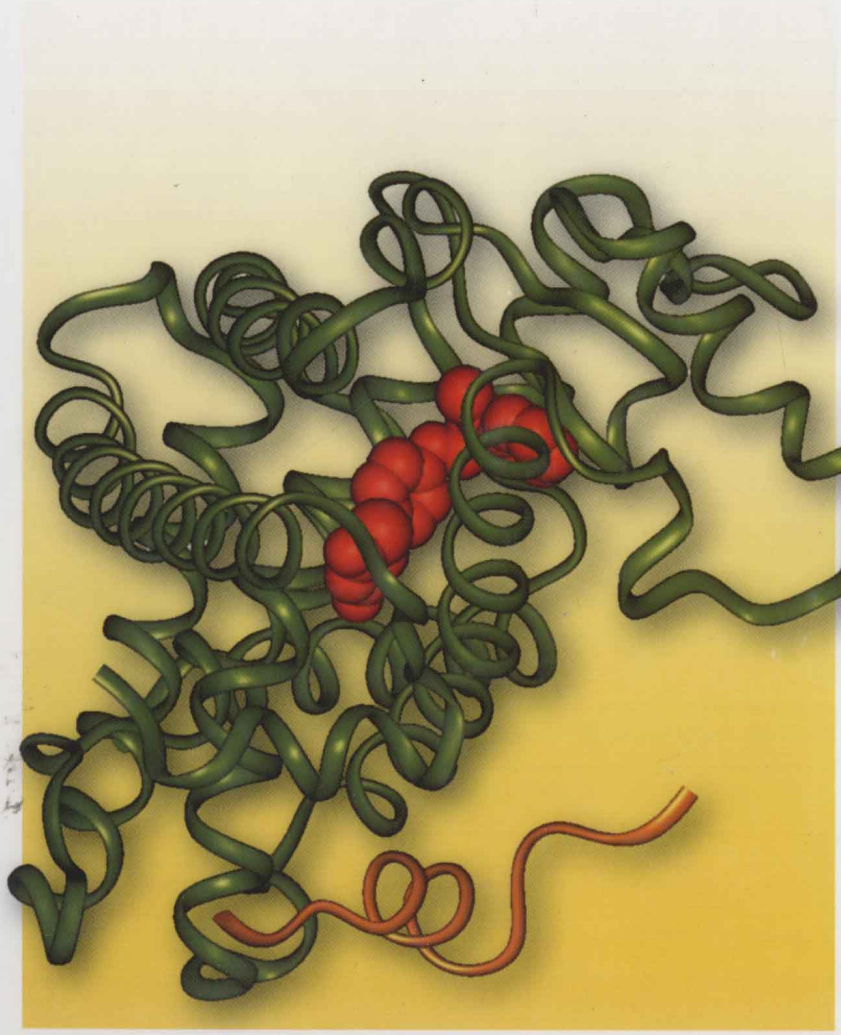
Methods and Principles in Medicinal Chemistry

Edited by
R. E. Babine and S. S. Abdel-Meguid

WILEY-VCH

Protein Crystallography in Drug Discovery



Volume 20

Series Editors:
R. Mannhold,
H. Kubinyi,
G. Folkers

Series Editors:

Prof. Dr. Raimund Mannhold
Biomedical Research Center
Molecular Drug Research Group
Heinrich-Heine-Universität
Universitätsstraße 1
40225 Düsseldorf
Germany
raimund.mannhold@uni-duesseldorf.de

Prof. Dr. Hugo Kubinyi
BASF AG Ludwigshafen
c/o Donnersbergstraße 9
67256 Weisenheim am Sand
Germany
kubinyi@t-online.de

Prof. Dr. Gerd Folkers
Department of Applied Biosciences
ETH Zürich
Winterthurerstr. 190
8057 Zürich
Switzerland
folkers@pharma.anbi.ethz.ch

Volume Editors:

Dr. Robert E. Babine
Suntory Pharmaceutical Research Laboratories
One Kendall Square
Cambridge, MA 02139
USA
robert.babine@sprlus.com

Dr. Sherin S. Abdel-Meguid
Suntory Pharmaceutical Research Laboratories
One Kendall Square
Cambridge, MA 02139
USA
sherin.s.abdel-meduid@sprlus.com

The cover illustration shows the antidiabetic compound rosiglitazone bound to the nuclear receptor PPAR-γ (see chapter 1).

Library of Congress Card No.: applied for

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Bibliographic information published by Die Deutsche Bibliothek
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at <http://dnb.ddb.de>

Printed in the Federal Republic of Germany
Printed on acid-free paper

Composition K+V Fotosatz GmbH, Beerfelden
Printing betz-druck gmbh, Darmstadt
Bookbinding Litges & Dopf Buchbinderei GmbH, Heppenheim

ISBN 3-527-30678-1

Protein Crystallography
in Drug Discovery
Edited by
R. E. Babine and S. S. Abdel-Meguid

Methods and Principles in Medicinal Chemistry

Recently published Volumes:

Vol. 12 Drug Targeting
G. Molema, D. K. F. Meijer (eds.)
2001, ISBN 3-527-29989-0

Vol. 13 Pharmacokinetics and Metabolism in Drug Design
D. Smith, D. Walker, H. van de Waterbeemd
2001, ISBN 3-527-30197-6

Vol. 14 Bioinformatics – From Genomes to Drugs
T. Lengauer (ed.)
2001, ISBN 3-527-29988-2

Vol. 15 Drug-Membrane Interactions
J. K. Seydel, M Wiese
2002, ISBN 3-527-30427-4

Vol. 16 BioNMR in Drug Research
O. Zerbe (ed.)
2002, ISBN 3-527-30465-7

Vol. 17 Quantum Medicinal Chemistry
P. Carloni, F. Alber (eds.)
2003, ISBN 3-527-30456-8

Vol. 18 Drug Bioavailability
H. van de Waterbeemd, H. Lennernäs, P. Artursson (eds.)
2003, ISBN 3-527-30438-X

Vol. 19 Protein-Ligand Interactions
H.-J. Böhm, G. Schneider (eds.)
2003, ISBN 3-527-30521-1

Preface

Among all new technologies in drug research, structure-based ligand design is one of the most powerful approaches. The drugs Captopril, Dorzolamide and Zanamivir, to mention only some prominent examples, resulted from rational design, based on the knowledge and analysis of protein 3D structures. The discovery of some other drugs, e.g. the more recent HIV protease inhibitors Nelfinavir and Amprenavir, was at least supported by protein crystallography studies. Many other drug candidates that resulted from structure-based design are in clinical development.

Whereas some early attempts of structure-based design failed due to inappropriate physicochemical and pharmacokinetic properties of the ligands, modellers are nowadays aware of the pitfalls in ligand design. Large, greasy ligands are avoided, as well as too polar compounds. According to favorable lead and drug properties, defined e.g. by the Lipinski rule of five, ligand design focuses on compounds with relatively low molecular weight, an intermediate lipophilicity range, and a limited number of hydrogen bond donors and acceptors.

In 1997, Klaus Gubernator and Hans-Joachim Böhm edited a volume on structure-based ligand design in this book series. A comprehensive review by Robert Babine, one of the editors of the current book, and Steven Bender, also published in 1997, discussed the design of aspartic, serine, cysteine, and metalloprotease inhibitors, and of immunosuppressants. The present book deals with some other families of important biological targets, e.g. nuclear receptors and kinases. In addition, several other attractive drug targets are reviewed. A special topic, the design of orthogonal protein-ligand pairs, will become important in personalized medicine. The book closes with chapters on recent progress in technologies: protein engineering to promote crystallization, micro-crystallization, and high-throughput crystallography. With its broad perspective, the book provides a state-of-the-art overview on important results and techniques that are relevant for protein 3D structure-based drug design.

The series editors are grateful to Robert Babine and Sherin Abdel-Meguid for their engaged work and to Frank Weinreich, Wiley-VCH, for ongoing support during the preparation of the book. We expect that volume 20 of "Methods and Principles in Medicinal Chemistry" will be another highlight of this successful series, which started only ten years ago.

September 2003

Raimund Mannhold, Düsseldorf
Hugo Kubinyi, Weisenheim am Sand
Gerd Folkers, Zürich

A Personal Foreword

When approached to initiate this book project the initial thoughts were to remember a review published in 1997 [1]. The first memories of that endeavor were quite painful, so the first instinct was to politely turn down the kind offer to initiate this project. On second reflection, researching and preparing the 1997 review was an educational and rewarding experience. Therefore, with the help of a co-editor, the invitation to initiate and complete this endeavor was accepted.

The objective of this book is to provide a forward looking overview of the use of protein crystallography in drug discovery. It has been organized so that the early chapters review and describe some mature and emerging topics that would fall under the 'traditional structure-based design' umbrella, the middle chapters provide focused accounts of specific works, and the final chapters delve into new and fertile areas of research. This book does not attempt to be comprehensive; thus the final lineup of chapters was a compromise between the interests of the editors and the willingness of the authors to contribute a chapter.

The first two chapters review nuclear hormone receptors and protein kinases. Both of these chapters provide an overview of the topic and shed some insight into how small molecule ligands can achieve selectivity between related protein targets. The next two chapters review topics that begin to 'push the limit' on the size of the complexes that can be used in drug design. Both the proteosome and the ribosome are very large biological macromolecules that are the targets for drug discovery. In the ribosome chapter, an important point is made regarding what conclusions are warranted, or not warranted, based upon the resolution of the x-ray diffraction data. The ribosome chapter also introduces the topic and challenges of antibiotic resistance in drug discovery. The next two chapters provide accounts of detailed structure-based design studies aimed at obtaining inhibitors of both cathepsin K and Cdk4. The cathepsin K chapter provides a nice account of the iterative structure-based design process. This work is especially notable for the use of an unexpected crystallographic result to move a project in a novel direction. The Cdk4 chapter also does an excellent job of introducing many computational methods that are used in drug discovery. For those readers interested in protein-based virtual screening of chemical databases, we also recommend the work by Bissantz *et al.* [2] in which they evaluate different docking/scoring combinations. Chapter 7 describes applications of the protease inhibitor ecotin, particularly its

use as a tool to obtain crystals of serine proteases and to study the interactions between serine proteases and their substrates. Chapter 8 reviews work on 'orthogonal ligand-receptor pairs' and the impact of crystallography on this area of research. While 'traditional structure-based design' uses the structure of a protein-ligand complex as a tool to design modified ligands, the work on 'orthogonal ligand-receptor pairs' uses the structure of the complex as a tool to design both modified proteins and modified ligands. This work has applications for deconvoluting some cellular processes and also provides useful tools in the area of chemical genetics. Protein/ligand pairs may also have future applications in gene therapy. Chapters 7 and 8 present examples that use crystallography to design mutant proteins for additional structural studies. Chapter 9 discusses the use of mutant proteins as an entry into obtaining high-resolution crystal structures. In numerous past examples, when a human protein proved difficult to crystallize, a protein from another species such as mouse, rat, or chicken was often used as a surrogate protein. A potential problem with this approach is that the active sites of the human and the surrogate enzyme might be different. The chapter on "Engineering Proteins to Promote Crystallization" reviews examples where the surface of a human protein is modified in a location remote from the active site to produce a mutant human protein that can give useful crystals. The next chapter discusses "High-throughput crystallography" which is an area of intense interest [3, 4]. This chapter provides a clear overview of the field, and has an informative section on crystallography in lead discovery. The final chapter describes miniaturization of crystallization utilizing the emergent technologies of microfluidics. This technology allows crystallization experiments to be performed on the nanoliter scale, thus conserving a very valuable resource in crystallography, the protein. Many of the technologies described in the final three chapters should impact crystallization and structure determination of the many new proteins identified in the human and other genomes.

We expect that this book will be a useful reference to practitioners of structure-based design. In addition, we hope that the technologies discussed in the later chapters will help researchers solve new problems in the next generation of structure-based design problems.

Cambridge, MA, April 2003

Robert E. Babine
Sherin S. Abdel-Meguid

1 Babine, R.E. and S.L. Bender, Molecular Recognition of Protein-Ligand Complexes: Applications to Drug Design. *Chem. Rev.* **1997**, *97*(5), 1359–1472.

2 Bissantz, C., G. Folkers, and D. Rognan, Protein-based virtual screening of chemical databases. 1. Evaluation of different docking/scoring combinations. *J. Med. Chem.* **2000**, *43*(25), 4759–4767.

3 Mountain, V., Innovation – Astex, Structural Genomix, and Syrrx. *Chem. Biol.* **2003**, *10*, 95–98.

4 Bertini, I., Structural genomics. *Acc. Chem. Res.* **2003**, *36*(3), 155.

List of Contributors

Jerry L. Adams
GlaxoSmithKline Pharmaceuticals
Medicinal Chemistry
1250 South Collegeville Road
Collegeville, PA 19426
USA

Robert E. Babine
Suntory Pharmaceutical Research
Laboratories LLC
Structural and Computational
Chemistry
One Kendall Square, Bldg 700
Cambridge MA 02139
USA

James M. Berger
University of California at Berkeley
Department of Molecular
and Cell Biology
229 Stanley Hall
Berkeley, CA 94720-3206
USA

Maxwell D. Cummings
Computational Chemistry
3-Dimensional Pharmaceuticals
665 Stockton Drive, Suite 104
Exton, PA 19341
USA

Donald F. Doyle
Georgia Institute of Technology
School of Chemistry and Biochemistry
770 State Street
Atlanta, GA 30332-0400
USA

Robert J. Fletterick
Departments of Biochemistry
and Biophysics and Pharmaceutical
Chemistry
University of California
at San Francisco
600 16th Street, Box 2240
San Francisco, CA 94143-2240
USA

Carl L. Hansen
California Institute of Technology
Department of Applied Physics
MS 128-95
Pasadena, CA 91125
USA

Jeffrey L. Hansen
Department of Molecular Biophysics
and Biochemistry
Yale University
266 Whitney Avenue, 415 Bass Center
New Haven, CT 06520
USA

Teruki Honma
Banyu Tsukuba Research Institute
in collaboration with Merck Research
Laboratories
Okubo-3
Tsukuba 300-2611
Ibaraki
Japan

Harren Jhoti
Astex Technology Ltd.
250 Cambridge Science Park
Milton Road
Cambridge, CB4 0WE
United Kingdom

Lei Jin
Suntory Pharmaceutical Research
Laboratories LLC
One Kendall Square, Bldg 700
Cambridge MA 02139
USA

Tsunehiro Mizushima
Japan Science and Technology
Corporation
Structure and Function
of Biomolecules
3-18-22 Honkomagome, Bunkyo-ku
Tokyo, 113-0021
Japan

Stephen R. Quake
California Institute of Technology
Department of Applied Physics
MS 128-95
Pasadena, CA 91125
USA

Lauren J. Schwimmer
Georgia Institute of Technology
School of Chemistry and Biochemistry
770 State Street
Atlanta, GA 30332-0400
USA

Kyle Self
Fluidigm Corporation
7100 Shoreline Court South
San Francisco, CA 94080
USA

Lisa M. Shewchuk
GlaxoSmithKline Pharmaceuticals
Discovery Research
5 Moore Drive
Research Triangle Park, NC 27709
USA

Morten Sommer
California Institute of Technology
Department of Applied Physics
MS 128-95
Pasadena, CA 91125
USA

Tomitake Tsukihara
Institute for Protein Research
Osaka University
3-2, Yamadaoka, Suita
Osaka 565-0871
Japan

James Veal
Serenex, Inc.
Informatics and Computational
Chemistry
323 Foster St.
Durham, NC 27701
USA

Daniel F. Veber
290 Batleson Rd.
Ambler, PA 19002
USA

Sandra M. Waugh
University of California
at San Francisco
Graduate Group in Biophysics
San Francisco, CA 94143-2240
USA

Contents

1
Molecular Recognition of Nuclear Hormone Receptor-Ligand Complexes

Robert E. Babine

1.1
Introduction

1.1.1
Nuclear Hormone Receptors: Ligand Binding Domains

Nuclear hormone receptors (NHR) are multidomain proteins that function as transcription factors. They contain a central DNA binding domain (DBD) responsible for targeting the receptor to highly specific DNA sequences comprising a response element. The DBD is surrounded by two activation domains: the activation function 1 (AF-1) domain that resides at the N-terminus and the activation function 2 (AF-2) domain that resides at the C-terminal ligand-binding domain (LBD) [1, 2]. NHRs for which no natural ligand is known are referred to as orphan nuclear hormone receptors. Regulation of gene transcription by nuclear receptors requires the recruitment of proteins characterized as co-regulators, with ligand-dependent exchange of co-repressors for co-activators serving as the basic mechanism for switching gene repression to activation [2]. Upon activation by a small molecule ligand, or hormone, NHRs dimerize with another ligated NHR and recruit co-activators to turn on target genes. The binding of a ligand thus can act as a switch between gene activation and gene repression. This biological property has made the ligand binding domains of nuclear hormone receptors important drug targets.

Our understanding at the molecular level of how nuclear receptor ligands exert their effects has been dramatically enhanced by the elucidation of the crystal structures of the apo- and/or ligand-bound LBDs of several nuclear receptors (Tab. 1.1). These structures have revealed a common fold among LBDs, consisting of an antiparallel α-helical sandwich of 11–13 helices. Conventional nomenclature refers to these helices as helix-1, and helices-3 through 12. The region between helices 1 and 3 is variable and may contain zero, one or two helices. The helices fold to form a hydrophobic cavity into which the ligand can bind. The position of helix-12 relative to the other helices is dependent upon the presence or absence of a ligand and the nature of that ligand. The binding site that comprises the AF-2 domain is determined by the position of the C-terminal α-helix (helix-12).

Protein Crystallography in Drug Discovery
Edited by R. E. Babine and S. S. Abdel-Meguid

ISBN: 3-527-30678-1

Tab. 1.1 PDB codes of all nuclear hormone receptor-ligand-binding domains in the protein data bank as of December 2002.

Protein	*Apo structure*	*Ligated structures*	*Complex dimer structure*
ER-α		1A52, 1ERE, 1ERR, 1QKT, 1QKU, 3ERD, 3ERT, 1L2I, 1GWQ, 1GWR	
ER-β		1QKM, 1QKN, 1HJ1, 1L2J	
AR		1I37, 1I38, 1E3G	
GR		1M2Z, 1NHZ	
TR-β		1BSX	
RAR-α			1DKF
RAR-γ		1EXA, 1EXX, 1FCX, 1FCY, 1FCZ, 2LBD, 3LBD, 4LBD, 1FD0	
RXR-α	1G1U, 1LBD	1FBY, 1G5Y, 1MV9, 1MVC, 1MZN	1DKF, 1FM6, 1FM9, 1K74
RXR-β		1H9U	
PR		1A28, 1E3K	
VDR		1DB1, 1IE8, 1IE9	
PPAR-α		1I7G, 1K7L, 1KKQ	
PPAR-γ	1PRG, 3PRG	1I7I, 2PRG, 4PRG, 1KNU	1FM6, 1FM9, 1K74
PPAR-δ	2GWX	1GWX, 3GWX	
PXR	1ILG	1ILH	
ROR-β		1K4W	
ROR-α		1N83	
HNF4-γ		1LV2	
ERR3	1KV6		

ER Estrogen Receptor, AR Androgen Receptor, GR Glucocorticoid Receptor, TR Thyroid Receptor, RAR Retinoic Acid Receptor, RXR Retinoie X Receptor, PR Progesterone Receptor, VDR Vitamin D Receptor, PPAR Peroxisome Proliferator-Activated Receptor, PXR Pregnane X Receptor, ROR Retinoic Acid-Related Orphan Receptor, NHF-4 Hepatocyte Nuclear Factor 4, ERR Estrogen-Related Receptor

1.1.2
Dimerization and Interactions with Co-activators and Co-repressors

The Glaxo group was the first to report the structure of a heterodimeric complex between two activated NHRs (peroxisome proliferator-activated receptor-γ) PPAR-γ and (retinoid X receptor-α) RXR-α. The structure (PDB entry: 1FM6) of this heterodimer complex contains six components: the two receptor LBDs, their two respective ligands, and two peptides derived from the steroid receptor co-activator-1 (SRC-1). These peptides contain a conserved LxxLL motif that is present in this class of co-activators. The complex is butterfly shaped, with both LBDs adopting the conserved "helical sandwich" fold previously reported for other ligand-bound nuclear receptors. PPAR-γ contains 13 α-helices and four short β-strands, while RXR-α is composed of 11 α-helices and two short β-strands [3]. This complex is illustrated in Fig. 1.1.