side) stereotactic injection of NSCs into the nigra is associated, within 1 week after grafting, with substantial recovery of TH synthesis within the ipsilateral DA nuclei (F) and their ipsilateral striatal projections (E). By 3 weeks posttransplant, however (G, H), the asymmetric distribution of TH expression disappeared, giving rise to TH immunoreactivity in the midbrain (H) and striatum (G) of both hemispheres that approached the immunoreactivity of intact controls (A, B) and gave the appearance of mesostriatal restoration. Similar observations were made when NSCs were injected 4 weeks after MPTP treatment (not shown). Bars: 2 mm (left), 1 mm (right). Note the ectopically placed TH⁺ cells in H. These are analyzed in greater detail, along with the entire SN, in II. II. Immunohistochemical analyses of TH, DAT, and BrdU-positive cells in MPTP-treated and grafted mouse brains. The presumption was initially that the NSCs had replaced the
dysfunctional TH neurons. However, examination of the reconstituted SN with dual βgal (green) and TH (red) ICC showed that 90% of the TH\(^{+}\) cells in the SN were host-derived [a c], and only 10% were donor-derived [d]. Most NSC-derived TH\(^{+}\) cells were actually just above the SN ectopically (blocked area in a, enlarged in b). These photomicrographs were taken from immunostained brain sections from aged mice exposed to MPTP, transplanted 1 week later with NSCs, and sacrificed after 3 weeks. The following combinations of markers were evaluated: TH (red) with βgal (green) [a–d]; NeuN (red) with βgal (green) [e]; GFAP (red) with βgal (green) [f]; CNPase (green) with βgal (red) [g]; TH (brown) and BrdU (black) [k]; GFAP (brown) with BrdU (black) [l]; CNPase (brown) with BrdU (black) [m]. Anti-DAT-stained areas are revealed in green in the SN of intact [h], mock-grafted [i], and NSC-grafted [j] brains. Three different fluorescence filters specific for Alexa Fluor 488 (green), Texas Red (red), and a double-filter for both types of fluorochromes (yellow) were used to visualize specific antibody binding: c, d, and h–j are single-filter exposures; a, b, and e–g are double-filter exposures. a shows a low-power overview of the SN/VTA of both hemispheres. The majority of TH\(^{+}\) cells (red cells in a) within the nigra are actually of host origin (~90%), with a much smaller proportion being donor-derived (green cells, ~10%) (representative close-up of such a donor-derived TH\(^{+}\) cell in d). Although a significant proportion of NSCs did differentiate into TH\(^{+}\) neurons, many of these actually resided ectopically, dorsal to the SN (boxed area in a, enlarged in b; high-power view of donor-derived (green) cell that was also TH\(^{+}\) (red) in c), where the ratio of donor-to-host cells was inverted: ~90% donor-derived compared with ~10% host-derived. Note the almost complete absence of a green βgal-specific signal in the SN\(^{+}\)VTA, whereas ectopically, many of the TH\(^{+}\) cells were double-labeled and thus NSC-derived (appearing yellow–orange in higher power under a red/green double-filter in panel b). c–g. NSC-derived non-TH neurons (NeuN\(^{+}\)) [e, arrow], astrocytes (GFAP\(^{+}\)) [f], and oligodendrocytes (CNPase\(^{+}\)) [g, arrow] were also seen, both within the mesencephalic nuclei and dorsal to them. h–j. Any proliferative BrdU\(^{+}\) cells after MPTP insult and/or grafting were confined to glial cells, whereas the TH\(^{+}\) neurons [k] were BrdU\(^{-}\). This finding suggested that the reappearance of TH\(^{+}\) host cells was not the result of neurogenesis but rather the recovery of extant host TH\(^{+}\) neurons. Bars: ~10 μm [a]; 20 μm [c, d, e]; 30 μm [f]; 10 μm [g]; 20 μm [h–j]; 25 μm, [k]; 10 μm, [l]; and 20 μm [m]. k–m. The green DAT-specific signal in j suggests that the reconstituted mesencephalic nuclei in the NSC-grafted mice (as in H) were functional DA neurons comparable to those seen in intact nuclei [h] but not in MPTP-lesioned sham-engrafted controls [i]. This further suggests that the TH\(^{+}\) mesostriatal DA neurons affected by MPTP are, indeed, functionally impaired. (Note that sham-grafted animals [i] contain only punctate residual DAT staining within their dysfunctional fibers, whereas DAT staining in normal [h] and, similarly, in engrafted [j] animals was normally and robustly distributed both within processes and throughout their cell bodies.) See color plates.
In addition to “replacing” the injured tissue, neural stem cells also appear to engage the host in a dynamic series of ongoing reciprocal interactions, each instructing the other. Under instruction from exogenous NSCs, the injured host nervous system also contributes to its own repair. The impact of NSCs in directly rescuing endangered host neurons was first evinced in a series of experiments in aged rodents in which the nigrostriatal system was impaired (Fig. 20.2). In the hope that NSCs might spontaneously differentiate into DA neurons when implanted into a DA-depleted region of the CNS, unmanipulated murine NSCs were implanted unilaterally into the substantia nigra of aged mice that had been exposed systemically to high-dose MPTP (1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine), a neurotoxin that produces a persistent impairment of mesencephalic DA neurons and their striatal projections (Ourednik et al., 2002). The NSCs not only migrated from their point of implantation and integrated extensively within both hemispheres, but also were, indeed, associated with a dramatic reconstitution of DA function throughout the mesostriatal system. While there was spontaneous conversion of a subpopulation of donor NSCs into dopaminergic neurons in DA-depleted areas, contributing to nigral reconstitution, the majority (~80–90%) of dopaminergic neurons in the “reconstituted midbrain” were actually host cells that had been “rescued” by factors produced constitutively by the NSCs with which they were juxtaposed and that had not, themselves, become neurons. These “chaperone cells” constitutively produced substantial amounts of neurosupportive agents. One such prominent molecule among many in this case was glial-cell-line-derived neurotrophic factors (GDNF), a factor known to be neuroprotective of ventrally located neurons (including DA neurons and spinal ventral horn cells). A similar observation is beginning to emerge from the implantation of human NSCs into the MPTP-lesioned subhuman primate model of PD.

A sense for the extent of the cross-talk also became evident when examining rodent models of hypoxia–ischemia (HI), a common cause of neurological disability in adults and children. HI causes much of its damage from extensive loss of cerebral parenchyma and of the cells and connections that reside there. When NSCs are implanted into these regions of extensive degeneration (particularly when transiently supported by biodegradable scaffolds), robust reciprocal interactions ensue spontaneously between the exogenous implant and the injured host brain that result in substantial reconstitution of parenchyma and anatomical connections as well as reduction of parenchymal loss, secondary cell loss, inflammation, and scarring. Similar results are observed in the hemi-resected adult rodent spinal cord in which evidence of an upregulated host neuronal regenerative response is noted, resulting in significant functional improvement. Indeed, the ability of engrafted NSCs to exert a protective and regenerative influence on degenerating host neural systems by virtue of their intrinsic expression of trophic factors is being observed in an increasing number of conditions. For example, the implantation of murine and human NSCs into the spinal cords of the SOD1 transgenic mouse model of ALS, a disease characterized by virulent motor neuron degeneration, has been pivotal in protecting these ventral horn cells from death and in preserving motor and respiratory function, blunting disease progression, and extending life. NSCs can similarly protect other neuron pools, promote motor axonal outgrowth following traumatic spinal cord injury, preserve
infarcted regions of cerebrum, induce vascularization of reconstituted regions of cortical parenchyma, and inhibit inflammation and scarring following traumatic or ischemic insult.

20.3.2 MSCs: Cardiac Illness

MSCs are relatively robust stem cells that are isolated from bone marrow, where they are present at a frequency of about 0.001% nucleated cells and of about a tenth as much as HSCs. While MSCs are usually cultured in FBS and resemble fibroblast cultures in appearance, unlike fibroblasts they have the capacity to differentiate into fates in the mesodermal lineage, including chondrocytes, adipocytes, and osteoblasts (Pittenger et al., 1999). Two properties of these cells make them attractive candidates for cellular therapies: their apparent lack of immunogenecity after transplantation and their potential to help in remodeling cardiac tissue after myocardial infarction (MI). While allogeneic MSCs have been used in animal models (rodent and porcine) with good functional and histological recovery, autologous bone marrow (presumably including several populations of cells in addition to MSCs) has also been used in the clinic with promising results (Silva et al., 2004). While MSCs have been injected into the site of cardiac ischemia resulting from MI, they have also been administered systemically and have been shown to home to the site of injury (Kraitchman et al., 2005). Subsequent histological examination of the site of transplantation indicates some incorporation and survival of the transplanted cells. However, the extent of recovery of the heart tissue is not concomitant with the numbers of surviving cells, suggesting that there is remodeling of the host cells and the extracellular matrix indicating possible paracrine effects originating from the transplanted cells. This is similar to the trend seen with the NSC transplantations. MSCs also appear to block T-cell responses in the host immune system (Aggarwal and Pittenger, 2005), thus leading to their possible use as adjunct therapy in transplantation of other tissues, including bone marrow. In addition, MSCs have been used with some success in therapies related to bone replacement (Caplan, 2005). Similar work is in progress with ADSCs, which also appear to ameliorate symptoms of MI and have the added advantage of ease of isolation of populations of interest from adipose tissue for autologous transplantation into the same patient.

20.3.3 Stem Cells as Anti-Cancer Agents

Tangential to the above discussion, NSCs and MSCs have been used to deliver therapeutics to tumors, particularly gliomas (Aboody et al., 2000; Lee et al., 2003; Nakamizo et al., 2005). Aboody et al. (2000) first reported that modified NSCs introduced into the parenchyma or cerebral ventricle as well as the systemic circulation could migrate over great distances to sites of intracranial pathology, as modeled by a glioma in rodent hosts and could position themselves in direct juxtaposition to glioma cells migrating away from the tumor bulk to invade normal tissue. This ability to track invading tumor cells signified a potentially powerful way to treat a phenomenon notorious to primary gliomas that has made their management so vexing. The same group observed reduction of tumor bulk and improved host survival with the use
of genetically modified NSCs. Combinations of NSCs from different sources have been shown to exhibit the same gliomatropic effect in experimental rodent brain tumor models and to effectively reduce tumor bulk and prolong host survival (Ehtesham et al., 2002, 2004; Barresi et al., 2003). Factors released and expressed by the glioma cells themselves, by the tumor stroma (composed of adjacent reactive astrocytes, microglia, and oligodendrocytes) and by tumor-derived endothelium as well as by the damaged surrounding normal brain itself all contribute to NSC gliompatibility. Some of these agents, such as stem cell factor (SCF) and monocyte chemoattractant protein-1 (MCP1), have been identified (Erlandsson et al., 2004; Widera et al., 2004; Sun et al., 2004), yet others are still to be characterized and their role in NSC gliompatibility defined (Werbowetski et al., 2004). Expression of SDF-1α by tumor-derived endothelium serves to attract the migration of NSCs (Allport et al., 2004; Fears et al., 2004). Blocking SDF-1α/CXCR4 interactions also prevents gliomatropic migration of NSCs (Ehtesham et al., 2004). SDF-1α/CXCR4 interactions appear to be pivotal as well to the gliompatibility of circulating adult hematopoietic progenitor cells (Tabatabai et al., 2005).

Exploiting the unique tropism of NSCs for gliomas, several groups have now confirmed the therapeutic efficacy of using genetically armed NSCs to target neoplasms in vivo in a variety of murine brain tumor models through the delivery of a variety of growth-regulating and anti-glioma gene products. Aboody et al. (2000) demonstrated in vivo efficacy of murine NSCs transduced with the gene for the enzyme cytosine deaminase (CDA). The enzyme converts the nontoxic pro-drug 5-fluorocytosine (5-FC) into the nucleoside analogue 5-fluorouracil (5-FU), which is then incorporated into the DNA of the neoplastic cell, causing chain termination and cell death. Tumor-bearing mice inoculated with CDA-expressing NSCs and given 5-FC demonstrated dramatic reduction of the intracranial tumor burden. This finding was subsequently corroborated in a different murine tumor model using a different NSC line (Barresi et al., 2003). Another approach utilizes NSCs as engraftable, mobile, gliomatropic viral packaging lines (Lynch et al., 1999). One study reported effective killing of tumor and escaping micro-deposits in a murine host by using murine NSCs to release replication-conditional HSV TK (Herrlinger et al., 2000), hence overcoming the typical low transduction frequency encountered in HSV glioma gene therapy by directing delivery of the virus to the intended cellular targets. This targeted cytotoxic effect, mediated by conversion of the pro-drug ganciclovir by TK into ganciclovir phosphate, is greatly amplified by virtue of the “bystander effect” (Li et al., 2005b). Gliomatropic NSCs engineered to be a viral packaging cell line for adeno-viral-based vectors also show similar efficacy (Arnhold et al., 2003). NSCs have also been used to deliver gene products that show therapeutic effects against gliomas such as IL4, IL12, and TRAIL (Benedetti et al., 2000; Ehtesham et al., 2002; Walczak et al., 1999) and a potent anti-angiogenic compound called endostatin (Bjerkvig et al., 2003).

It should be noted that although the tropism of stem cells for cancer was first unveiled by observing the behavior of NSCs, this phenomenon has also been observed in MSCs. MSCs engineered to express interferon-β (IFN-β) when injected into the carotid artery of brain-tumor-bearing mice also appeared to migrate in a gliomatropic fashion and destroyed the tumor via direct cytotoxicity, resulting in prolonged...
survival of the hosts (Nakamizo et al., 2005). Others have shown that MSCs over-expressing the immunomodulatory cytokine IL-2 migrate to the contralateral tumor-bearing hemisphere via the corpus callosum, helping promote tumor destruction (Nakamura et al., 2004). Lee et al. (2003) also suggested that cells derived from MSCs, which have some characteristics of NSCs, have the capacity to migrate toward an injury or glioma in the brain. Furthermore, MSCs administered systemically appeared to localize to prostate and breast cancers metastatic within the periphery. The question of which type of stem cell is best suited for which type of tumor within which region will need to be determined empirically. At present, we favor the view that stem cells derived from the lineage-of-origin of the cancer are best suited for “hunting it down” and eradicating it.

A brief spate of papers introduced the possibility of somatic cells differentiating into lineages removed from that of their origin—for instance, NSCs differentiating into skeletal muscle and hematopoietic lineages (Bjornson et al., 1999; Galli et al., 2000). Similarly, MSCs differentiate into mesodermal fates such as chondrocytes, adipocytes, and osteocytes as expected (Pittenger et al., 1999), but can also be induced to differentiate into ectodermal fates such as neurons and smooth muscle and endodermal fates such as liver. Whole bone marrow and MSCs have been shown to remain viable when transplanted into the CNS in models of spinal cord and ischemic injury and are thought to provide beneficial effects (Kopen et al., 1999; Chopp and Li, 2002; Akiyama et al., 2002; Hofstetter et al., 2002). While this “transdifferentiation” is an intriguing biological question in its own right, it may also provide alternative therapeutic options, especially in the case of somatic stem cells that are challenging to propagate, such as HSCs, or where the tissue in which they reside is relatively inaccessible, such as NSCs in the brain. However, it is also possible that the starting population of stem cells may not be as homogeneous as originally thought—that is, that the MSCs had some neural progenitors that were carried over from the bone marrow, or that some of the cells scored positive for a particular mature phenotype were scored aberrantly due to experimental artifacts (Daley et al., 2003). In addition, the occurrence of cell fusion of stem cells with other mature cells in the vicinity in vitro (Ying et al., 2002; Terada et al., 2002) and in vivo (Weimann et al., 2003a,b) introduces the possibility that some of these transdifferentiation events could be a result of cell fusion. In spite of all these caveats, the possibility of transdifferentiation of somatic cells merits further scrutiny because it has not been conclusively disproved.

In summary, the science of transplantation has made significant progress in the past decade, but several of the details remain to be investigated and conditions standardized. While transplantation of adult stem cells appears to be a more mature discipline than that of ESCs at the present time, mechanistic details of related phenomena will greatly enhance the efficacy of the treatment. These phenomena include incorporation and survival of transplanted cells in the compromised tissue, homing of transplanted cells to injured tissue, differentiation of stem cells to pertinent host tissue if any, paracrine/trophic effects of transplanted cells on host tissue (with or without differentiation), fraction of transplanted cells that are retained in the tissue of interest and location, and survival of remaining cells. The method of delivery of the cells may be determined somewhat by the disease being treated, but introduction of
cells may be at the site of the injury or systemic. In either scenario, the cells could be delivered with several refinements that enhance survival and perhaps differentiation/incorporation after transplantation, such as addition of growth factors that promote survival, genetic manipulation of the transplanted cells that may allow prolonged expression of genes of interest and may permit control of the transplanted cells in the patient, and, particularly in the case of ESCs, transplantation of cells that have been partially differentiated into appropriate lineages. Perhaps the site of transplantation may be pretreated to be favorable to receive and enhance the efficiency of the graft. Finally, knowledge of the precise mechanism(s) by which symptoms of the illness are reversed—that is, differentiation of transplanted cells and integration of host and transplanted tissues, trophic and paracrine functions of the transplanted cells, and cell fusion—will aid in the design of treatment protocols and in the possible use of combination therapies that include transplantation.

### 20.4 Uses of Stem Cells in Research Applications

#### 20.4.1 Experimental Systems to Study Basic Biology

Since the functional definition of a stem cell lies in its capacity to differentiate into several fates, the obvious follow-up to that question is the determination of the mechanism of these differentiation events. Stem cells provide a tractable system for a study of dynamic signaling processes that occur during fate choice and other events important in mammalian development, such as cell proliferation, differentiation, cell motility, cell survival, senescence, cell death, and formation of functional three-dimensional tissues. A wealth of literature exists describing (a) some of the molecular mechanisms of the complex differentiation phenotypes presented by stem cells and (b) the use of these cells to model three-dimensional cultures that resemble tissues and to model disease states so that the mechanism of diseased phenotypes may be studied. Three-dimensional tissues reminiscent of liver, capillaries, and blood have been described where stem or progenitor cells were cultured in three-dimensional matrices, including those composed of collagen (Imamura et al., 2004), matrigel or scaffolds composed of lactic acid/glycolic acid, and self-assembling peptide matrices (Semino et al., 2003). Ishikawa et al. (1996) studied hESCs, which were deficient for the NFκB precursor, and showed that NFκB is important in inflammatory responses involving p50 and RelA complexes. It is also conceivable that diseases caused by mutations in single genes, such as Lesch–Nyhan’s disease, could be modeled in hESCs. With the advent of recent successes in cloning of hESCs, it is conceivable that the generation of mutant lines could expedite mechanistic studies of certain disorders, especially developmental disorders. The regulation of proliferation and differentiation of stem cells may also borrow a few lessons from research progressing in the field of cancer, such as the observation that dicer miRNA circumvents the G1-S boundary, thus allowing stem cells to proliferate in an environment that generally does not promote proliferation, a fact that may be pertinent to cancer (Hatfield et al., 2005). Alternatively, miRNA have been shown to be involved in lymphoma, gastric, and lung cancers (Croce and
Calin, 2005) and in the differentiation of cardiac and fat tissue (Esau et al., 2004; Zhao et al., 2005). The use of stem cells as in vitro models to study the mechanism of action of such observations is feasible, especially when differentiation-/development-related events are involved.

We have studied glial differentiation of NSCs in relative detail and will discuss some of these findings below. NSCs differentiate into cells characteristic of the CNS—namely, neurons, astrocytes, and oligodendrocytes—upon withdrawal of the mitogen bFGF (Johe et al., 1996). The addition of specific factors also causes the directed differentiation of NSCs into CNS and non-CNS fates. While the addition of platelet-derived growth factor (PDGF) enhances neuronal differentiation, triiodothyronine (T3) causes oligodendrocyte differentiation, and ciliary neurotrophic factor (CNTF) along with the other gp130 family of cytokines is a robust inducer of the astrocytic fate (Johe et al., 1996; Bonni et al., 1997; Rajan and McKay, 1998). While the neuropoietic cytokines are robust inducers of glial differentiation of neural stem cells, other factors such as addition of bone morphogenetic proteins (BMPs), serum, and (to a much lesser extent) epidermal growth factor (EGF) also cause glial differentiation. Physical parameters are also important: NSC cultures that are maintained and passaged at high density are biased toward a glial fate. BMPs are a large family of soluble factors that have several important functions during development and are involved in the differentiation of several cell fates, including neural crest, cortical neurons, generation of germ cells, lung, tooth, kidney, cartilage, and bone, as well as in the formation of dorsalizing gradients during early differentiation (Liem et al., 1995, 1997; Nguyen et al., 2000; Sela-Donenfeld and Kalcheim, 1999; Weaver et al., 1999; Pizette and Niswander, 2000). There is only one major cytoplasmic signaling pathway described for BMPs, which seems paradoxical in light of the various BMP-mediated effects described. One mechanism by which a single factor may have such pleiotropic effects is by acting in conjunction with other factors present in the local environment. Specific cellular fates may be caused by the interaction of the signaling intermediates of two or more cytoplasmic pathways, leading to different transcriptional profiles and thus different phenotypes. Although this idea is conceptually simple, the details of the molecular interactions and their specific outcomes are not clearly defined in most cases. The complex differentiation phenotype that BMP exerts on NSCs was studied as a model system of this phenomenon.

BMPs cause differentiation into three fates in NSCs in our cultures: neurons, smooth muscle, and glia. The culture conditions used in this study do not support the survival of neurons, and thus the mechanism of differentiation of that lineage was not pursued. However, the differentiation of NSCs into smooth muscle and glia is mutually exclusive and can be regulated in culture by density. Dense cultures treated with BMP yield glia as a result of Stat3 activation, and sparse cultures differentiate into smooth muscle as a result of SMAD activation. BMP-mediated Stat3 activation occurs in the presence of a prior signal initiated by cell–cell contact in dense cultures and is blocked by the drug rapamycin. In dense cultures, the glial fate is dominant and occurs at the expense of smooth muscle (Rajan et al., 2003). In a developmental context, the increase to threshold levels of a differentiation signal by the concerted action of two or more pathways provides a model for the generation of various
phenotypes by intersecting gradients of factors. This example shows the involvement of a clinically relevant molecule in regulation of fate choice. It is thus not surprising that rapamycin analogues have been used on occasion in the treatment of glial tumors (Galanis et al., 2005). Also, even though FRAP and Stat3 are relatively ubiquitous proteins, the complex of these two proteins could be considered a marker of glial differentiation in stem cells (Fig. 20.3).

Recent data suggest that acetylation is possible on lysine 685 present on Stat3. In overexpression studies, it was suggested that acetylation is required for optimal activation of Stat3 in addition to phosphorylation (Yuan et al., 2005). It has also been shown that lysine 685 is a substrate for the CBP/p300 protein that forms a complex with Stat3 (Wang et al., 2005). Although the physiological significance of this observation has already been called into question (O’Shea et al., 2005), it is an interesting additional level of control that may be relevant to the glial differentiation of stem cells. The gp130 family of cytokines including CNTF, LIF, oncostatin M, and IL6 cause the activation of STAT proteins in several cell types including liver, kidney, neurons, astrocytes, and neural stem cells. However, the outcome of STAT activation in these various cell types is distinctive. Also, neural stem cells isolated from rats of different ages exhibit varying capacities for astrocytic differentiation. NSCs isolated

![Figure 20.3 Schematic of intracellular signals activated by BMP4 in neural stem cells](https://example.com/schematic.png)

This model is based on the data presented in Rajan et al. (2003). BMP4 causes the activation of at least two signaling pathways in neural stem cells. Upon ligand–receptor binding, the BMPR receptor complex releases FKBP12 and activates SMAD proteins, resulting in smooth muscle differentiation. The released FKBP12 may bind with rapamycin to inhibit FRAP or may act in some modified form to activate FRAP. FRAP then catalyzes serine phosphorylation (S*) of STAT to augment its prior activation by tyrosine phosphorylation (Y*) by another density-mediated signal. High cell density acts to promote basal STAT activation and DNA binding by an unknown signaling mechanism. This enhanced activation of STAT (Y*S*) causes efficient glial differentiation. Levels of activated STAT proteins in the cell receiving the BMP signal dictate whether the STAT or SMAD signal acquires precedence in the fate choice between smooth muscle and glia. See color plates.
from E12 rat embryos do not respond to LIF by differentiating into glia, while those isolated from E15 embryos do (Molne et al., 2000). The differences in outcomes could possibly be due to the epigenetic state of the cell receiving the signal, as has been suggested by Song and Ghosh (2004). There is also increasing evidence that genes such as bmi1 (which is one of the Polycomb Group of genes, an “oncogene” that regulates cell-cycle-related genes such as INK4a and p16ARF) are involved in the maintenance, proliferation, and differentiation states of hematopoietic and neural stem cells, while bmi1 forms part of a complex called the polycomb repressive complex 1 (PRC1), thought to maintain stable maintenance of gene expression by regulation of epigenetic chromatin modifications (Valk-Lingbeek et al., 2004). Mice deficient in bmi1 display deficits in neural stem cells and cerebellar neurons (Leung et al., 2004). Interestingly, they also show an increase in astrocytes (Zencak et al., 2005). These observations, though fragmented, strongly suggest that epigenetic modifications play an important role in fate choice of glia and, in a broader sense, in the decision of a cell to remain stem-like or to differentiate.

Epigenetic modification includes processes such as DNA methylation, histone modification, and ATP-driven chromatin remodeling. These processes function cooperatively to establish and maintain active or inactive chromatin states in cellular development. In general, acetylation of core histones and methylation of K4 of histone H3 correlate with transcriptional active (“open”) chromatin state, whereas deacetylation of core histones and methylation of K9 of histone H3 correlate with transcriptional repressed (“closed”) chromatin state. The proteins that mediate these interactions include DNA cytosine methyltransferases, histone acetyl transferases, histone deactylases, histone methyltransferases, and other specialized proteins that mediate the interaction of proteins with nucleosomal DNA. The precise function of these proteins in plasticity of stem cells is not clear. It appears that the expression of Oct-4 is required for the maintenance of pluripotency, self-renewal, and survival (Kehler et al., 2004). A global study of hESC promoters that bind the three markers Oct-4, Nanog, and Sox2 yielded clues as to genes that remain inactive when bound to these transcription factors and others that appear to be activated (Boyer et al., 2005). Follow-up studies may delineate the epigenetic reasons for this observation, including a possible mechanism for the role of these three factors in global regulation of “stemness” (genes of promoters that are activated) and phenotypic commitment (genes of promoters that are repressed) in ESCs. While it is generally agreed that global transcriptional, translational, and epigenetic characteristics of ESCs are unique, the factors that orchestrate these changes are still elusive. In a preliminary experiment to identify these factors, it was determined that characteristics of ESCs could be conferred on somatic cells by fusion (Cowan et al., 2005). These authors showed that the nucleus from the mature fusion partner could be appreciably “reprogrammed” such that it acquired some of the epigenetic features resembling an undifferentiated hESC. It appears that cytoplasmic factors play an important role in the initiation and maintenance of the epigenetic state of the nucleus.

Thus, as expected, there appears to be an intimate relationship between cytoplasmic and nuclear signals. Another facet of regulation acquiring importance is stable, nontranslated RNA molecules called micro RNAs (miRNAs), which lead to reduced translation of specific proteins by either mRNA degradation or inhibition of
translation (Alvarez-Garcia and Miska, 2005). Although small interfering RNAs (siRNAs) were first discovered in C. elegans (Lee et al., 1993) and have since been used extensively as research reagents that inhibit eukaryotic—including mammalian—gene expression, functionally similar miRNAs have been isolated from mammalian cells, and two have been shown to be involved in cardiac muscle and adipocyte differentiation. Zhao et al. (2005) have shown that miR-1 targets a transcription factor hand-1 that controls the proliferation of ventricular cardiomyocytes during development. Similarly, miR-147 has been shown to regulate adipocyte differentiation, possibly by inhibiting translation of ERK5 (Esau et al., 2004). On a more global level, mutation of the enzyme dicer, which is part of the cellular machinery responsible for the generation of miRNAs, causes a G1-S block in stem cells in Drosophila, implying that miRNAs are essential for the proliferation of stem cells in their native environment (Hatfield et al., 2005).

The initial use of mouse ESCs was in the creation of “knock-out” and “knock-in” mice, which have contributed immensely to the study and understanding of mammalian developmental biology. The past decade has seen a revolution in this field with the isolation of murine, rodent, and human stem cells from a variety of sources, including blastocysts, the brain, and bone marrow. In spite of the obvious relevance of the human stem cells to the medical sciences, important discoveries about the basic biology of the stem cells themselves and some of the associated developmental phenomena in mammals are continuing to be made in the rodent and murine systems, and to a lesser extent in invertebrates and lower vertebrates such as Drosophila and Zebrafish. All these studies—molecular, cellular, systems biology, and clinical—will contribute to the gestalt of knowledge about stem cells and will aid immensely in the successful exploitation and manipulation of this powerful system to our advantage.

20.4.2 Creation of Applied Systems for Drug Discovery

Stem cells are potentially useful in several stages of the drug discovery process, including target discovery and validation, secondary screens of candidate molecules, and toxicological studies of potential drugs. Alternatively, stem cells could be used as in vitro experimental systems to determine the precise mechanism of action of a candidate drug. Using genetic manipulation technology, ESC lines may be created that express specifically mutated proteins or lines that lack specific genes, either of which may be used to study the mechanism of action of particular targets or as a platform for screening drugs that reverse a particular phenotype. Alternatively, “humanized” mice may be produced that express the human version of the gene of interest, which may then be used for some in vivo testing of drug action and ADME-TOX (administration, distribution, metabolism, excretion, and toxicity of candidate drug molecules). Mouse ESCs are already being used by companies in drug screening applications (McNeish, 2004). The creation of person-specific and disease-specific hESCs would also aid drug discovery efforts. The former has potential for autologous cellular therapies as well as for pharmacogenomic studies where genomic profiles are correlated with the most effective therapeutic regimen. Disease-specific hESCs could lead to the design of experiment
systems for target validation and subsequently to screening platforms that may be used for secondary screens of compounds of interest.

In addition to the advantages that stem cells present in the areas of drug discovery, they also represent a potentially limitless source of “normal,” untransformed, diploid cells. This is especially relevant to human systems, because most primary human cells do not propagate beyond a few generations in culture. This property of stem cells could potentially be exploited for the generation of large quantities of cells needed to generate screening platforms for high-throughput screening of chemical libraries. Secondly, drug discovery platforms that would otherwise be challenging to generate on a large scale, such as platforms that employ neuronal cells, could be derived from stem cells or precursors that have been differentiated \textit{in vitro}. Some groups have been successful in the generation of neurons from hESCs (Reubinoff et al., 2001; Schuldiner et al., 2001), which suggests that this may be possible in the future. Even specific types of neurons, such as dopaminergic or serotonergic, could possibly be derived \textit{in vitro} stem cell differentiation: The former has particular relevance to Parkinson’s disease, and the latter has relevance to psychiatric illness. Finally, stem cells and their derivatives could be used for toxicity studies of candidate drugs. They would provide a preferable alternative to the transformed cell lines on which these studies are currently being performed. In addition, they could potentially be differentiated into liver cells, which would provide an ideal platform for toxicity studies to be performed \textit{in vitro}, and for which a viable alternative does not exist.

\section*{20.5 Projections and Conclusions}

While the enormous potential presented by stem cells exists without a doubt, there are scientific and other pragmatic problems that need to be addressed before this potential is realized. Procedures need to be developed for the reproducible scale-up of all kinds of human stem cells, with precise definition of the resultant cells. The cells could be undifferentiated stem cells (for possible use in transplantation and drug screening assays) or differentiated sufficiently along a particular lineage (as in hESCs differentiated into neuronal precursors of a specific kind for use in transplantation) as dictated by the end use. The scale-up will have to take into account the following: optimization of culture of stem cells in suspension or adherent cultures; culture systems that use defined media and are devoid of animal-derived components; cost-effective culture of undifferentiated stem cells that may then be used as such or further differentiated into fully or partially differentiated cultures; and panels of markers sufficient to characterize the desired endproducts, be they a single population of undifferentiated stem cells or a complex population of cells that have been partially differentiated cells along a particular lineage. Also, the ability of the stem cells to remain stable in their phenotypes through the rigors of the drug screening process will have to be ensured. Understanding the molecular intracellular signals that mediate the proliferation, survival, differentiation, senescence, and motility of stem cells will be integral to these efforts and to the ability to manipulate these cells with precision and reproducibility. When used as cellular therapeutics, the ability to control these cells
post-transplantation will be an additional desirable manipulation. The ultimate use of these cells to better the human condition, both directly in the clinic as cellular therapeutics or as surrogates in the pharmaceutical industry, will be dictated by our understanding of these complex cells themselves and by the reconciliation of philosophical differences pertaining to ethical issues surrounding embryonic stem cells, particularly in the United States.

In spite of these considerable hurdles, scientists worldwide are tackling this problem with interest and determination, due to both the interesting scientific challenge it poses and the significant clinical gains it presents. In addition to appreciable efforts in the academic quarter, several companies are investing substantial amounts of manpower and resources to achieve practical usability of these cells. These include Amgen, Cell Genesys, Chemicon, Cytogen Therapeutics, Geron, Invitrogen, NeuralStem, Novocell, Osiris Therapeutics, Pfizer, ReNeuron, StemCells, Inc., and ViaCell, to name a few. While transplantation of bone marrow as an adjunct for cancer therapies is not uncommon, neural, mesenchymal, and hematopoietic stem cells are increasingly being used in clinically relevant and experimental scenarios. Although initial neural stem cell trials were related to Parkinson’s disease, stem cell transplantation is now being thought of as a viable option for replacement therapies in spinal cord injuries. Companies such as Geron, NeuralStem, and Stem-Cells, Inc., have invested substantial resources into determining the practical use of these cells in animal models (Vogel, 2005; Cummings et al., 2005; Keirstead et al., 2005) and hope to progress to clinical trials. Transplantation of whole bone marrow and selected cell populations, including mesenchymal stem cells, has been revisited several times in therapies relating to heart cell replacement (Pittenger and Martin, 2004; Shim and Wong, 2004), and Osiris Therapeutics has initiated clinical trials with human MSCs. Geron is also anticipating the release of cultures derived from hESCs to be used for toxicity studies of candidate drugs (Sinha, 2005). The nature and complexity of the challenge presented by stem cells makes it an ideal discipline for active and extensive collaborations between the academic and corporate scientific establishments at national and international levels.

ACKNOWLEDGMENTS

We would like to acknowledge the A-T Children’s Project, NIH, and the Michael J. Fox Foundation, who have supported some of the experiments described here.

REFERENCES


Villa et al. (2000).


INDEX

Aberrant glycosylation, 343
ACD database, 42
Acetylation, 348
Acetylcholine (ACh), 39
Acetyl cholinesterase (AChE) inhibitors, 38
Achiral compound, 231
Achiral drugs, 241
Activator ubiquitylation, 339
Active sites, 7
Activity-dependent neurotrophic factor (ADNF), 205
Acute myeloid leukemia (AML), 373
Acute promyelocytic leukemia (APL), 373, 374, 376
Acute promyelomonocytic leukemia (APML), 416
Acutolysin-C, 42
ADAM12, 42
ADAM17, 43
Adamalysin, 42
ADAM & EVE, 38
ADAMs, 42
Adenine moiety, 177
Adenosine deaminase (ADA), 394
Adenosine diphosphate(hydroxymethyl) pyrrolidinediol (ADP-HPD), 169
Adenovirus protease, 14
Adipose derived stem cells (ADSCs), 480
ADME, 52
ADME/Tox, 64, 82, 498
outcome, 77
problems, 77
ADP-carbocyclic-ribose (cADPcR), 189
ADP-ribose, 179
analogues, 181
hemiacetal, 172
monomer, 176
β-Adrenergic block agents, 234
Affinity, 40
AICAR transformylase, 13, 50
Airway hyper-responsiveness (AHR), 316
Alanine screening, 8
Alcipain-3, 51
All-d-amino acid analogue, 311
Aldose reductase (AR), 13, 41
Allosteric agonists, 311
Allosteric sites, 7
All trans-retinoic acid (ATRA), 374
All trans-retinoid acid (ATRA), 416
Alpha1a antagonists, 48
homology model, 48
receptor, 48
Alzheimer's disease (AD), 203
AMBER, 9
AMD3100, 316
Ameliorative effects, 480
Aminohydroxylation, 268
Aminoxypentane (AOP)-RANTES, 309
AMPA/kainite receptor, 205
AmpC β-lactamase, 14
Amphetamine, 234, 260
β-Amyloid, 204
accumulation of, 204
Amyotrophic lateral sclerosis (ALS), 440
Anchor and grow, 11
Anchors, 5
Annotated ligands, 44
Antagonist CP-96345, 48
Anti-apoptotic proteins, 411
Anti-cytokine therapy, 316
Antioxidants, 205
Antisense oligodeoxynucleotides (ODNs), 417
Antisense oligonucleotides, 287
Aplysiatoxin, 132

Drug Discovery Research: New Frontiers in the Post-Genomic Era, Edited by Ziwei Huang
Copyright © 2007 John Wiley & Sons, Inc.

507

SOFTbank E-Book Center Tehran, Phone: 66403879,66493070 For Educational Use.
Apogossypol, 419
Apoptosis, 130, 292, 408
Apoptosis-blocking proteins, 411
Apoptosis inducing factor (AIF), 410
Apoptotic pathway, 293
AQSAR, 91
Aquaculture, 131
Arabidopsis kryptonite mutant, 346
Arabidopsis PARG, 169
Arabidopsis thaliana, 169
Area under curve (AUC), 473
Arg–Asp salt bridge, 284
Arginine Methylation, 346
A-ring, 140
A-ring-modified analogues, 146
A-ring-modified spacer domains, 143
Arsenic, 383
Artificial intelligence approach, 67
ASLW, 311
Aspergillus nidulans, 347
AstraZeneca compound, 56
Asymmetric synthesis, 255
of L-DOPA, 263
Asymmetric transformation, 261
Asymptomatic stage, 302
ATP binding cassette (ABC), 482
transporters, 179
ATRA, 381
Atropine, 233
Atypical isozymes, 131
AutoDock, 9, 10, 12, 26, 73
Azobenzene (AZO) linkage, 283
Bacteroides fragilis, 44
Baculoviral IAP repeat (BIR), 37, 424
Bak BH3 peptide, 29
Baker’s yeast, 247
Basic resolving agents, 259
Batrachotoxins, 261
Bax, 130
B-cell lymphoma-2 (Bcl-2), 14, 130, 292
family, 29
Bcl-2-dependent pathway, 293
Bcl-2 homology-3 (BH3)
  binding site, 36
domain, 419
Bcl-xL binding affinities, 36
Bcl-xL/Bak complex, 293
Benzoate ester, 151
BH3I-1SCH3, 37
BH3-only proteins (BOPs), 423
Binary fingerprints, 68
Binding site(s), 7, 8
Bioavailability, 473
Biocatalysis, 245
BioSpecs, 42
Biotransformation, 245, 247
BJH pore size distribution curves, 363
Bone morphogenetic proteins (BMPs), 495
B-ring oxygen, 134
Broad-spectrum anti-HRV agents, 52
Brown’s allylation, 141
Brunauer–Emmett–Teller (BET) surface
area, 363
Brute-force method, 8
Bryostatin, 1, 128–130, 154, 157
analogues, 14, 156
functional analogues of, 157
mode of action of, 130
substructures of, 14
BTB02809, 41
Bugula neritina, 128, 131
BUILDER, 94

Ca2+
  channel inhibitors, 204
  mobilization, 186
  mobilizing messengers, 186
  release channels, 187
  releasing activity, 192
  signaling, 186
    membrane-permeant agonist of, 189
    spatiotemporal, 198
    signals, 195
      local, 186
cADPR, 187, 192
  analogues, 188, 200
  model for the interaction of, 200
  structure of, 187
  structure–activity relationship of, 195
cADPR-mediated amplification, 195
Calcineurin, 207
Calcium-dependent conventional isozymes,
  131
Calcium-release activities, 195
Calmidazolium, 54
Calmodulin (CaM), 187, 290
CaM-binding proteins (CaMBP), 290
Cambridge Crystallographic Database, 10
CypA inhibitors, 119
Cysteine proteases, 51, 408
Cytochrome P450 oxygenases (CYPs), 472
Cytokine acting, 340
Cytokine-stimulated glial cells, 313
Cytolytic T-cells (CTLs), 410
Cytoplastic markers, 481
Cytosine deaminase (CDA), 492
Database queries, 48
Daxx, 377
Daylight fingerprint, 65
Death domain (DD), 409
structure, 412
Death effector domain (DED), 409
Death inducing signaling complex (DICS), 409
Death-promoting BH3 domain, 293
Decoy ligand, 16
Dehydroative cyclization, 138
Deimination process, 347
D-enantiomer, 235
(R)-(--)-Denopamine, 248
De novo design, 90
programs, 94
De novo drug design, 109
fragment-based, 110
De novo ligand design, 102
Deoxycytidine kinase (dCK), 394
Designed host–guest interactions, 283
Desolvation term (DS), 12
Diacylglycerol kinase (DGK), 155
Diastereoisomer(s), 232
composition, 232
excess, 232
Diastereomer crystallization, 257
interconversion, 261
z-Dicarbonyl amide, 223
Diels–Alder cycloaddition, 136
Dihydro-dipicolinate, 14
Dihydrofolate reductase (DHFR), 18
Dihydroxylation, 267
Dimer interface, 7
Dimethylidoxirane (DMDO), 138
Dipeptidyl peptidase IV (DPP-IV), 471
Directed tweak algorithm, 51
Direct strategies, 90
Discrete substructure analysis, 67
Discriminant function analysis, 69
Disseminated intravascular coagulation (DIC), 374
Distance-compatible match searches, 100
Distance-dependent dielectric (DDD), 12
DiverseSolution (DVS), 77
Diversity analysis, 76
Dms program, 8
DNA
  damage checkpoint kinases, 56
  gyrase, 13
  intercalators, 172, 181
  methylation inhibitor, 7
  repair mechanisms, 166
DOCK, 26, 72, 98
  algorithm, 99
  energy score, 27
  energy scoring function, 9
  program, 8, 11, 37
DOCK 3, 12
DOCK3.5, 29, 41
DOCK4.0, 46, 47, 100, 115
DOCK4.01, 54
DOCK 4/5, 12
Docking, 5, 18, 25, 52, 71, 81, 98
  algorithms, 25, 26
  library, 18
  methods, 64
  programs, 26
  rigid-body, 100
  soft, 74
Docking-based screen, 28
Docking-based virtual screening, 109
Dopamine D2 receptor, 46
Dose-limiting toxicity, 130
DPP-IV, 40
(S)-Dropropizine, 237
Drosophila, 426
Drug chirality, 230, 240
Drug design process, 4
  structure-based, 4, 7, 8, 40, 104
Drug development, antineoplastic, 50
Drug discovery
  computer-based, 75
  structure based, 108
  task of, 108
Drug likeness, 113
Drug receptor sites, 233
DrugScore, 27, 74
DSA, 67
  method, 79
  visualizer, 67
Dual-tropic strains, 303
DUBs, 337
Dynamic histone acetylation, 348
Dynamic post-translational modification, 337
Effector phase, 316
Embelin, 38
Embryonic stem cells (ESCs), 480
Empirical scoring functions, 74
Enantiomer(s), 231, 232
  composition, 232
  of dobutamine, 237
  excess, 232
(S)-Enantiomers, 236
(R)-Enantiomers, 236
Enantioselective technologies, 262
Encephalomyelitis (EAE), 313
Endogenous activator DAG, 132
Endogenous cannabinoid system (ECS), 47
Endoplasmic reticulum (ER), 410
Enthalpy of interaction, 5
Enzymatic synthesis of eugenyl a-glucoside, 254
Enzymatic transesterification, 252
Enzyme-based catalysis, 246
Epidermal growth factor (EGF), 495
Epidermal growth factor receptor (EGFR), 54
Erβ binding assay, 40
Erβ-selective compounds, 39
E-state, 68
Ester-based spacer domain, 145
Esterification esterases, 251
Ester linkage, 142
Estrogen receptor(s) (ER), 39, 468
  antagonists, 416
Ethinicline, 172
Ether linkage, 142
Euclidean space, 112
EUDOC, 45
EUDOCK, 12
Excitotoxicity, 204, 440, 442, 446
Extracellular regulated kinase (Erk), 324
Extrinsic apoptotic pathway, 409
Extrinsic pathway, 411
5-FC, 492
5-FU, 492
INDEX

FADD, 409
Falcipain-2, 3, 51, 52
False-negative hits, 27
False-positive hits, 27
Farnesyl protein transferase (FPT) inhibitors, 93
Farnesyl pyrophosphate (FPP), 45
Farnesyl transferase (FT), 45
 inhibitor, 102, 345
FasL, 413
FDA approved drugs, 241
FDA-approved NMEs, 241
Fermentation, 246
Fetal bovine serum (FBS), 486
Fingerprint, 66
 scaling, 66
Fisher log score, 67
FK506, 207, 212
 acyclic analogues of, 217
 binding domain of, 221
 binding protein (FKBP) ligand, 102
 neuroprotective effect of, 207
 structure of, 217
FK506 bind proteins (FKBPs), 206
FKBP12, 187, 206, 215
 structure of, 215
FKBP51, 206
FKBP52, 211
FKBP12–FK506 complex, 219
Flexible ligand docking, 101
FlexX, 10, 26, 51, 73, 98
FlexX/FlexE, 12
FlexX-Pharm, 74
 program, 56
FlexxScore, 74
FLIP, 411, 415
proteins, 415
Fluorescence-activated cell sorting (FACS), 481
Fluorescence-based method, 172
Fluorescence polarization assay, 294
Fluorescence polarization (FP) binding assay, 36
Fluoxetine, 248
Focused library, 111, 116, 118
Focused virtual library, 42
Force-field-based scoring functions, 73
Force-field derived scoring functions, 102
Force-field parameters, 45
Force-field scoring functions, 27
Four-step RBVS protocol, 49
Fragment libraries, 116, 111, 118
Fragment-linking approach, 110
FRED, 26
Free energy, 5
F-score, 27
Functionalization, 150
Function-oriented synthesis, 128
Fusion protein, 154
G.38, 247
GABA transport, 271
Gallotannin, 178, 180
GAP-43, 207
GAPDH, 445
GB/SA method, 10
GB/SA scoring function, 10
GAT1, 271
GCN4, 285
GDNF gene-therapy, 206
Geldanamycin, 211
G262E mutant PARG, 169
Generalized born (GB), 10, 12
Genetic algorithm(s) (GA), 12, 69, 111
GGG motif, 169
Glial cell line-derived neurotrophic factor (GDNF), 205, 490
Glide, 10, 12, 26, 40, 73, 75
GlideScore, 74
Gliomas, 491
Glucagon-like peptide-1 (GLP-1), 471
Glucocorticoid receptor (GR), 468
Glucoside scopolin, 39
Glutamate, 440, 442
 receptor antagonists, 204
Glutamaterigic neurodegeneration, 204
Gly-262, 169
Gly-263, 169
Gly-264, 169
Glycation end products (AGEs), 204
Glycoprotein, synthesis of, 342
Glycosylation, 341, 343, 350, 395
 enzymes, 344
 sites, 342
Gnidimacrin, 132
GOLD, 10, 12, 26, 39, 73, 74, 98
GOLD2.0, 47
Goldscore, 27, 73
 fitness function, 40
Goodness-of-fit, 72
Gossypol, 419
GPCRs, 325, 50
GPI, 177, 214, 215
GPI anchor, 342
attachments, 341
G-protein coupled receptor(s) (GPCR), 46, 49, 75, 300, 303, 379
Green fluorescent protein (GFP), 154
GRID, 94
GROMACS, 47
Growth factors, 223
G-score, 27
Guanidinium-based receptors, 283
Guanine nucleotide exchange factors (GEFs), 50
Guvacine, 272
GX-15-070, 419
H2A, 339
HA14-1, 29
HAD, 322, 324
H2B, 339
H2O2-scavenging properties of tannins, 180
Hairy cell leukemia (HCL), 400
Hansch equation, 91
HAUSP, 339
hCypA-focused library, 119
H-dTMP incorporation, 178
HeLa cells, 130
α-Helical mimetics, 287
α-Helix, 281, 282, 286
Helix stabilization, 283
Hematopoietic stem cells (HSCs), 480
Hemiketal functionality, 139
Hepatitis C virus (HCV) helicase inhibitor, 91
hESCs
disease-specific, 498
person-specific, 498
Heteroatomic triads, 133
Hexamethonium, 454
High-throughput (HTP) strategies, 24
High-throughput screening (HTS), 24, 58, 64, 90
library, 18
results, 69
technology, 79
Hinge-region backbone, 54
Histone, 346
acetylation, 346, 348
arginine, 346
lysine methylation, 346
methylation, 346, 347
Histone deactylases (HDACs), 417
Hit(s), 3, 10, 40, 63
Maybridge, 53
physicochemical profile of, 79
Hit rates, 17
HIV-associated dementia (HAD), 312
HIV-1 IN inhibitor, 94
HIV-1 RNATAR inhibitor, 102
HIV-1 strains, 301
H-labeled thymidine 5'-triphosphate (H-dTTP), 178
Homology modeling, 6, 7, 53
ligand-supported, 48
Homology models, 51
rhodopsin-based, 49
Homology module, 42, 47
Homology proteins, 112
HPLC, 54
HPRTase, 14
Hsp90 inhibitor, 49
Human double minute 2 (HDM2), 287
Human embryonic kidney 293 (HEK293) cells, 296
Human NSCs (hNSCs), 314
Human rhinovirus (HRV), 52
Huntington’s disease (HD), 203
Hydrocyanation, 249
Hydrogenation, 134, 136, 263
Noyori asymmetric, 138
of ketones, 265
Hydroformylation, 266
Hydrogen bonding groups, 70
Hydrogen phthalate, 260
Hydrolysis
enantioselective, 251
enzymatic, 251
Hydrolitic enzymes, 254
Hydrolyzable tannins, 176
Hyperacetylation, 348
Hyperleukocytosis, 382
Hyperphosphorylation, 340
HypoGen, 93
Hypotheses, 71, 92
Hypoxia-ischemia (HI), 490
IBS virtual-screening approach, 45
IC50 values, 56
ICM, 26, 73
score, 54
virtual-screening program, 54
ICM-based VS procedure, 54
ICM-Dock, 12
IDD594, 41
IFN-β, 492
Imino sugars, 344
Immobilization, of antibiotics, 358
Immunophilin(s), 206
ligands, 210
Immunotherapy, 130
IMPDH, 13
Incremental construction (IC), 12, 11
Incretion, 340
Indirect strategies, 90
Induced fit, 200
problem, 74
Ingenol, 132
Inhibitor-binding pocket, 50
Inhibitor of apoptosis proteins (IAPs), 37, 424, 425
Intellectual property (IP), 79
Interferons (IFNs), 379, 381
Intracellular Ca²⁺, 204
Intramolecular cyclization, 193
Intraneuronal inclusions, 339
Intrinsic pathway, 408
Invaginations, 8
ISIS/base search program, 92
Isoprenaline, 236
Isotherms analysis, 363
JNK kinase ligands, 67
Jurkat cells, 38
Jurkat T-lymphocytes, 193
Key-lock principle, 72, 82
Kinase targets, 74
Kinetic resolution, 261
with enzymatic method, 261
L-685818, 213, 214
30’s loop, 320
Lactate dehydrogenase (LDH), 417
LCP apoptotic step, 340
LD1.0 software, 12, 111
Lead(s), 3, 63
Lead-structure searches, 47
LeapFrog, 94
Leave-N-out method, 70
LEGEND, 94
Lennard-Jones potential, 36
Leptin, 467
Lesch–Nyhan’s disease, 494
Leucine-rich nuclear export signal (NES), 167
Levocetirizine, 239
Levorotatory alkaloid hyoscyamine, 233
Ligand
docking, 45
flexibility, 26
free energy, 9
Ligand-based approach, 25
Ligand-based method, 64
Ligand–receptor docking, 98
LigBuilder, 94
Lipid protein(s), 344, 345, 350
Lipinski’s rule of five, 40, 43, 52
Lipodation, 344
Lipoprotein
synthesis of, 344
Liver receptor homolog 1 (LRH1), 473
Liver X receptor (LXR), 473
Loracarbef, 268
LSD1, 346
Luche reduction, 136
Ludi, 27, 74, 94, 219
Lysine demethylation, 346
Lysine methylation, 346
MACCS keys, 65
Macrophage inflammatory protein (MIP)-1 β, 304
Macrophage (M)-tropic strains, 302
Macrotransacetalization, 138
MAPK, 445
Massively parallel signature sequencing
(MPSS), 480, 484
Matriptase, 14
Matsuda’s strategy, 193
Maximum dissimilarity selection, 76
Maximum similarity elimination, 76
Maximum-tolerated dose (MTD), 401
MaxMin function, 77
Maybridge, 42
database, 48
MCH-1R, 49
MDL drug data report (MDDR), 68, 79
databases, 78
MDL toxicity database, 79
Melanin concentrating hormone (MCH), 49
receptor, 308
Memantine, 449, 452, 454, 456
Membrane-permeant agonists, 198
Merged target structure ensemble, 12
Mesenchymal stem cells (MSCs), 480, 491
Metabolic compromise, 204
Metabolizing enzymes, 233
Metalloenzymes, 44
Methadone, 239
Methadone’s enantiomers, 239
1β-Methylcarbapenem, 266
(5)-Metolachlor, 264
Me too compounds, 66
Metropolis algorithm, 9
Microbial hydroxylation of progesterone, 251
Microstructure parameters, 359
Microtubule-modulating drugs, 423
Midbrain dopamine (DA), 487
ML-IAP, 424
MM-GB/SA, 10
MM-PB/SA, 10, 12
MOBILE approach, 48
MODELLER6.0, 53
MOE, 26
Molecular diversity, 76
Molecular docking, 5, 98, 115
Molecular dynamics (MD), 12, 11, 26
multicanonical, 53
simulations, 10, 47
Molecular fingerprint, 65
Molecular recognition, 25
MOLT-4, 136
Molybdate, 211
Molybdenum enzyme, 44
Monoclonal antibodies (MAbs), 316
Monocyclic β-lactam, 266
Monocyte chemo attractant protein-1
(MCP1), 492
Monoubiquitylated proteins, 339
Monte Carlo (MC), 12
genetic algorithm, 26
simulations, 26
techniques, 11
MPTP, 490
lesioning model, 215
MSK-039, 54
Multidrug resistance (MDR), 130
Multidrug resistance-associated proteins
(MRP), 472
Multiple copy simultaneous search (MCSS), 8
Multiple sclerosis (MS), 312, 313, 402
Munc proteins, 155
Murine astrocytes, 178
Muscle-specific PPARγ knockout, 469
Myocardial infarction (MI), 491
Myristate analogues, 145
Nacetylglucosamine, 342
NAD+ consumption, 179
Naproxen, 245, 268
Natural chemokines, 303
Natural killer (NK), 410
Natural products, 243
isolation of, 244
NCI-H460, 136
Nerve growth factor (NGF), 203, 205
Neural networks, 69
Neural stem cells (NSCs), 314, 480, 487, 490
Neural transplantation, 223
Neurodegenerative disease, 203, 440
mechanism of, 223
Neuroimmunophilin, 206
ligands, 203, 212, 224
neuroregenerative mechanism of, 210
Neurokinin-1 (NK1) receptor, 13, 48
Neurokinin peptide receptor (NK1R), 308
Neurologic diseases, 440
Neuronal cell death, 204
Neuronal sprouting, 223
Neuropoietic cytokines, 495
Neuroprotectants, 203, 204
Neuroprotection, 164
Neuroregenerative mechanism,
calcineurin-involved, 210
Neurotoxic poisons, 204
NFκB, 412, 469, 494
NGF-like organic compounds, 206
NH-analogue of ADP ribose, 177
Nickel-containing enzyme, 44
Nicotinamide, 164
NINA A, 207
Nitrogen adsorption–desorption isotherms,
361
NitroMemantines, 458–460
N-linked glycosylation, 341
NMDA receptors (NMDARs), 204, 440, 444, 446
NMR-derived constraints, 36
NMR spectroscopy, high-resolution, 6
NMR structure, high-resolution, 7
Nobotanin, 174
Nobotanin B, 176, 178, 179
Nobotanin E, 178
Nobotanin K, 178
Non-Hodgkin’s lymphoma (NHL), 417
Non-insulin-dependent diabetes mellitus (NIDDM), 470
Nonpeptide scaffolds, 285
Nonradiometric assay, 172
Novel chemical entities (NCE), 57, 429
Novel ligands, 15, 235
Novobiocin, 93
NRB 04079, 48
NSC23766, 51
Nuclear bodies (NBs), 377
Nuclear localization signal (NLS), 167
Nuclear magnetic resonance (NMR), 4, 58
Nuclear mitotic apparatus (NuMA), 375
Nuclear Overhauser effects (NOE), 286
Nuclear RA receptors, 375
Nucleophosmin (NPM), 375
Nucleoside permeation, 398
Oblimersen sodium, 417
Octanoate ester, 151
Oen B, 178
Ofloxacine, 237
O-GlcNAc residues, 342
O-GlcNAcylation, 342
Olefin metathesis, 283
Oligoaamid foldamer, 294
Oligomeric ellagitannins, 174
inhibitory activity of, 175
Oligosaccharides, 342
oligosaccharide scaffold, 286
O-linked glycosylation, 341
Omapatrilat, 251
Omega (OpenEye), 11
Omeprazole, 262
O-methylation, 195
OMI, 425
Oncogenic mutations, 293
OPC-29030, 262
Open-channel block, 447
Optical activity, 232
Optical rotation, 232
Organic anion transporting polypeptide 2 (OATP2), 472
Orientation searching, 100
Oxidation, 250
Oxidative cleavage, 136
Oxidoreductases, 254
Palmitoyltransferase inhibitor, 345
Pargamicin, 177
PAR-deficient animals, 170
Parkinson’s disease (PD), 487
PARP-1, 164
inhibitory compounds, 166
isozymes, 164
overactivation of, 164
overstimulation of, 164
pindle-associated, 164
PARP-deficient mice, 166
Partial least squares (PLS), 68
Pathologically activated therapeutics (PAT), 449
PB/SA methods, 10
PCA-based clustering, 53
PD 124, 177, 966
Peak-trough fluctuation (PTF), 473
Pentasaccharide, 286, 287
scaffold, 285
Peptidylprolyl cis–trans isomerases (PPIases), 206
Peroxisome proliferator-activated receptor δ (PPARδ), 474
Peroxisome proliferator-activated receptor gamma (PPARγ), 117, 467, 468
focused library, 118
knockout, 468
Peroxisome proliferator-activated receptors (PPARs), 117
P-glycoprotein (P-gp), 130
Pharmacokinetic outcome, 77
Pharmacokinetics, 421
Pharmacophore, 70, 92
analysis, 64, 70
constraint, 82
fingerprinting, 71
fingerprints, 66
generation of, 92, 93
hypothesis, 134
model, 41, 48, 52, 92
screen, 56
searches, 4
structure-based, 39
of type II 5α-reductase inhibitors, 93
Pharmacophore-based design, 90
Pharmacophore-based virtual screening, 92
Pharmacophoric models, 102
Pharmacopoeia, 240
PharmPrint method, 71
Phe–Lys cation–π interaction, 284
Phenanthridinones, inhibitory effects of, 166
Phorbol esters, 132, 156, 157
Phorbol-12-myristate-13-acetate (PMA), 156
Phosphate buffered saline (PBS), 359
Phosphodiesterase type 1 (PDE-1), 43
Phosphodiesterase type 5 (PDE-5), 43
Phosphorylation, 339, 341, 350
Phytanoyl-CoA-α-hydroxylase (PAHX), 211
Picenadol, 236
Picolog recognition domain, 145
Pivalate analogue, 145
PKCδ, 155
P-labeled NAD+, 171
P-labeled PAR, 171
Plakoside A, 265
Plasmodium falciparum, 51
Platelet-derived growth factor (PDGF), 495
PLP function, 10
PMF, 27, 74
PML, 377
oncogenic domains (PODs), 377
Poisson–Boltzmann (PB) terms, 10
Polyacrylamide gel electrophoresis, 171
Poly(ADP-ribose) (PAR), 163
accumulation, 164
catabolism of, 178
hydrolysis, 168
substrate analogues of, 176
Poly(ADP-ribose)glycohydrolase (PARG), 163, 164, 171
inhibitors, 172
major classes of, 181
+/−(−) mice, 170
enzymatic activity, 171, 178
hydrolytic activity of, 180
inhibition of, 164
by tannin-related drugs, 179
isoforms, 167
recombinant bovine, 169
second-generation intercalator-like, 177
silencing of, 179
spindle-associated, 164
Poly(ADP-ribose)/polymerases (PARPs), 163
family, 167
Poly(ADP-ribosylation), 170, 179
Polycomb repressive complex 1 (PRC1), 497
Polyglutamine diseases, 349
Pose, 73
Post-translational modification process, 337
P-PAR assay, 172
PREDICT, 46
3D model, 46
algorithm, 46
Preferential crystallization, 256, 257
Pregnane X receptor (PXR), 472
Primary sclerosing cholangitis (PSC), 403
Principal component analysis (PCA), 53, 76
Principal component regression (PCR), 68
Proapoptotic subfamily, 293
proteins, 293
PROCHECK, 47
Profavine, 172
PROLIGAND, 94
Promiscuous inhibitors, 79
Promyelocytic leukemia zinc finger (PLZF), 375
Promyelocytic leukocyte zinc finger gene (PLZF), 375
Propafenone (PPF), 234
Prosurvival subfamily, 293
Protein acetylation, 348
Protein-binding sites, 233
Protein glycosylation, 341, 343, 344
Protein kinase C (PKC), 130, 207
isozymes, 131
translocation assay, 154
Protein kinase D (PKD), 155
Protein lipodation, 344, 345
Protein post-translational modification, 336, 351
Protein tyrosine phosphatase-1B (PTP-1B), 41
Pseudomonas diminuta, 45
PTP-1B, 17
Putative ligand(s), 5, 8, 41, 98
Pyrrholidine cis-hydroxyls, 177
Pyruvate decarboxylase (PDC), 254
Quantitative structure–activity relationship (QSAR), 4, 5, 64, 68, 78, 90
approaches, 70
classical, 91
models, 47, 71
modern, 91
Quantitative structure–property relationships (QSPR), 68
Quercetin, 180
(+)-Quinidine, 244
(−)-Quinine, 244
QXP, 9, 12
Racemates, 241
Racemic dropropizine, 237
Rac1-specific inhibitors, 50
Rac1–Tiam1 complex, 50
RAIDD protein, 410
RANTES, 308
Rapamycin, 207, 208
RARs, 376
RasGRP, 155
activation of, 155
Ras protein, 345
RCP168, 318, 319, 321
Receptor-based approach, 25
Receptor-based virtual screening (RBVS), 25–27, 56
work-flows, 28
Receptor(protein)-based method, 64
Recognition domain, 138
Recognition sites, 7
Recursive partitioning (RP), 69
5α-Reductase inhibitors, 95
Resolution, 255
of 2-aryalkanoic acids, 260
lipase-catalyzed, 245
of racemic sulfoxides, 262
via preferential crystallization, 257
Resolving agents, 258
Retinoic acid receptor, 14
Retinoic acid receptor α (RARα) gene, 375
Retinoic acid response elements (RAREs), 376
Retinoic acid syndrome (RAS), 382
Retinoid X receptors (RXRs), 376
Retrosynthetic analysis, 67
Reversible phosphorylation, 339
Rheumatoid arthritis (RA), 315, 403
Rhizopus arrhizus, 251
Rho GTPases, 50
RmtA, 347
RmtB, 347
RmtC, 347
RNAPII, 341
Rodent cerebrocortical cultures (RCCs), 323
Root mean square (RMS) deviation, 285
Rotamer libraries, 26
RP algorithm, 69
RP approach, 70
RSVM, 311
Rubicon, 11
Ru-BINAP catalysts, 265
complexes, 265
RyR-associated proteins, 187
Saccharomyces cerevisiae, 247
SANDOCK, 219
SC58635, 116
Scaffold(s), 5, 29
synthetic 284, 285
terephthalamide 296
Scopoletin, 39
SCORE, 27
Scoring, 25
Scoring algorithms, 26
Scoring function(s) (SF), 9, 53, 73, 101, 109
first-principles-derived, 9
knowledge-based, 27, 74, 102
regression-based, 102
Scoring methods, 26
Screening, 3, 5, 64
structure-based 12, 17
SDF-1α/CXCR4 interactions, 492
Secnidazole, 245
Second extracellular loop (ECL2), 304
Selective GABA reuptake inhibitor (SGRI), 272
Semisynthetic drugs, 243
Septic-shock-like syndrome, 177
Sequential-growing approach, 110
Serial analysis of gene expression (SAGE), 484
Serotonin receptor, 46
Serotonin 5-HT1A receptor, 47
Serotonin 5-HT4 receptor, 47
Severe acute respiratory syndrome (SARS), 53
    coronavirus, 53
    3C-like proteinase, 53
SH2 region, 45
Shape scores, 102
Sharpless asymmetric dihydroxylation, 138
Side population cells, 482
Sigmoid function, 115
Signal conduction, 345
Silica matrix delivery approach, 356
Simmons-Smith cyclopropanation, 265
Simulated annealing, 9
Simulation methods, 26
Single-isomer drug, 240
Site-directed mutagenesis, 169, 304
SMAC, 425
Smac bind, 37
Small heterodimer partner (SHP), 473
Small molecule, 82
Small-molecule database, 28, 38
SmithKline Beecham results, 223
SmithKline Beecham’s compound, 218
SMoG, 74
Smonoester, 251
Smooth muscle myosin light chain kinase (smMLCK), 290
Snake venom phosphodiesterase, 173
Sol–gel materials, 356
Sol–gel synthesis parameters, 356
Somatic stem cells, 480
Sources of chiral drugs, 243
Spacer domains, 141, 142
    modification, 14
    synthesis of, 136
Sphere exclusion, 76
SPHGEN, 8
Spontaneous resolution, 256
SPROUT, 94
SPSS10.0, 43
STA-21, 46
π–π stacking, 283
Stat3, 45
    dimerization, 46
    dimerization, 46
    DNA binding activity, 46
Stat3-dependent luciferase activity, 46
Stat3 signaling, blockade of, 45
Stem cell factor (SCF), 492
Stemness genes, 481
Stereoelectronic properties, 93
Stereoisomers, 231
    potency of, 236
Stereoselective, 233
Stereoselective syntheses of (R)-terbutaline, 250
Steric considerations, 282
Stochastic search, 26
Stromal cell-derived factor (SDF)-1α, 304, 308
Structural search, 91
Structure–activity analysis, 177
    of natural bryostatins, 150
Structure–activity relation (SAR), 58, 64, 221, 421
Structure–activity relationship(s), 14, 134, 188
    of cADPR, 195
Structure-based design, 5, 90
Substituted cysteine accessibility method (SCAM), 452
Substructure analysis, 64
Substructure constraints, 82
Sulfoxidation, 268
Sumoylation of PML, 377
Superimposition of minimized conformations, 199
Surface area (SA) term, 10
Surface plasmon resonance (SPR), 119
    assay, 287
Survivin, 424
SYBYL6.9, 47
SYBYL software, 42, 46
Symmetric synthesis, 262
Synphilin-1/SIAH complex, 339
Synthetic chemokine analogues, 317
Synthetic molecules, 292
Synthetically and modularly modified (SMM) chemokines, 317, 318, 320–323, 325
Systemic search, 26
Tabu search, 26
TACE, 43
Tachykinin NK1 receptor, 46, 47
Tachykinin receptors, 285
TAFII250, 339
Tail-vein injection, 470
Tanimoto coefficient, 65, 66
Tannin cytoprotection, 179
Tannins, 174, 181
  in vivo effects of 178
Target-dependent methods, 64, 78
TBHP, 262
TBP-associated complex, 339
T-cell (T)-tropic strains, 303
Tej mutants, 169
Teleocidin B-4, 132
(R)-Terbutaline, 250
Terephthalamide derivatives, 296
Tetraethyl orthosilicate (TEOS), 357
Tetra-galloyl-glucose, 180
Tetramethyl orthosilicate (TMOS), 357
TFIIA, 340, 341
TFIIE, 340, 341
TGB, 272
tgl, 14
Thermal gravimetric analysis (TGA), 359
Thiazolindinedione (TZD), 468, 469
Thin-layer chromatography (TLC), 171
Third extracellular loop (ECL3), 304
Thymocyte aggregation, 156
Thyroid hormone receptor (TR), 376
Tilorone, 172
Time to progression (TTP), 417
(S)-(−)-Timolol, maleate of, 237
TLC-based methods, 171
TNF-family cytokines Fas-ligand (FasL), 413
TNF-family ligands, 411
TNF-family receptors, 409, 412
Toll-like receptor-9 (TLR9), 418
TOPAS program, 94, 97
Topliss tree, 91
Topological state, 68
Torpedo californica AChE, 38
Torsional Search (TS), 12
TR3, 422, 423
TRADD, 412
Traditional drug discovery, 58
TRAIL, 413
Tramadol, 237
  racemate of, 237
Transcriptome analyses, 482
Transferrin, 343
Transient receptor potential (TRP), 441
Transmembrane (TM) helices, 300
TreeDock, 36
Trichosporon sp. (TSL), 245
Trimipramine, 239
  metabolites of, 239
Trisubstituted indanes, 285
Tromboxane A2, synthesis of, 269
Tubulin phosphorylation, 340
Tumor necrosis factor (TNF), 409, 412
  cytokine family, 412
Tumor necrosis factor alpha (TNF-α), 130
Ubiquitin, 337
  activating enzyme, 337
  conjugating enzyme, 337
  proteasome system, 337
  protein ligase, 337
Ubiquitin-dependent proteolysis, 338
Ubiquitin-regulated proteolysis, 336
Ubiquitylation, 337, 338, 339
UBPs, 337, 338
UCHs, 337
UNITY, 51
UNITY4.4, 47
UNITY fingerprint, 65
UNSATP, 114
UVB irradiation, 179
V-10, 214, 215, 367
V-13661, 210
Verapamil, 237
Vinyl ether formation, 136
Vinyl sulfone-based peptides, 51
Viral capsid, 52
Viral macrophage inflammatory protein
  (vMIP)-II, 308
Virtual filtering, 26
Virtual library, 5, 220
Virtual screening, 25, 45, 49, 64, 103, 109, 220
  structure-based, 50
Vitamin D3 receptor (VDR), 376
Vlevofoxacin, 264
vMIP-II, 310, 311
VX-853, 214, 215
W191G cavity, 15
Waldenstrom macroglobulinemia, 401
WDI, 79
World drug index (WDI) database, 78
Xanthen-9-ones, 173
XIAP BIR3 domain, 37
X-linked inhibitor of apoptosis (XIAP), 37
X-ray crystallography, 6, 58, 295
X-SCORE, 46
Yamaguchi esterification, 135
protocol, 143

ZINC, 11
Zinc-β-lactamase, 44
Zinc metalloenzymes, 44
Zn-dependent phosphotriesterase, 45
Zymosan, 177
Figure 1.2  Predicted complexes versus X-ray crystallographic structures that were subsequently determined. See text for full caption.

Figure 1.3  Docking predicted ligands from virtual screening against simple cavity sites. See text for full caption.

Figure 1.4  Predicted versus experimental structures from virtual screening. See text for full caption.
Figure 2.1  Typical receptor-based virtual screening scheme consisting of many steps in repeated cycles. See text for full caption.

Figure 2.3  Structural model of MSK-039 docked at the EGFR kinase active site (only selected residues are shown). Dashed lines represent hydrogen bonds between ligand and active site residues.

Figure 2.4  Structural model of a representative hit docked at the Chk-1 kinase active site (only selected residues are shown). Dashed lines represent hydrogen bonds between ligand and active site residues.
**Figure 3.1** The chemical structures of a number of 5-HT reuptake inhibitors showing clear similarities in structure, illustrating the basis for the similarity and substructure methods described in the text. The common substructures are highlighted.

**Figure 3.2** The DSA Visualizer for exploring substructure fragments in a set of compounds active for a particular target. See text for full caption.
Figure 3.3  Example of a hypothetical recursive-partitioning tree. See text for full caption.

Figure 3.5  An example of molecular docking using the Glide program. See text for full caption.
Selection of 16 structurally different 5α-reductase inhibitors

1. Different conformations computed
2. Pharmacophore generated by HypoGen

Pharmacophore model

Molecules in training set mapped to pharmacophore

Screening NCI database

Figure 4.5 Isoflavone derivatives as 5α-reductase inhibitors from pharmacophore screening.
Figure 5.3  Structure of SC58635 and fragment division for COX-2 inhibitors.

Figure 5.5  Common structural characteristic of PPARγ agonists and possibly structural fragment division.

Figure 5.6  Three-dimensional structure of hCypA with designed inhibitors displayed. The CypA surface was colored by electrostatic potential.

Figure 6.4  The preferred pharmacophoric correlation of 1,2-diacyl-sn-glycerol, the phorbol esters, and bryostatin 1.
Figure 7.1  (a) In response to mild DNA damage, PARP-1 binds to DNA and catalyzes the formation of poly(ADP-ribose) polymers onto protein acceptors including itself. See text for full caption.

Figure 7.2  Human PARG110 has a two-domain structure, with both a regulatory domain (1–417) and a catalytic domain (418–976). See text for full caption.
**Figure 8.4** Effect of the cADPR mimic cIDPRE on \( \text{Ca}^{2+} \) signaling in T cells. See text for full caption.

**Figure 8.6** cIDP-DE-mediated local \( \text{Ca}^{2+} \) signaling and global \( \text{Ca}^{2+} \) waves in intact mouse cardiac myocytes. See text for full caption.
Figure 8.7  Superimposition of minimized conformations of cADPR and cIDP-DE. The conformations of cADPR (yellow) and cIDP-DE (pink) were minimized using the density functional theory (DFT) method. (A) Superimposition based on nucleobases; (B) superimposition based on diphosphates; (C) superimposition based on whole molecules.

Figure 9.5  Complex structure of FKBP12-FK506; the protein is displayed in cartoon model, and FK506 is shown in stick model. See text for full caption.

Figure 9.7  Complex structure of Holt’s compound and FKBP12. See text for full caption.

Figure 9.8  Ludi-based ligand design pipeline.
**Figure 9.9** Site points in FKBP. See text for full caption.

**Figure 9.10** Complex of compound 308 and FKBP12. See text for full caption.

**Figure 10.4** Crixivan bound to HIV protease (taken from www.biotech100.com).

**Figure 10.11** The process of normal release and reuptake of GABA.
Figure 11.1  (A) An \( \alpha \)-helix that is stabilized by internal hydrogen bonding. (B) Ribbon diagram showing rise per turn. (C) Amide backbone showing dihedral angles.

Figure 11.2  Some examples of helical stabilization by covalent linkages: disulfide (1), lactam (2), azobenzene (3), olefin (4).

Figure 11.3  Some examples of helical stabilization by non-covalent interaction: \( \pi-\pi \) interaction (5), salt-bridge (6), cation-\( \pi \) interaction (7).

Figure 11.4  A 1,6-disubstituted indane molecule, 8, that mimics of the \( i \) and \( i + 1 \) side chains (shown in yellow) of an \( \alpha \)-helix.

Figure 11.6  (A) X-ray crystal structure of the HDM2/p53 complex. (B) \( \beta \)-hairpin inhibitor that disrupts the HDM2-p53 interaction.

Figure 11.7  An \( \alpha \)-helix backbone can be replaced by a synthetic scaffold that projects the functionalities in a similar fashion.
Figure 11.9 X-ray crystal structure of the CaM-binding region of smMLCK (residues 796-815, shown in red) in a complex with CaM. Side chains of the critical residues (Trp800, Thr803 and Val807) are shown in stick representation.

Figure 11.11 X-ray crystal structure of the gp41 fusion active core C34 (a.a. 628-661, shown in yellow helical tubes representation) bind to N36 (a.a. 546-581) by inserting the side chains of Trp628, Trp631 and Ile635 residues into a hydrophobic cleft on the protein surface.

Figure 11.12 Results of the $^{15}$N-HSQC experiments of 13 binding to Bcl-xL. See text for full caption.

Figure 11.13 (A) X-ray crystal structure of an oligoamide foldamer 14. (B) Overlay of a poly-alanine $\alpha$-helix and 14, showing good agreement of the side chains.
Figure 11.14  (A) Terephthalamide helical mimetics. (B) The superimposition of 15 on the $i$, $i + 4$, and $i + 7$ positions of a poly-alanine α-helix (RMS deviation = 1.03 Å). (C) The superimposition of terphenyl 13 and terephthalamide 15 (RMS deviation = 0.34 Å). The functional side chains are highlighted.

Figure 11.16  Superimposition of the top-ranked result of the molecular-docking studies of 17 and the Bak peptide. The residues that showed significant and moderate chemical shift changes upon the addition of 17 were shown in magenta and yellow, respectively.

Figure 12.4  Distinct functional sites for HIV-1 gp120 and SDF-1α on a hypothetical structural model of CXCR4. See text for full caption.
Figure 12.5 General design concept for SMM-chemokines. The N-terminal (1–10) sequence module of vMIP-II was modified by introducing unnatural D-amino acids or specific sequences grafted from other chemokines to chemically engineer receptor selectivity and/or signaling property.

Figure 12.7 Distinct functional sites for SDF-1α and SMM-chemokines highlighted on a hypothetical structural model of CXCR4. See text for Full caption.
Figure 13.3  Protein phosphorylation regulating cellular processes.

Figure 17.2  Structure of Bcl-2 and associated BH3 peptide. The 3D structure of Bcl-X_L is depicted (blue) with Bak BH3 peptide bound (red).
Figure 20.2 NSCs possess an inherent mechanism for rescuing dysfunctional neurons. Shown here is evidence from the effects of NSCs in the restoration of mesencephalic dopaminergic function. [Modified from Ourednik et al. (2002)]. See text for full caption.

Figure 17.3 Structure of BIR domain and inhibitory SMAC peptide. BIR domain of XIAP is presented in space-filling model, while the SMAC tetra-peptide is presented in stick mode. The left structure gives the entire Bir3 domain, while the right structure is a subdomain of Bir3. The structure is derived from PDB #1G3F.