

**Protein Crystallography
in Drug Discovery**

Edited by

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

Protein Crystallography in Drug Discovery
Edited by R. E. Babine and S. S. Abdel-Meguid
Copyright © 2004 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 3-527-30678-1

**Pharmacyclics LLC
Exhibit 2121
BeiGene v. Pharmacyclics
PGR2024-00009
EX2121.0001**

Exhibit #

Veal 02

11/22/24 BK

exhibitsticker.com

Methods and Principles in Medicinal Chemistry

Edited by

R. Mannhold

H. Kubinyi

G. Folkers

Editorial Board

H.-D. Höltje, H. Timmerman, J. Vacca, H. van de Waterbeemd, T. Wieland

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

Protein Crystallography in Drug Discovery

Edited by

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



WILEY-VCH Verlag GmbH & Co. KGaA

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

This book was carefully produced. Nevertheless, authors, editors and publisher do not warrant the information contained therein to be free of errors. Readers are advised to keep in mind that statements, data, illustrations, procedural details or other items may inadvertently be inaccurate.

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>>

© 2004 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

All rights reserved (including those of translation in other languages). No part of this book may be reproduced in any form – by photoprinting, microfilm, or any other means – nor transmitted or translated into machine language without written permission from the publishers. Registered names, trademarks, etc. used in this book, even when not specifically marked as such, are not to be considered unprotected by law.

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

Contents

Preface *XI*

A Personal Foreword *XIII*

List of Contributors *XV*

1	Molecular Recognition of Nuclear Hormone Receptor-Ligand Complexes	1
	<i>Robert E. Babine</i>	
1.1	Introduction	1
1.1.1	Nuclear Hormone Receptors: Ligand Binding Domains	1
1.1.2	Dimerization and Interactions with Co-activators and Co-repressors	2
1.2	Steroid Receptors	5
1.2.1	The Role of the Ligand	5
1.2.1.1	Estradiol Estrogen Receptor Complex	5
1.2.1.2	Other Estrogen Receptor Agonists Complexes	8
1.2.1.3	Estrogen Receptor Antagonists Complexes	9
1.2.1.4	Genistein – An ER- β Partial Agonist	12
1.2.1.5	<i>R,R-5,11-cis</i> -Diethyl-5,6,11,12-tetrahydrochrysene-2,8-diol: An ER- α Agonist and ER- β Antagonist	14
1.2.2	Structural Basis for Agonism and Antagonism and Partial Agonism: The Role of the Ligand	15
1.2.3	Progesterone/Progesterone Receptor Complex	16
1.2.4	Androgen Receptor Complexes	17
1.2.5	Glucocorticoid Receptor	18
1.2.6	Steroid Ligand Selectivity	19
1.3	The Vitamin D Receptor-Ligand Complexes	22
1.4	The Retinoic Acid Receptors RAR and RXR	23
1.4.1	Introduction	23
1.4.2	RAR- γ and RXR- α Retinoid Complexes	25
1.4.3	Selectivity of RAR Ligands and RAR Isotypes	27
1.4.4	RXR Complexes with Unnatural Ligands	33
1.5	PPAR: Isotype-Selective Ligands	34
1.6	Summary	41
1.7	Acknowledgements	43

1.8	References	43
2	Kinases	47
	<i>Jerry L. Adams, James Veal, and Lisa Shewchuk</i>	
2.1	Introduction	47
2.2	Structure and Function	48
2.2.1	Tertiary Structure	48
2.2.2	Catalysis and Substrate Binding	49
2.2.3	Regulation and Conformational Flexibility	50
2.2.3.1	Activation Loop Conformation	50
2.2.3.2	Glycine Rich Loop	51
2.2.3.3	C-Helix Orientation	51
2.2.3.4	Lobe Orientation	53
2.2.3.5	Solvent Channel	53
2.3	Crystallization	54
2.3.1	Defining the Construct	54
2.3.2	Mutagenesis	55
2.3.3	Phosphorylation	55
2.4	Inhibitor Design	56
2.4.1	Binding in ATP Cleft	56
2.4.1.1	ATP Binding Sites	57
2.4.1.2	Gatekeeper-Dependent Binding Pocket	60
2.4.1.3	Lipophilic Plug	61
2.4.1.4	Polar Surface Site	61
2.4.2	Conformational Considerations	63
2.4.2.1	Inhibitor-Induced Binding	63
2.4.2.2	What is the Most Appropriate Enzyme Form for Crystallography?	64
2.4.2.3	Homology Models and Surrogate Kinases	66
2.4.3	Paradigms for Kinase Drug Discovery	68
2.4.3.1	High Throughput Screening	68
2.4.3.2	Structure-Based Design	69
2.4.3.3	Mechanism-Based and Ligand Mimetic Design	72
2.4.3.4	Computational Chemistry and Virtual Screening	72
2.4.4	Selectivity	73
2.5	Conclusion	75
2.6	References	76
3	The Proteasome as a Drug Target	83
	<i>Tsunehiro Mizushima and Tomitake Tsukihara</i>	
3.1	Introduction	83
3.2	The Ubiquitin-Proteasome System	83
3.2.1	Role of the Ubiquitin-Proteasome System	84
3.2.2	26S Proteasome	85
3.2.3	20S Proteasome	85
3.3	Structure of the 20S Proteasome	86

3.3.1	Active Sites of Eukaryotic 20S Proteasomes	88
3.3.2	Novel Ntn-Hydrolase Active Site of the $\beta 7$ Subunit	88
3.3.3	Predicted Structure of Immunoproteasome and Substrate Specificities	89
3.4	Proteasome Inhibitors	93
3.4.1	Structure of Proteasome Inhibitor Complexes	95
3.5	Conclusions	96
3.6	Acknowledgements	96
3.7	References	96
4	Antibiotics and the Ribosome	99
	<i>Jeffrey L. Hansen</i>	
4.1	Introduction	99
4.2	The Ribosome	99
4.2.1	Introduction	99
4.2.2	Binding of tRNA	100
4.2.3	Peptidyl Transferase Activity	101
4.2.4	Structure of the Ribosome	102
4.3	Antibiotics	103
4.3.1	Introduction	103
4.3.2	Antibiotics that Bind to the 50S Subunit	104
4.3.3	MLS _B Antibiotics	105
4.3.4	Macrolides	107
4.3.4.1	Macrolides, 15- and 16-Membered	107
4.3.4.2	Binding Interactions Between the Lactone Ring and the Ribosome	108
4.3.4.3	Sugar Interactions with the Ribosome	110
4.3.4.4	A Covalent Bond	111
4.3.4.5	Macrolides, 14-Membered	112
4.3.4.6	Rational Drug Design of Macrolides	112
4.3.5	Lincosamides	114
4.3.6	Streptogramins	114
4.3.7	Chloramphenicol	115
4.3.8	Nucleoside Analogue Antibiotics	117
4.3.8.1	Puromycin	117
4.3.8.2	Aminoacyl-4-aminohexosyl-cytosine Antibiotics	118
4.3.9	Other Antibiotics that Bind to the 50S Subunit	119
4.3.9.1	Sparsomycin	119
4.3.9.2	Anisomycin	120
4.4	Prospects for Rational Drug Design of Antibiotics that Bind to the Ribosome	120
4.5	Acknowledgements	122
4.6	References	122

5	Structure-Based Design of Cathepsin K Inhibitors	127
	<i>Daniel F. Veber and Maxwell D. Cummings</i>	
5.1	Introduction	127
5.2	Background and Issues to be Addressed Using Protein Structure	128
5.3	Cysteine Protease Inhibitors: Historical Perspective	131
5.4	Diaminoketone-Based Inhibitors	133
5.5	Cyclic Diaminoketone Inhibitors	136
5.6	Alkoxyethyl and Thiomethyl Dipeptidyl Ketone-Based Inhibitors	140
5.7	Diacylcarbohydrazides	142
5.8	Conclusions	143
5.9	References	145
6	Structure-Based Design of Potent and Selective Cdk4 Inhibitors	147
	<i>Teruki Honma</i>	
6.1	Introduction	147
6.2	Homology Modeling of Cdk4	148
6.3	Analysis of the ATP Binding Pocket	151
6.4	Strategies for Structure-Based Lead Identification: Virtual Screening of Known Compounds and de Novo Design	152
6.4.1	New de Novo Design Strategies	153
6.4.2	Evaluation of Chemical Availability by SEEDS	155
6.5	Structure-Based Generation of a New Class of Potent Cdk4 Inhibitors	157
6.5.1	Identification of New Scaffold Candidates Using LEGEND and SEEDS	157
6.5.2	From the Identified Scaffold to Lead Compounds	159
6.5.3	Confirmation of the Binding Mode	159
6.6	Structure-Based Design of Cdk4 Selective Inhibitors	160
6.6.1	Identification of Cdk4-Specific Amino Acid Residues	161
6.6.2	Library Design Based on the Locations of Cdk4-Specific Amino Acid Residues	163
6.7	Conclusion	167
6.8	Acknowledgements	167
6.9	References	168
7	Crystallization and Analysis of Serine Proteases with Ecotin	171
	<i>Sandra M. Waugh and Robert J. Fletterick</i>	
7.1	What is Ecotin?	171
7.2	Methods	173
7.2.1	Expression of Wild Type Ecotin	173
7.2.2	Purification of Wild Type Ecotin	174
7.2.3	Crystallization of Ecotin and Protease Complexes	175
7.3	Representative Examples of Ecotin and Protease Structures	175
7.3.1	Ecotin Defines the S7 Through S4' Subsites of Collagenase	176

7.3.2	Ecotin as a Tight Binding Substrate	177
7.3.3	Ecotin Defines Regions Distal to the Factor Xa Protease Domain	178
7.4	Crystallization and Structure Determination of E ₂ P ₂ Complexes	180
7.5	Conclusion	183
7.6	References	184
8	X-ray Crystallography in the Development of Orthogonal Ligand–Receptor Pairs	187
	<i>Donald F. Doyle and Lauren J. Schwimmer</i>	
8.1	Introduction	187
8.2	Applications of OLRPs	188
8.2.1	Applications in Basic Research	188
8.2.2	Applications in Applied Research	189
8.3	Early Work	189
8.4	Structure-Guided Mutagenesis and OLRPs	190
8.4.1	Binding Pocket Mutations	191
8.4.1.1	Nuclear Receptors	191
8.4.1.2	Estrogen Receptor OLRPs	193
8.4.1.3	Retinoic Acid Receptor OLRPs	195
8.4.1.4	Retinoid X Receptor OLRPs	195
8.4.2	Distant Mutations	197
8.5	Other Examples of OLRPs	200
8.5.1	Chemical Inducers of Dimerization to Control Transcription	200
8.5.2	OLRPs and ATP Analogues	202
8.6	Summary	204
8.7	References	206
9	Engineering Proteins to Promote Crystallization	209
	<i>Lei Jin and Robert E. Babine</i>	
9.1	Introduction	209
9.2	Removing Protein Heterogeneity by Truncation	210
9.3	Removing Protein Heterogeneity by Point Mutation	211
9.4	Improving Crystal Packing by Point Mutation	212
9.5	Acknowledgements	214
9.6	References	214
10	High-throughput Crystallography	217
	<i>Harren Jhoti</i>	
10.1	Introduction	217
10.2	Technological Advances	218
10.2.1	Clone to Crystal	218
10.2.1.1	Protein Production	218
10.2.1.2	Crystallization	219
10.2.2	Crystal to Structure	220
10.2.3	Progress in Structural Genomics	221

x | Contents

10.3	High-throughput Crystallography in Lead Discovery	223
10.3.1	Protein-Ligand Crystal Structures	224
10.4	Fragment-Based Lead Discovery	225
10.4.1	Fragment-Based Lead Discovery Using X-ray Crystallography	226
10.4.2	Structure-Based Optimization of Fragment Hits	229
10.5	Conclusions	230
10.6	Acknowledgements	231
10.7	References	231
11	Micro-Crystallization	235
	<i>Carl L. Hansen, Morten Sommer, Kyle Self, James M. Berger, and Stephen R. Quake</i>	
11.1	Introduction	235
11.2	Microfluidics – Method and Design	237
11.3	Utility of Microfluidics for Crystallization	242
11.4	References	253

Subject Index	257
----------------------	------------

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 proteasome 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

2

Kinases

JERRY L. ADAMS, JAMES VEAL, and LISA SHEWCHUK

2.1

Introduction

The regulation of glycogen phosphorylase function by reversible phosphorylation was first described in the mid-1950s. However, the potential importance of protein kinases was not fully appreciated until the advent of DNA cloning and sequencing in the mid-1970s. The initial estimate of 1000–2000 human protein kinases made in 1987 has now been refined using the full human genome to 518 putative protein kinase genes or approximately 1.7% of all human genes [1]. With nearly a third of all proteins containing at least one covalently bound phosphate, it is easy to envisage an important regulatory role for the enzymes that attach (kinases) and remove (phosphatases) inorganic phosphate from proteins. Because kinases are involved in the regulation of all aspects of cellular function, protein kinase inhibition may be a widely applicable strategy for the treatment of disease. Protein kinases now figure prominently among the molecular targets currently being pursued by the pharmaceutical industry to treat a wide spectrum of diseases, these include cancer, diabetes and its complications, rheumatoid arthritis, Alzheimer's, hypertension and stroke.

The realization of the therapeutic promise of protein kinase inhibitors has proven to be a daunting task. Two major obstacles to be overcome were achieving access to these intracellular targets and selectivity of inhibition. With over 500 kinases in the human genome, it is not surprising that selectivity has proven to be the more difficult of the two problems. However, as will be discussed in this review, X-ray crystallography has proven to be the central technology driving our understanding of the structural features governing inhibitor binding and providing direction in the search for selective kinase inhibitors.

2.2

Structure and Function

2.2.1

Tertiary Structure

A great level of structural knowledge has accumulated for protein kinases over the past decade (for various reviews see [2–9]). The minimal kinase domain consists of approximately 300 amino acids: for example, Cyclin-Dependent Kinase 2 (CDK2) has 298 residues (Fig. 2.1) [10, 11]. The kinase domain can be further divided into two lobes or sub-domains, N-terminal and C-terminal, with the lobes connected via a five to six residue “hinge” or “linker” region. The N-terminal lobe is the smaller subunit and consists of approximately 80 amino acids. It is primarily beta sheet in nature but does contain a significant alpha helix segment, termed the C-helix (nomenclature derives largely from the initial Cyclic AMP-Dependent Kinase, PKA, structure [9, 12]). By contrast the C-terminal lobe, in addition to being substantially larger, is largely α -helical in nature.

A binding cleft is formed between the two lobes, and ATP docks into this region. Hydrogen bonds are formed between the N6 exocyclic amino group of ATP and a backbone carbonyl oxygen of the hinge region noted above, as well as between the N1 of ATP and a backbone NH, again from the hinge region. Note

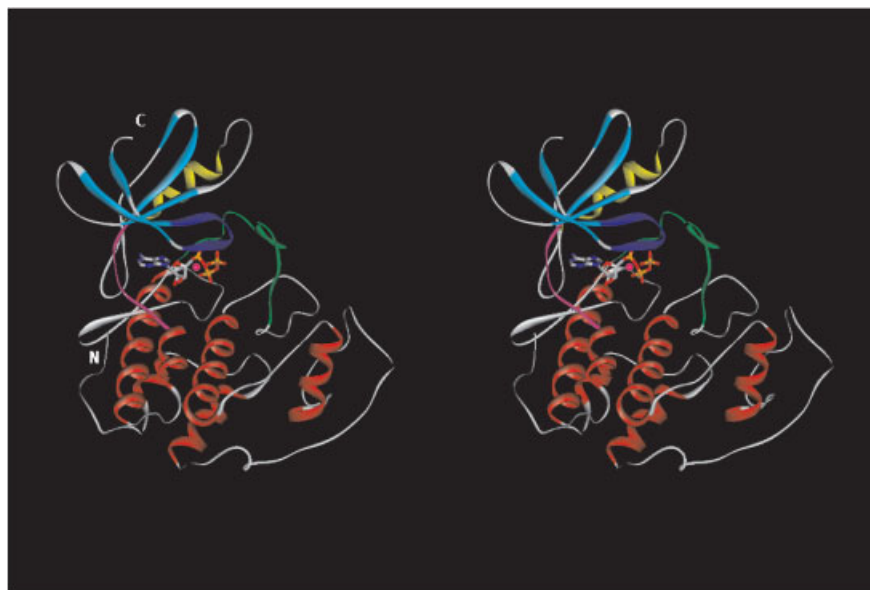


Fig. 2.1 Stereo ribbon diagram of human CDK2 with bound ATP (1HCK.PDB). Structural elements are colored as follows, glycine-rich loop in dark blue, N-terminal beta sheet

in cyan, the C-helix in yellow, linker or hinge region in magenta, activation loop in green and C-terminal lobe helices in red.

that, in addition to these two groups, the hinge region also projects a second backbone carbonyl oxygen into the binding site in the region of adenine H2. This functionality is significant for synthetic ligand-hydrogen bonding interactions as discussed later. Beyond hydrogen bonds to the adenine ring, the ribose hydroxyls of ATP are often observed to show hydrogen bonds to a polar residue (Asp 86 of CDK2 [10, 11], Asp 1083 of insulin receptor kinase [13], and equivalent residues) that resides at the beginning of the C-terminal lobe as defined. The adenine and ribose rings are also sandwiched by a series of hydrophobic residues from both the N- and C-terminal lobes. In particular for CDK2, residues Val-18, Ala-31, Val-64, and Leu-134 have van der Waals contacts with ATP as do equivalent residues of other kinases. An additional residue, Phe-80 for CDK2, is also significant and contacts ATP in the vicinity of C5 and the N6 exocyclic amino group and has been termed the gatekeeper. This residue shows high variability across kinases. As brief examples, it is Met for insulin receptor kinase (IRK), Thr for p38, Gln for ERK2, and Val for VEGFR2 [14–19]. The gatekeeper nomenclature derives from the fact that the residue straddles both the ATP binding site and an additional region deep within the kinase that small molecule inhibitors often utilize (see Sections 2.2.3.1 and 2.4.2.2). The final component of ATP, the triphosphate group, projects as would be expected into a highly hydrophilic and open region framed by a set of strictly conserved residues involved in catalysis. It should be noted at this point that although the small molecule inhibitors designed for protein kinases are predominantly ATP competitive, they commonly occupy regions of substantial volume in addition to where ATP would normally bind. As a result, the “ligand” binding site, when referring to inhibitors, is not synonymous with the ATP binding site.

2.2.2

Catalysis and Substrate Binding

Crystallographic studies, in particular with non-hydrolyzable ATP and transition state analogues as well as substrate analogues, have provided a detailed understanding of the likely nature and mechanism of substrate binding and catalysis [13, 20–24]. As noted, the triphosphate group interacts with conserved residues of the protein kinase. More specifically, the α and β phosphates are oriented via a salt bridge with a lysine (Lys-33 of CDK2) and the lysine itself is further fixed in an appropriate location for catalysis by a conserved glutamic acid residue (Glu-51) residing on the C-helix. Asp-145 and Asn-132 of CDK2 and equivalent residues complete a coordination sphere between the phosphate groups and accompanying magnesium counterions.

Peptide substrate then docks onto the protein kinase, in general presumably occupying a cleft along the C-terminal lobe, as exemplified by the peptide inhibitor in the PKA-AMPPNP-PKI and IRK-ATP structures. Catalysis appears to be via a dissociative transition state mechanism and a planar phosphate intermediate [20, 22]. The incoming peptide hydroxyl is oriented via Asp-127, which is in turn further stabilized via a hydrogen bond to Asn-132. These latter two residues that

are part of the signature-catalytic loop DxxxxN motif that designates proteins as potential protein kinases.

2.2.3

Regulation and Conformational Flexibility

Protein kinases are highly regulated owing to their central role in signal transduction [3, 5, 25]. While a full description of regulatory mechanisms is outside the remit of this review, it is nonetheless important to consider in the context of how it affects ligand binding. More specifically, regulatory control of protein kinases often involves significant conformational changes in structure that alter the shape of the binding site afforded to a small molecule inhibitor. Four mechanisms of regulation and their conformational effects will be mentioned: activation loop conformation, glycine-rich loop conformation, C-helix position and lobe orientation.

2.2.3.1 Activation Loop Conformation

The original unliganded IRK structure [14] revealed a protein in which the ATP binding site was occluded by residues from the so called “activation loop” or “T-loop”. This C-terminal extended region is present in all protein kinase structures to date and, as defined here, begins with a well conserved Asp-Phe-Gly sequence motif (DFG starting with Asp-145 of CDK2 and Asp-1150 of IRK) and extends until just upstream of a consensus Ala/Thr-Pro-Asp/Glu triplet. The loop contains one or more phosphorylation sites, i.e., it is a substrate for another protein kinase or autophosphorylation via protein dimerization. It is useful to consider the activation loop as having a range of conformations, in equilibrium, with two relative extremes. One end of the equilibrium is as seen in the IRK-AMPPNP structure [13] and represents the catalytically active conformation. Specifically, the aspartic acid of the DFG motif is strictly conserved and involved in catalysis (Section 2.2.2), and so must be properly positioned to coordinate magnesium ions. The adjacent phenylalanine occupies a position adjacent to the C-helix and packs against a series of hydrophobic residues from other regions of the protein. The ATP site is unhindered by protein such that ATP can freely dock and form suitable hydrogen bonds to the linker strand with the γ phosphate fully accessible. By contrast, at the other end of the equilibrium, a conformation is observed as represented by the initial IRK structure. The phenylalanine of the DFG motif is translocated by approximately 10 Å to roughly a position where, in an ATP -liganded structure, the ribose would reside. Additionally, the aspartic acid residue conformation is altered via rotation of the peptide backbone ζ angle.

Phosphorylation of the activation loop appears to drive the equilibrium towards a catalytically active conformation due to the ability of the phosphorylated residue to form stabilizing salt bridges with a series of lysine or arginine residues. However, it should be noted that, based on the evidence to date, the activation loop can occupy a range of conformations including one close to an active conforma-

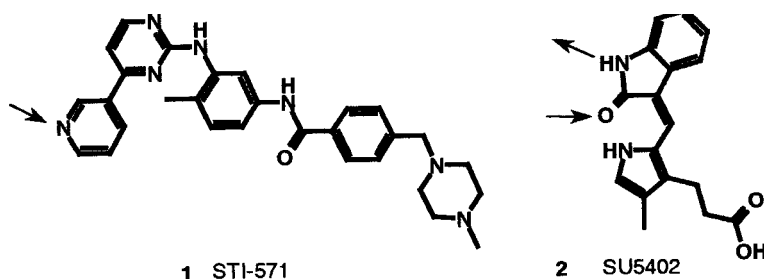
tion in the absence of phosphorylation and an inactive one despite phosphorylation. In short, it is a true equilibrium with energetics and relative populations of different forms varying from kinase to kinase depending on a range of factors including phosphorylation, additional protein and substrate binding, and internal energy of the conformation. The significance for ligand binding, particularly for small molecule inhibitors, is that the nature of the ATP pocket and adjacent regions is significantly altered. For example, STI-571 **1** (Gleevec, Scheme 2.1), as described in more detail later, binds solely to an inactive type conformation [26, 27] with a portion of the ligand residing where the phenylalanine of the Abl DFG triad would reside in an active conformation. Inactive conformations have now been observed for a range of both serine/threonine and tyrosine kinases including Abl, IRK, p38, Tie2, and Twitchin kinase [14, 26–29] indicating that the phenomena is likely widespread across kinases, and so again emphasizing the importance to ligand design. One final point to note is that the region immediately upstream of the DFG motif is relatively constant to date across X-ray structures. In particular, the carbonyl oxygen of the residue preceding the aspartic acid is typically involved in a buried hydrogen bond to either a tyrosine phenol or histidine epsilon nitrogen. This hydrogen bond holds the carbonyl group fixed, and consequently, the backbone NH of the aspartic acid is also fixed via the rigid peptide bond. The major movement of the loop is, as a result, initiated via rotation around the ζ torsion of the aspartic acid.

2.2.3.2 Glycine Rich Loop

The glycine-rich loop (G-rich loop) is defined by the GxGxxG motif that is highly prevalent in protein kinases and occurs in the first 10–20 residues of the kinase domain. Other residues within this region, e.g., Thr-14 and Tyr-15 of CDK2, may be targets for phosphorylation and consequential inhibition of kinase activity. With respect to inhibitor ligand design, the important feature of this region is its inherent flexibility. In particular, ligand structures such as FGFR-SU5402 **2** [30] and Abl-Gleevec **1** (Scheme 2.1) [26, 27] demonstrate the ability of the loop to mold or collapse onto the ligand to form favorable contacts, particularly with the aromatic residue that typically precedes the final glycine. It is difficult to determine based on structural evidence to date, but one can speculate that there is less likelihood of seeing a recurrent pattern for how the loop adjusts in conformation, as opposed to the activation loop conformational changes where patterns are clearly present.

2.2.3.3 C-Helix Orientation

Recall that the C-helix contains a strictly conserved glutamic acid residue. It is apparent from multiple structures of different protein kinases that modulation of the conformation of this helix is a common regulatory theme. Additionally, there are multiple ways of regulating C-helix conformation, and as with other modes of regulation, ligand binding is affected. For CDK2 [10, 11], the C-helix is properly



Scheme 2.1 Arrows indicate the hydrogen bonding pattern of the inhibitor to the kinase backbone determined by X-ray crystallography.

positioned by the binding of cyclin proteins to give a tertiary complex, upon ATP binding, that is catalytically active. In the absence of these proteins, the C-helix rotates on its axis, projecting the glutamic acid more towards the solvent, and translates away from the catalytic lysine. This conformational change creates additional space in a region adjacent to the ATP binding site that a small molecule inhibitor can occupy without steric interference. Similar patterns of C-helix displacement are also observed in Hck and Src Kinase X-ray structures [6, 31–35], but the degree of change is distinct. For Src and Hck, C-helix conformation is not modulated via interactions with cyclins but instead through phosphorylation and/or interaction with other protein domains.

Another mechanism of modulating the C-helix conformation is employed in common for the so called AGC family of kinases: PKA, PKB, PRK, SGK1 and MSK1 are among the members of this family. These kinases utilize a hydrophobic motif (FxxF) beyond the kinase C-terminus to pack adjacent to the C-helix and lock it into an active conformation [12, 36–40]. For proper packing to occur, the evidence to date is that an acid functionality must be immediately adjacent and downstream of the second phenylalanine of the hydrophobic motif. This acidic group can be in the form of either a phosphorylated residue (e.g., PKB/SGK1), an aspartic acid (PRK2), or a C-terminus-free carboxylate (PKA). Crystallographic studies [38–40] for PKA and mutated PKB (S474D) show the acidic functionality to form a salt bridge or hydrogen bonds with residues from the N-terminal lobe [41]. This interaction stabilizes the C-helix location. There are two key features for ligand design considerations. First, as with the other mechanisms described above, the nature of the ligand binding site in the region of the C-helix is affected by the regulatory state. However, there is also a second feature, unique to AGC family members. In addition to the hydrophobic motif, they also have other conserved sequence motifs, which based on PKA and recent PKB structures, impact the shape of the front of the ATP binding site in the vicinity of C2 of ATP. Namely, a phenylalanine (PKA 327, PKB 439), that is upstream of the hydrophobic motif, projects into the front of the ATP site, stacking over a glycine–glycine step that is at the end of the hinge region, and forms an aromatic herringbone type interaction with the C2/H2 of the adenine ring conformation [36, 38–41]. In

other kinases this region is fully accessible, and commonly occupied by inhibitors, so a substantially different binding profile is thus possible for AGC family members, relative to other protein kinases.

2.2.3.4 Lobe Orientation

An additional feature of protein kinase structure and function that impacts ligand binding is the orientation of the N-terminal lobe relative to the C-terminal lobe. Catalysis is, as noted, dependent on residues from both lobes, and as such, the proper spatial orientation between the domains is required for effective catalysis. In the absence of ATP, substrate, or other needed co-factors and adapter proteins, the domains are observed to orient differently via conformational changes in the hinge region as well as a few other confined sequence elements (e.g., an extended region just downstream of the C-helix). The net effect is to alter the shape of the ATP binding cleft and, consequently, the shape of the synthetic ligand needed for inhibition. This effect is perhaps observed to be most pronounced to date for ERK2 and p38 MAPK kinases [16–19, 42]. For p38, the binding site is substantially more open, relative to, e.g., CDK2 or IRK, and this feature as a result alters the profile for the type of ligand needed to inhibit p38 [43].

2.2.3.5 Solvent Channel

The lysine–glutamic acid salt bridge has been described above, but one aspect that has not been emphasized is that this interaction occurs relatively deep in the binding site, in the vicinity of the N7/C8 positions of ATP. Additionally, the backbone NH of the aspartic acid that is part of the DFG motif also projects into the sterically accessible space adjacent to N7 of ATP. These two consistent features across protein kinases (Lys–Glu salt bridge and Asp NH) create a polar region buried within the ATP binding site. The hydrophilicity of this site can be further enhanced by the presence of polar amino acids in this region, for example at the gatekeeper residue or the residue immediately preceding the DFG motif. The net result is to favor and stabilize the presence of a solvent channel of two to three water molecules deep within the ATP binding site. The precise positioning of the water molecules is modulated by the types of amino acids in this region, and the number of water molecules can be expanded significantly for catalytically inactive conformations. Additionally, sterically bulky residues such as Phe-80 of CDK2 may block the presence of some of the water molecules. Inhibitors may either interact favorably with the solvent located in this region or displace it in part or in whole. An increasingly common motif is to target the Asp NH as will be discussed. Regardless of whether the solvent is to be displaced or retained, it is a key aspect of inhibitor design.

2.3

Crystallization

2.3.1

Defining the Construct

One of the most challenging aspects of obtaining a kinase crystal structure is designing the initial constructs. Recent advances in high-throughput cloning techniques, affinity purification tags and crystallization robots have however allowed for multiple constructs to be tested rapidly in parallel.

Initial constructs are often designed based on previously solved crystal structures and usually contain only the catalytic kinase core. However, many kinases are multi-domain proteins and inclusion of additional domains can increase both protein stability and solubility. The minimal kinase domain typically contains five β -strands and one α -helix in the N-terminal lobe and 7–8 α -helices in the C-terminal lobe. In tyrosine kinases, the N-terminal lobe usually starts near a highly conserved tryptophan, 14 residues upstream of the GxGxxG motif. The start of the N-terminus is less well conserved in serine/threonine kinases.

Limited proteolysis has proven useful for identifying the boundaries of the N- and/or C-terminal lobes as well as identifying additional segments or domains that may provide stability. Typically the full length protein is first expressed and purified. The protein is then digested with a panel of proteases, both in the presence and absence of known ligands or inhibitors, and the digests are analyzed by SDS-PAGE. The N- and C-terminus of the stable fragments are identified by N-terminal sequencing and mass spectrometry. In the case of EphB2 and T β R-1, an N-terminal juxtamembrane or regulatory domain was included in the crystallography construct in order to obtain data quality crystals [44, 45]. Regions C-terminal to the kinase domain were included in the CaM and Twitchin kinase constructs and were found to be pseudosubstrates in the crystal structure [46, 47].

The majority of kinases solved to date have been expressed using a baculovirus expression system in insect cells. Only a few kinases, including p38, ERK2, PAK and EphB2, have been successfully expressed in *E. coli* [17, 19, 45, 48]. Constructs typically contain an affinity tag to aid in the purification. Both His₆ and cleavable GST tags [45, 48, 49] have been used successfully. Newer expression systems, including the Invitrogen Echo[®] system, Gibco/Life Technologies Gateway[®] system or the Novagen pTriEx-1 cloning system, allow one to test multiple affinity tags in both *E. coli* rapidly and insect cells with minimal subcloning [50].

To improve the stability and/or solubility of the protein during expression and purification, peptide substrates, binding partners and ligands have been added. The structures of many cyclin-dependent kinases (CDKs) have been solved by co-expression and purification with protein activators and inhibitors. The structure of CDK5 was solved by co-expression with the protein activator p25 [51]. The structure of CDK6 was solved by reconstitution of a CDK6/cyclinD/p18^{INK4c} complex *in vitro* during the purification [52]. The solubility and stability of GSK3, during purification and crystallization, was significantly improved by addition of a

39 amino acid peptide derived from the C-terminus of its binding partner FRAT1 [53].

2.3.2

Mutagenesis

In some cases, mutation of surface or active site residues has been required to improve stability and solubility. Mutation of surface cysteines prevented the formation of disulfide-linked aggregates in FGFR1 [54]. Mutation of highly conserved, active site residues, to create a kinase-dead mutant, was required for a number of kinases whose over-expression proved to be toxic to cells. The conserved aspartic acid in the DFG motif was mutated to an asparagine in CDK5 while the conserved lysine in PAK was mutated to an arginine [48, 51].

Removal of flexible loops within the kinase domain has also been required for some kinases. Many receptor tyrosine kinases contain an additional domain within their C-terminal lobe, referred to as the kinase-insert domain. Removal of this highly charged domain was required to obtain data quality crystals of VEGFR2 [15].

Some proteins have not crystallized despite considerable effort. In such cases, mutation of a closely related kinase to resemble the kinase of interest has provided valuable structural information. Ikuta and co-workers mutated three residues within the ATP binding pocket of CDK2 to the corresponding residues in CDK4. The resulting mutant protein was crystallized and used to design CDK4-selective inhibitors [55].

2.3.3

Phosphorylation

Heterogeneous phosphorylation is often a problem when kinases are expressed in insect cells. Multiple approaches have been used to solve this problem. Proteins have been completely dephosphorylated by incubation with λ protein phosphatase or alkaline phosphatase [38, 39, 56]. Ion exchange and isoelectric focusing chromatography have been used to separate proteins with multiple phosphorylation states. An γ -aminophenyl ATP-sepharose column was used to separate different phosphorylated states of human c-Src [34]. Alternatively, serine/threonine or tyrosine phosphorylation sites can be mutated to alanine or phenylalanine, respectively [42]. For tyrosine kinases with multiple autophosphorylation sites, the active site aspartic acid can be mutated to an asparagine, creating a kinase dead mutant [57].

In many cases the active, phosphorylated form of the protein is preferred for structure-based drug design studies. Quantitative autophosphorylation has been achieved for many kinases including IRK, IGF-1 and VEGFR2 [13, 15, 49]. Residual ATP and ADP must be removed following phosphorylation to ensure homogeneity.

Many serine/threonine kinases require phosphorylation by an upstream kinase for activation. This requires cloning the upstream kinase as well as additional purification steps after phosphorylation [38, 39, 58]. In some cases, the active

form of the protein can be mimicked by replacing phosphorylated serine or threonine residues with aspartic or glutamic acid [38, 39].

2.4

Inhibitor Design

2.4.1

Binding in ATP Cleft

The region of the ATP binding site is the current focus of kinase inhibitor drug discovery research. At first blush the pursuit of selective kinase inhibitors competitive with a substrate common to all kinases, ATP, and which is present at intracellular concentrations (2–5 mM) that are 10–1000-fold in excess of the K_m , would appear to be a poor strategic approach. This belief was widely held into the mid-1990s at which time several reports emerged of kinase-selective ATP-competitive inhibitors. The structural basis for both binding in the ATP site and features governing kinase selectivity was rapidly developed in subsequent X-ray crystallographic studies and as discussed below provides a framework for current kinase discovery efforts. While many alternative binding sites exist which could provide the basis for non-ATP competitive approaches to kinase regulation, none are likely to prove as universally applicable as the compounds that bind into the ATP site. The primary attractiveness of the ATP site is that it affords a deep hydrophobic

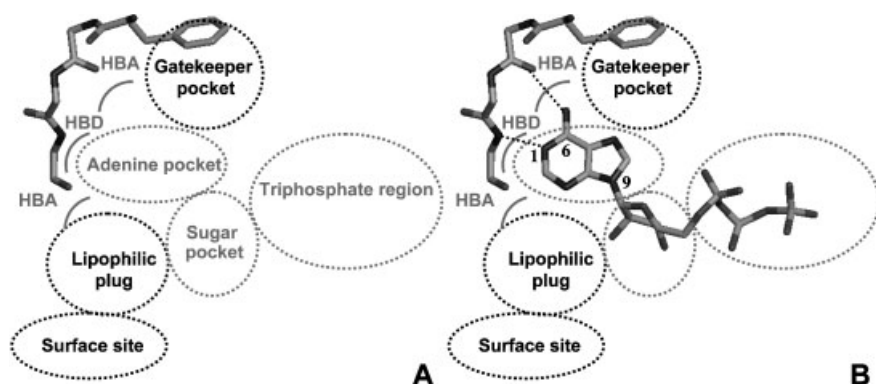
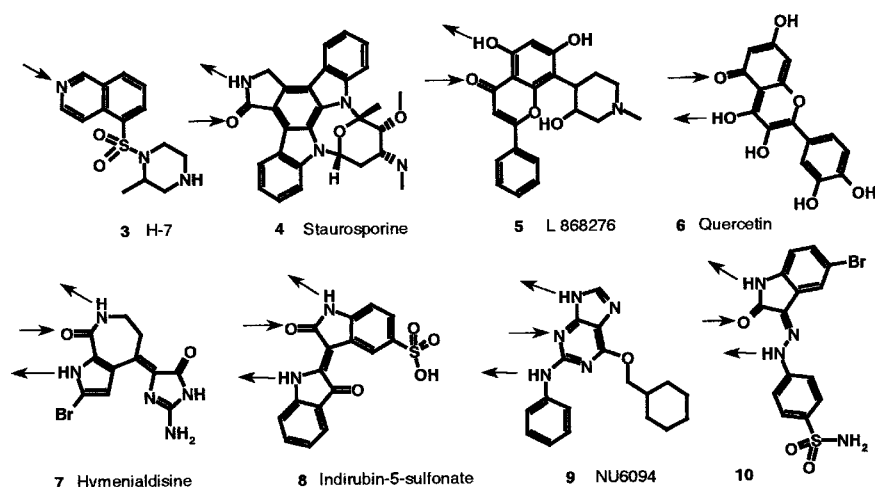


Fig. 2.2 (A) Traxler/Bower Kinase inhibitor pharmacophore. The perspective given is a top down view from the N-terminal to C-terminal domain in which the plane of the adenine ring matches that of the page. A short segment of the kinase linker region is shown which includes the three hydrogen bonding residues (minus side chains) plus the gatekeeper residue. In this orientation the top of

the graphic corresponds to the back of the pocket and the bottom of the figure to the front or solvent. HBA and HBD denote hydrogen-bond acceptor and donor, respectively. The graphics were generated using WebLab ViewerLite from MSI and the indicated PDB files. (B) Complex of ATP, cyclin A and human CDK2 (PDB:1FIN).



Scheme 2.2 Arrows indicate the hydrogen bonding pattern of the inhibitor to the kinase backbone determined by X-ray crystallography.

cleft for the adenine of ATP (or in some cases the guanine of GTP) that is common to all kinases and yet highly varied in shape. The commonality of binding elements provides inhibitor design features that can be applied to all kinases, whereas the exploitation of unique structural features provides the basis for achieving selectivity.

At present there are >100 protein kinase-inhibitor structures in publicly available structural databases which span 28 kinases and a variety of inhibitor structural classes. Based upon an analysis of these data a classification of ATP binding regions has been proposed by Traxler and Furet [59] and Bower (personal communication from Michael J. Bower, December 1999). In the subsequent discussion a slightly modified version of this classification will be used to organize the trends seen across kinases and inhibitor classes (see Fig. 2.2). In the Traxler model, five sites were proposed of which three (adenine, sugar and phosphate-binding sites) can be directly related to ATP and two additional lipophilic sites which lay outside of the region occupied by ATP. In the Bower model, an additional polar site on the surface of the protein was proposed.

2.4.1.1 ATP Binding Sites

As previously described, the purine binding site is well engineered to complement both the planar aromatic nature and hydrogen bonding pattern of adenine. That the great majority of known inhibitors mimic these features highlights the essential role the purine binding site plays in the binding of both ATP and inhibitors. The key features of a purine mimetic are hydrogen bonding functionality, which at a minimum mimics the hydrogen bond formed by N1 of adenine and a planar aromatic or heteroaromatic ring system (Fig. 2.2B). Many inhibitors also

contain a group which occupies the ribose-binding site and engages residues in the protein which either hydrogen bond to a ribose hydroxyl or to groups involved in the ligation of Mg and the α -phosphate of ATP. The isoquinoline sulfonamide **3** (Fig. 2.3A) is good example of an inhibitor which incorporates these binding elements [36]. While inhibitors must minimally contain a single hydrogen bond acceptor, more commonly seen are inhibitors which mimic the dual hydrogen bond donor–acceptor properties of adenine. Examples of donor–acceptor inhibitors (Scheme 2.2) are staurosporine **4** (NHCO of lactam) [60] and des-chloroflavoripridol **5** (carbonyl and phenolic hydroxyl) [61]. Also observed are alternate hydrogen bond donor–acceptor complexes in which the protein acceptor and donor atoms are located on the same amino acid residue (N+3 from the gatekeeper). Quercetin **6**, whose phenolic oxygens bind to the kinase amide backbone, is an example of this alternative pattern [62] (Fig. 2.3B). Interestingly, in casein kinase 2 which utilizes GTP, as well as ATP, as the phosphate donor, this alternate hydrogen bond donor–acceptor pair has also been observed for the binding of GTP [63]. Finally, the structures of hymenialdisine **7** [64], indirbin-5-sulfonate **8** [65], NU6094 **9** [66] and **10** (Fig. 2.5B) [67] in CDK2 provide examples of inhibitors that make all three possible hydrogen bonding contacts within the hinge region.

The triphosphate binding site of kinases presents a conserved set of polar side chains to bind the phosphate portion of Mg–ATP and to position the γ -phosphate for transfer to the acceptor residue (Section 2.2.2). In general, the triphosphate site is largely unoccupied in the kinase–inhibitor complexes thus far described. A notable exception to this trend is balanol **11a**, a potent polyphenolic natural product inhibitor of PKA and PKC. In the PKA structure, balanol occupies the entire ATP binding site, making extensive hydrogen bonding contacts to the glycine-rich loop and other polar phosphate-binding residues (Fig. 2.4A) [68]. A detailed investigation of synthetic balanol analogues revealed that the removal of polar groups which bound in the triphosphate pocket (**11b–d**, OH and CO₂H) had no effect on PKA potency, although PKC inhibition was reduced by 10–1000-fold (Tab. 2.1)

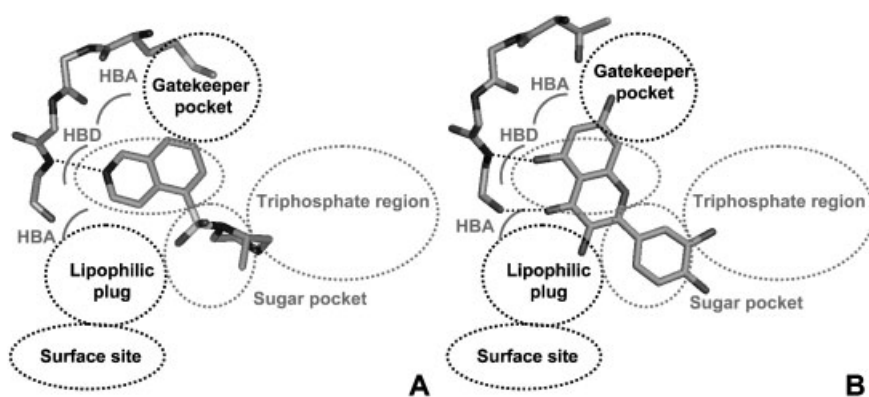


Fig. 2.3 (A) Complex of the catalytic subunit of bovine PKA with **3**, H-7, (PDB.1YRD).

(B) Complex of quercetin, **6**, with the Src family kinase, human HCK (PDB.2HCK).

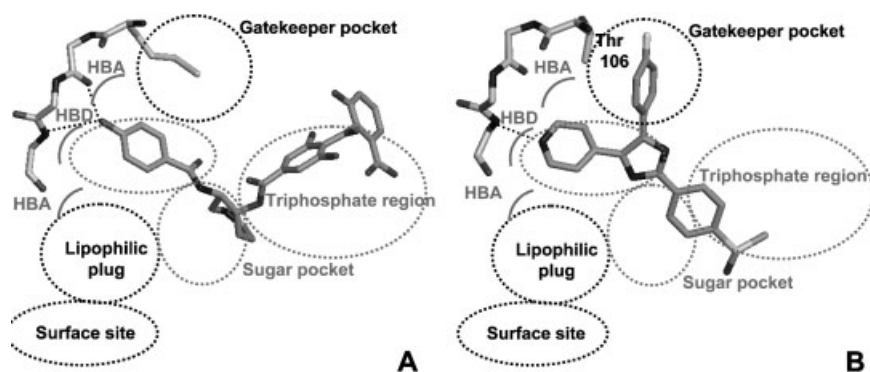


Fig. 2.4 (A) Complex of the catalytic subunit of activated murine PKA with balanol, **11a**,

(PDB.1BX6). (B) Complex of SB 203580, **12**, in human p38 (PDB.1AU9).

[69]. Also prepared was a 5'-deoxybalanol analogue, **11e**, which lacks the phenolic group that hydrogen bonds the amide backbone in the hinge region and again this change had no effect on PKA inhibition. Using these balanol analogues and a number of the isoquinoline sulfonamides, a theoretical model was developed to calculate ligand-binding affinities to PKA [70]. These calculations, which were able to reproduce most of the experimental data, found that the nonpolar binding interactions made the primary contribution to high-affinity binding. Hydrogen bonding interactions in general provided no net contribution to binding as the energetics of binding to the kinase were negated by the requirement to desolvate these polar groups. However, for the poorly solvated isoquinoline inhibitors, the single hydrogen bond made to the amide backbone in an otherwise lipophilic

Tab. 2.1 Potency of balanol analogues prepared by Koide et al. against PKA and PKC [69].

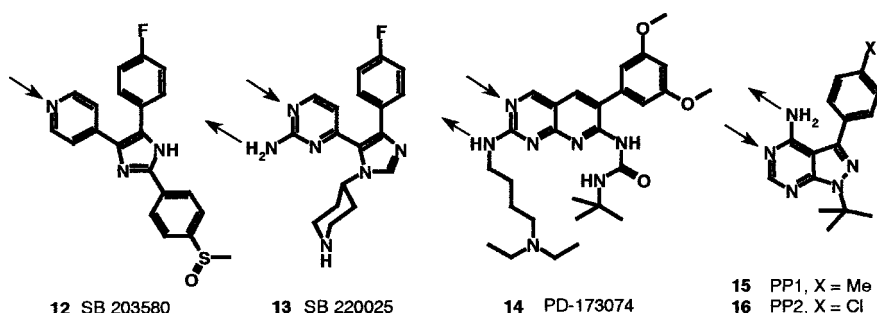
Compound	PKA K_i (nM)	PKC K_i (nM)	X	Y	Z	W
11a	4.7	5.3	OH	OH	CO ₂ H	OH
11b	3.9	604	OH	H	CO ₂ H	OH
11c	11	5000	OH	OH	H	OH
11d	3.4	80	OH	OH	CO ₂ H	H
11e	3.5	69	H	OH	CO ₂ H	OH

binding pocket does make a positive electrostatic contribution to binding. These conclusions, which are consistent with binding studies in other protein-small molecule systems, may provide an explanation for the paucity of kinase inhibitors which make extensive contact in the highly polar triphosphate-binding pocket.

2.4.1.2 Gatekeeper-Dependent Binding Pocket

The solution by Tong and co-workers of the co-crystal structure of SB 203580, **12**, (Scheme 2.3) in p38, a serine/threonine kinase, provided the first structural data to explain how kinase selectivity could be achieved with an ATP-competitive inhibitor [71]. This structure revealed the binding of the fluorophenyl group into a pocket behind and orthogonal to the adenine of ATP. Residues Thr-106 and Lys-53 provide the most extensive contacts with the fluorophenyl group by forming the walls of the pocket above and below the plane of the aromatic ring (Fig. 2.4 B). Recognizing that a limited number of kinases have side chains at position 106 the size of threonine or smaller, Tong proposed a potential role of Thr-106 in determining kinase selectivity. For example, the closely related ERK2, which has the larger glutamine side chain at position 106, is 1000-fold less sensitive to SB 220025, **13** ($IC_{50}=20\text{ }\mu\text{M}$) relative to p38 α [43].

The publication by Tong et al. was quickly followed by similar findings for several tyrosine kinases. Specifically, the co-crystal structures of PD 173074, **14**, in FGFR1-14 (Fig. 2.5 A) [72], PP1, **15**, in Hck [32], PP2, **16**, in Lck [73] and **15** in c-Src [74] all highlighted the significance of this newly discovered binding pocket. Further compelling evidence that a single amino acid residue, which is often termed the gatekeeper, as it controls access to this binding pocket, was the primary factor governing access to this pocket was provided by mutagenesis studies with p38 γ [75], p38 α [76] and c-Src [74]. Two important conclusions from these studies were: (1) that the introduction of a small gatekeeper residue into any protein kinase was sufficient to confer sensitivity to inhibitors that occupied the gatekeeper pocket, and (2) that kinase sensitivity to these inhibitors was closely related to the size of the gatekeeper residue, being greatest for glycine, followed by alanine and then serine and valine (Tab. 2.2).



Scheme 2.3

Tab. 2.2 The effect of site-specific mutagenesis of gatekeeper residues Met-106 (p38 γ) and Thr-338 (Src) on inhibitors which occupy the gatekeeper specificity pocket [74, 75].

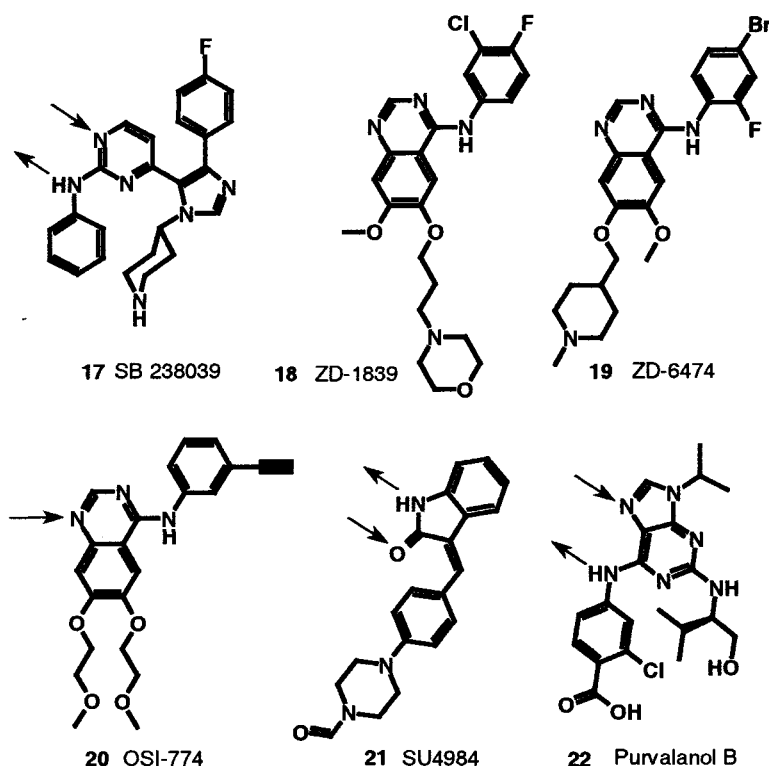
<i>p38γ/SAPK3-12</i>		<i>Src-15</i>	
<i>Gatekeeper</i>	<i>IC₅₀ (nM)</i>	<i>Gatekeeper</i>	<i>IC₅₀ (nM)</i>
Gly	30	Gly	5
Ala	10	Ala	5
Ser	50	Ser (wildtype)	400
Thr	300	Thr	100
Leu	45 000	Val	100
Gln	50 000	Ile	5000
Met (wildtype)	> 100 000	Met	8000
Phe	> 100 000	Phe	8000

2.4.1.3 Lipophilic Plug

The lipophilic environment created by residues that form the adenine-binding pocket extends sufficiently beyond the opening of the pocket to create an additional lipophilic binding area in front of the adenine. This pocket may not be accessible to inhibitors that bind into tightly closed adenine pockets. Inhibitor accessibility to this pocket may also be restricted for some kinase families. For example, the AGC family kinases (PKA, PKC, AKT) have a C-terminal tail that extends back into N-terminal domain and caps the adenine site with a phenylalanine (see Section 2.2.3.3). Inhibitor-kinase structures that bind an aryl ring in this site include **10** in CDK2 (Fig. 2.5 B) [67], **9** in CDK2 [66], and **17** (Scheme 2.4) in p38 [77]. For these later two examples, the substitution of the inhibitor NH₂ with NHPH, which places the aryl group in this lipophilic pocket, affords a 10-fold increase in potency.

2.4.1.4 Polar Surface Site

The extension of the inhibitor beyond the lipophilic site that caps the adenine pocket has been successfully employed either to gain access to solvent (water) or to sites on the surface of the kinase. In general the inhibitor scaffolds that have been pursued as potential drug candidates are relatively flat lipophilic heteroaromatics, and the resulting compounds have very poor aqueous solubility. The attachment of a polar water-solubilizing group, typically a tertiary amine, to take advantage of this access to the solvent has been a key strategy for many kinase drug discovery efforts. Published examples for the anilinoquinazoline class of inhibitors are ZD1839 **18**, an EGFR inhibitor [78] and ZD6474 **19**, a VEGFR2 inhibitor [79]. While no crystal structures of these examples have been published, the recent publication of OSI-774, **20**, in complex with EGFR confirms the anticipated binding mode of these compounds [80]. The crystal structures of **14** (Fig. 2.5 A) [72] and SU4984, **21**, [30] in FGFR1 confirm that the water-solubilizing tertiary amines introduced onto these scaffolds project into the aqueous environment.



Scheme 2.4

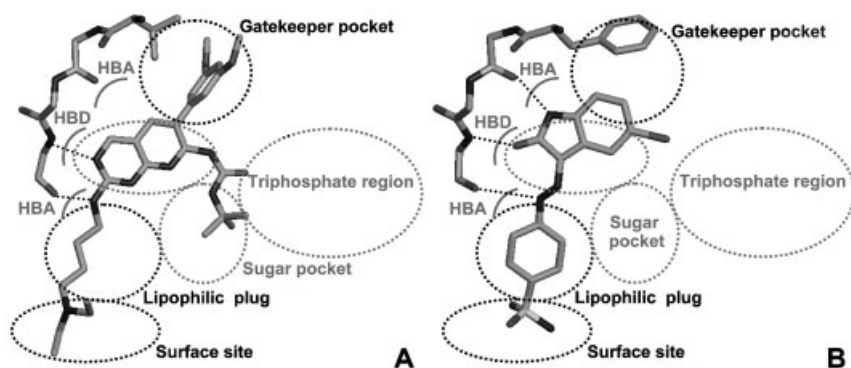


Fig. 2.5 (A) Structure of unphosphorylated FGFR1 soaked with a solution of 14

(PDB.2FGI). (B) Structure of 10 in complex with inactive human CDK2 (PDB0.1 fvt).

As would be expected for functionality that remains water-solvated, the introduction of these amines often has little impact on inhibitor potency. Conversely, the observation of a significant effect on potency for functionality which extends be-

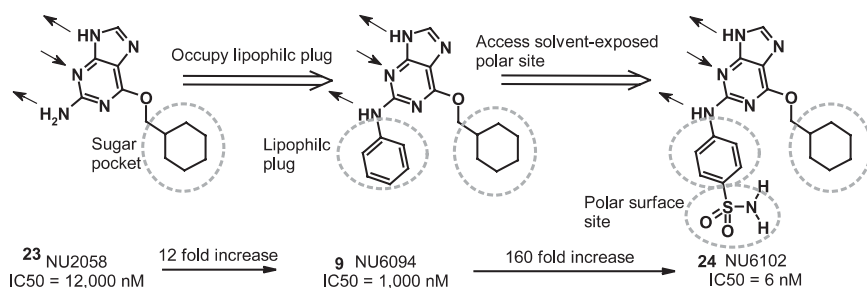


Fig. 2.6 Structure-based design guided by iterative crystallography using activated cyclin A-bound Thr-160-phosphorylated CDK2.

yond the lipophilic plug raises the expectation of a specific binding site on the protein surface. The hydrogen bonding interaction of an aryl sulfonamide with a backbone amide NH of Asp-86 appears to be the main driver for the >100-fold enhancement in binding affinity seen for **10** (Fig. 2.5 B) [67] and **24** (Fig. 2.6) [66].

2.4.2

Conformational Considerations

The conservation of kinase structure and mechanism across the entire protein kinase family dictates a conservation of shape and electrostatics for the ATP binding site when ATP is bound to the catalytically active form of the kinase. However, no such commonality of features is required for the ATP binding site of kinases in their inactive state. In the last few years protein crystallography has provided many insights into the structural biology of protein kinase regulation. The picture that has emerged is complex. Multiple mechanisms have been described for regulation of kinase activity and the catalytic state of the protein is often the sum of competing mechanisms of activation and de-activation. The distortion of the kinase away from its active conformation is the common theme for these regulatory mechanisms. In many cases the conformational change required for activation is the direct result of a phosphorylation in the activation loop [81]. For the cyclin-dependent kinases an additional event, the binding of an accessory protein, cyclins, is required to obtain an active conformation [82]. An in depth account of these mechanisms can be found in recent reviews by Huse and Kuriyan [5], Johnson and Lewis [83] and Engh and Bossemeyer [2]. The following sections consider the impact of conformational flexibility on inhibitor design.

2.4.2.1 Inhibitor-Induced Binding

The structures of two potent PKA inhibitors, staurosporine **4** [84] and balanol **11a** [68], reveal a protein binding surface complementary to that of the inhibitor. Noting the large protein–inhibitor contact surfaces in these structures, Engh and Bossemeyer investigated the relationship of buried surface area to inhibitor potency.

The intuitive notion, that inhibitor potency should be related to binding area, can be theoretically derived as a logarithmic relationship between inhibitor potency and buried surface area [85]. When the data were plotted for PKA (log IC₅₀ versus contact area), all of the compounds fell close to the theoretical line. One remarkable finding is that this correlation held for compounds whose binding induced a significant conformational change in PKA, for example **4** and **11a**. Staurosporine is a particularly instructive example as it is a potent inhibitor of many kinases and in all cases induces a protein conformation complementary to itself. The most likely explanation is that the induced conformations are energetically easily accessible, an assumption well-supported by structural studies which have examined the activation mechanisms of kinases. Several authors noting the tight fit of staurosporine have made reference to the phenomena of hydrophobic collapse. This picture, that the kinase envelopes and molds to the shape of inhibitor, fits well with the conclusions of Hunenberger et al. in which they noted that hydrophobics and not hydrogen bonding was the chief driver of inhibitor affinity [70].

When Engh expanded the analysis to include all the available kinase-inhibitor structures, the fit was poorer. Outliers included STI-571 **1**, which has the greatest buried surface area, but is only moderately potent and a number of compounds that bound more potently than would be expected based upon their surface contacts. An explanation for the relatively poor binding affinity of **1** is suggested from its structure as determined in the un-activated form of Abl [27]. This structure reveals that binding of the inhibitor forces the activation loop into a catalytically inactive conformation which appears to be energetically unfavorable relative to other active conformations. This conclusion was supported by inhibition studies demonstrating **1** to be a potent inhibitor of the un-activated kinase ($K_I = 37$ nM), but a weak inhibitor of the fully activated Tyr-393 phosphorylated Abl ($K_I = 7000$ nM). Moreover, the insensitivity of the fully activated kinase was predicted by modeling which suggested that Tyr-393 phosphorylation would favor the catalytically active conformation of the activation loop. That inhibitor binding can force the kinase into higher energy conformations is a further consequence of conformational plasticity that promises to play an important role in the design of selective kinase inhibitors.

Finally, what about those inhibitors that bind better than predicted based upon an analysis of buried surface area? Examples given by Engh and Bossemeyer include CDK2-22, Hck-15, Lck-16 and p38-13 [2]. This tighter binding could reflect a higher degree of inhibitor–enzyme complementation. The structures of **13** and related pyridinylimidazoles in p38 offer support for this explanation. Unlike many of the structures discussed thus far, the binding of pyridinylimidazoles appears to induce very little change in p38 over that of the catalytically inactive apo-structure [43].

2.4.2.2 What is the Most Appropriate Enzyme Form for Crystallography?

The most pervasive mechanism for the regulation of kinases is that effected by other kinases and phosphatases that add and remove phosphate at multiple sites. Another common mechanism by which kinase activity and substrate binding (ATP and acceptor protein) are regulated is through the binding to proteins that

lie both upstream and downstream to the kinase transmitting the signal. In most of these examples, the regulation of kinase activity is probably effected by conformational changes that either favor or disfavor the attainment of the catalytically active ATP-bound conformation. In principle, consideration of these inactive conformations in designing an inhibitor is not required, if the sole mechanism of the inhibitor is to compete with ATP for the binding to the catalytically active form of the kinase. In practice, it is not that simple, as there are several examples known in which the inhibitors bind equally well to both the un-activated and activated kinase. Also, unfortunately in most cases it has proven much more difficult to purify and crystallize the active protein. Hence, many of the inhibitor-kinase structures have been solved with an inactive form of the kinase.

Recently Davies and colleagues have reported their studies on the design of CDK2 inhibitors. Whereas previous structure-based design efforts had used the structure of inactive monomeric CDK2, prior to initiating their design efforts, the Davies group undertook a comparative study of both the inactive and activated cyclin A-bound Thr-160-phosphorylated structures of CDK2 [86]. This comparison revealed two major changes which might affect inhibitor binding. They also solved the structures of an inhibitor, indirubin-5-sulfonate, bound to both kinase forms and performed calculations to quantify the effect of structural differences between these two forms on inhibitor binding. Having established an empirical basis for using the activated kinase as a tool in structure-based design, in a subsequent publication the authors presented a successful example of structure-based inhibitor design based upon the activated CDK2 structure [66]. Starting with a 12 μ M inhibitor, NU2058 **23**, a >1000-fold improvement in potency was achieved for the optimized compound, NU6102 **24** (Fig. 2.6).

While the above example supports activated kinases as the most relevant form for structure-based design, the behavior of STI-571 **1**, which binds some 200-fold more potently to the un-activated form of Abl, suggests that inactive kinase structures may also have value [27]. The Abl-1 structure revealed that inhibitor binding locks the activation loop in an unusual inactive conformation opening up a new pocket deeper in the ATP site. Potential advantages of inhibitors that reform the ATP binding site are enhanced selectivity and the ability to lock the kinase into a non-activatable conformation. However, these advantages are not fully realized for **1**, as it is not uniquely selective (Abl, c-Kit and PDGFR are all equally inhibited) [87] and there are no data demonstrating that **1** inhibits activation. Another potential advantage is that the inhibitor may not have to compete with ATP for binding to the un-activated kinase. Given the high intracellular concentration of ATP, the cellular potency of inhibitors that are strictly competitive with ATP should be greatly attenuated, and as such binding of the inhibitor to an un-activated enzyme form with low ATP affinity might circumvent the disadvantage inherent to ATP-competitive inhibitors [88].

Recent Abl structural studies comparing the binding of **1** to PD 173955, **25**, suggest that for this example the unique binding conformation may come at a high cost [26]. PD 173955, which has a reduced contact surface with the enzyme, does not bind in the pocket formed by movement of the activation loop. However,

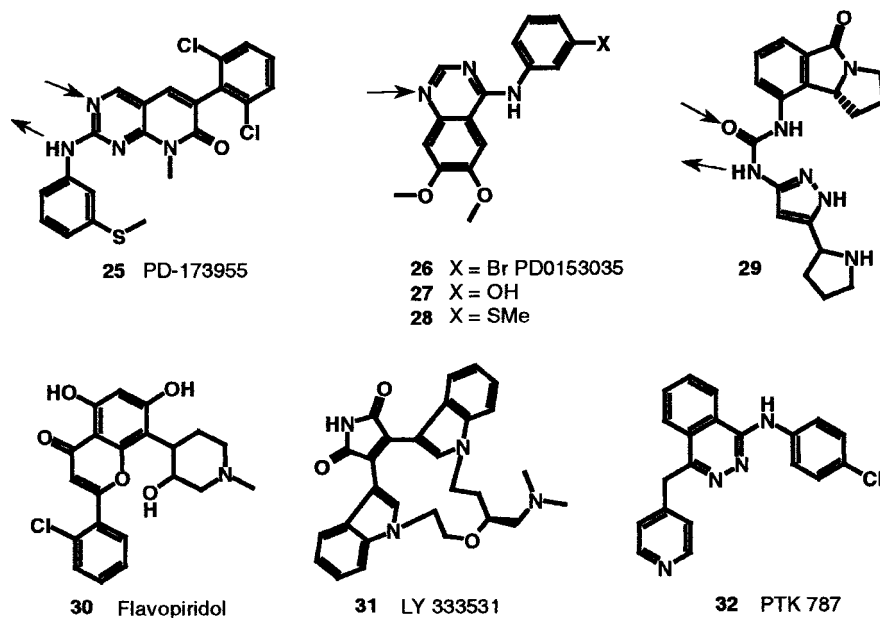
inhibition studies demonstrate that **25** inhibits both the un-activated and activated kinase equally well. The explanation proposed by the authors is that the larger binding surface of **1** is required to pay the penalty for binding to the novel inactive conformation. In contrast **25**, which does not bind to this novel activation loop conformer, is an equipotent inhibitor of both un-activated and activated Abl. The authors propose that the differences in cellular potency of these two inhibitors (**25** is some 10-fold more potent in a cellular assay) may be a reflection of their differing binding modes and conclude an advantage for the inhibitor which can bind to multiple enzyme forms (Scheme 2.5).

The rationale of using the activated CDK2 structure to design **24** are clear. However, these studies leave unanswered the question as to whether this structure provided a result superior to that based upon the inactive kinase. Indeed, many groups have used the inactive form of CDK2 to successfully aid their lead optimization efforts. The Abl example illustrates how inactive kinase structures can lead to unique design possibilities. To re-iterate, there are several mitigating factors that make any blanket statement on choice of enzyme form unwise. As will be outlined in the subsequent section, much progress has been made in kinase design irrespective of the activation state, a result which suggests that any structure is better than none.

2.4.2.3 Homology Models and Surrogate Kinases

The conservation of tertiary structure and mechanism, which complicates the discovery of highly selective kinase inhibitors, provides for a portability of learning from one kinase to the next. Furthermore, the more closely two kinases are related, the more likely the information can be applied to the second kinase. One way in which structural knowledge is extrapolated from one kinase to the next is the process of homology modeling [89]. To construct a homology model the backbone tertiary structure of a known kinase is used as the template upon which the side chain residues of the target kinase are attached. Using various minimization algorithms, the homology model is refined and thereby provides a starting point for inhibitor docking or design. As discussed elsewhere in this chapter, homology modeling is a routinely used tool that can provide valuable insights.

The application of inhibitor-binding modes from one kinase to the next is another example of the portability of structural information. The universal commonality of an adenine mimetic in inhibitor templates, and the corresponding engagement of common backbone-binding elements in the kinase provides a powerful rationale for this process. Structural proof of a common binding mode in PKA, c-Src, Lck, CDK2 and Chk1 has been demonstrated for staurosporine, which potently inhibits all of these kinases [31, 73, 84, 90, 91]. Structural confirmation of a common binding mode for **15** and **16** to several members of the Src family of cytosolic tyrosine kinases has also been obtained [32, 73, 74]. In the above examples it can be argued that the unique structure of staurosporine or the close similarity of Src family members predisposes these systems to exhibit a common binding mode. However, even in highly unfavorable cases a “conservation” of binding



Scheme 2.5

modes has been observed. The anilinoquinazoline **26**, which has a sub-nanomolar IC_{50} against EGFR, weakly inhibits p38 ($IC_{50}=6.3\ \mu M$) and is all but inactive against CDK2 ($IC_{50}=250\ \mu M$). Nonetheless, the binding mode observed for **27** in CDK2 ($IC_{50}=1\ \mu M$) and **28** in p38 ($IC_{50}=5\ \mu M$) [42], is the same as that recently observed for the related anilinoquinazoline, **20**, in EGFR [80].

There are several ways in which one may want to exploit the conservation of binding mode to advance inhibitor design. Two very common situations are the lack of protein crystallography for the target kinase at the initial stages of a new effort and the failure to obtain protein suitable for crystallography. To illustrate the later problem, despite the continued efforts of several groups to purify and crystallize Raf and MEK, the kinases upstream of ERK, no structures have been forthcoming. On the other hand, multiple groups have succeeded in crystallizing CDK2, p38 and members of the Src family. In our own labs when the above situations arise we routinely use the binding mode obtained from the more structurally tractable kinase as a “surrogate” for the target kinase. Ikuta et al. have published a clever application of the surrogate kinase concept to assist in the design of CDK4-selective inhibitors. Their approach was to design and crystallize a CDK2 mutant containing a region of the CDK4 ATP binding site in place of that of CDK2. This CDK4 mimic, which retained activity and demonstrated CDK4-like inhibitor selectivity, was used to design an inhibitor **29** that demonstrated high selectivity for CDK4 versus CDK2. The crystal structure of the CDK4 mimic-**29** complex was consistent with their design which identified a single amino acid as the

primary determinate for inhibitor selectivity [55]. Ideally, as in the case above, one would like to have available a surrogate kinase closely related to the target of interest, as this will make it more likely that the inhibitor will bind to the chosen surrogate and that information obtained will be more readily applied to the target. However, this is not a requirement. When faced with a new structural class of unknown binding mode, the solution of the first structure in any kinase is of great value. The process is not foolproof and as with any other co-crystal structure the utility of the information is dependent upon its predictive value.

2.4.3

Paradigms for Kinase Drug Discovery

One gauge of the impact of protein crystallography on kinase inhibitor design is the increasing frequency of its use. The majority of recent publications employ molecular modeling and report the use of protein kinase-inhibitor structures in the design process. Further evidence of the growing importance of crystallography is the increasing frequency of new kinase structures. In 2002 the list of newly solved structures included EFGR [80], AKT2 [38, 39], MAPKAPK2 [92], Aurora A [93], EphA2, and FAK [94]. The following discussion illustrates how this information is being applied to inhibitor design.

2.4.3.1 High Throughput Screening

Where to start for the design of a kinase inhibitor? Does one follow the example set for proteases using a knowledge of mechanism and substrate to design inhibitors and then use iterative cycles of crystallography and design? For now and into the foreseeable future the answer is no. Instead, with rare exception, the starting point for kinase design has been a high throughput screening (HTS) lead. As has been true for other drug targets, natural products have proven to be a rich source of kinase inhibitor scaffolds. The solution of kinase crystal structures with natural products, such as staurosporine **4**, L868276 **5**, quercetin **6**, hymenaldisene **7**, and balanol **11a** (Scheme 2.2 and Tab. 2.1) have provided an understanding of binding in the ATP site which underpins much of the current design efforts. Furthermore, modifications of these natural products has led to the discovery of compounds which are being studied in clinical trials for the treatment of cancer (flavopyridol, **30**) [95] and diabetic retinopathy (LY33351, **31**) [96]. The screening of large compound libraries has also provided a rich source of new inhibitor scaffolds. An unusually successfully example is PTK 787 **32**, a KDR inhibitor screening lead which has advanced into clinical trials [97]. Major inhibitor classes found as kinase leads using HTS that have led to clinical candidates are the quinazolines [98], oxindoles [99], anilinoypyrimidines [100] and the recently discovered diaryl ureas [29].

2.4.3.2 Structure-Based Design

In a 1995 publication Furet et al. at Ciba Geigy used the available PKA structures to predict the binding interactions of PKA with staurosporine [101]. This prediction, which was subsequently confirmed with a co-crystal structure, stands as an early illustration of how crystallography would guide the discovery of kinase inhibitors. With the increasing success of crystallographers, many inhibitor design efforts for a novel kinase can now begin with a similar level of knowledge. In the absence of such information, inhibitor-docking experiments using a homology model or alternatively, crystallography data from a surrogate kinase can be used to guide inhibitor design efforts. The following section details several examples illustrating the current status of structure-based design. For more comprehensive reviews on inhibitor design see Traxler and Furet [59], Toledo and Lydon [60], Scapin [102], and Garcia-Echeverria et al. [103].

The staurosporine 4 binding model developed by Furet et al. correctly identified the lactam amide as the donor–acceptor that bound to the linker backbone [101]. Furthermore, the authors hypothesized the indole rings to occupy the sugar and back hydrophobic pocket (referred to as the sugar and gatekeeper pockets in Fig. 2.2A). This model led the group to design less rigid analogues which retained a cyclic imide to maintain the essential backbone-binding element identified in staurosporine (Fig. 2.7). While the bisindolyl maleimides, 33, maintained the potent PKC inhibition seen with 4 [104], the dianilinophthalimides, 34, proved to favor inhibition of EGFR kinase [105]. The success of these new templates provided support for the original modeling and design hypothesis, which at this time had yet to be structurally verified, and furthermore demonstrated the possibility of achieving more selective kinase inhibition with an ATP-competitive inhibitor. Docking studies using the dianilinophthalimides in a homology model of the EGFR kinase led to the development of a model similar to that previously proposed for 4. Based upon this model, a sulfur-aromatic interaction of a meta-halo aryl group of the inhibitor with Cys-773 present in the ribose-binding pocket was proposed as a key-binding element [106]. Furthermore, since Cys-773 is uniquely present at this position in the EGFR family, this interaction could explain the enhanced selectivity of these compounds for EGFR kinase. Continuing their focus

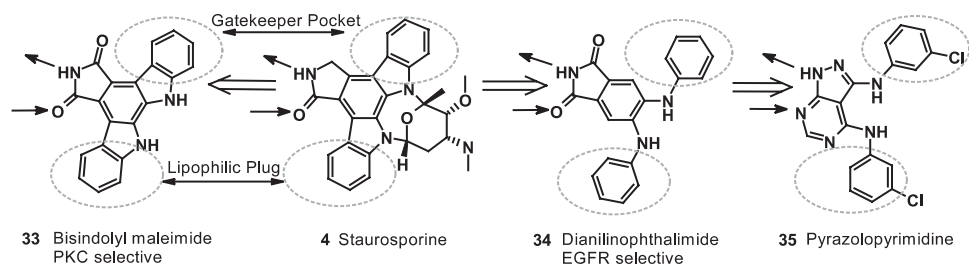


Fig. 2.7 The structure-based design of new inhibitor templates based upon a model of staurosporine bound to PKA.

on EGFR, which had been identified as a therapeutically attractive target, the Novartis (Ciba-Geigy) group pursued two more related series, the pyrazolopyrimidines, **35** [106] and the pyrrolopyrimidines, **36** [107]. In both series a preference for a meta-halo aryl group was demonstrated for the ring which was presumed by modeling to fit in the lipophilic pocket adjacent to Cys-773 (Fig. 2.8).

Independently, 4-anilinoquinazolines were being pursued as EGFR inhibitors at Parke-Davis [108] and Zeneca [109]. The obvious structural similarities between the Novartis series and the quinazolines extended to the SAR, which indicated in both series a preference for a meta-halo group on one of the aryls. These similarities led the Parke-Davis group to propose that a common binding mode for **26** and **36** would be expected (Fig. 2.8). However, docking studies on the EGFR homology model developed at Parke-Davis led this group to propose an alternative to the Novartis model, in which the meta-halo aryl group was positioned in an inner lipophilic pocket next to the gatekeeper. In a comparison of both docking models against all the Novartis data, the Parke-Davis model was better able to fit all the data [110]. A recent X-ray crystal structure of EGFR with anilinoquinazoline **20** matched the Parke-Davis model [80]. The successful design of CI-1033, **37**, a selective inhibitor which binds irreversibly to EGFR through reaction of the arylamide with Cys-773, provides additional evidence to indicate that the Parke-Davis model will hold for other anilinoquinazolines and closely related scaffolds [98, 111]. However, further extrapolation of this result is premature. Changes in template substitution and differing interactions between kinases have been shown to affect the binding orientation of kinase inhibitor templates [43, 112].

It is important to note that in the 3–5 year period prior to crystallographic confirmation of these models, they were used successfully by several research groups to design new templates and to advance lead optimization programs. The modeling predictions of Trumpf-Kallmeyer at Parke-Davis stand out in this regard [113]. In a publication outlining the development of a binding model for two series of 6-aryl pyridopyrimidines **38** (Fig. 2.9), the authors present in detail the methodology

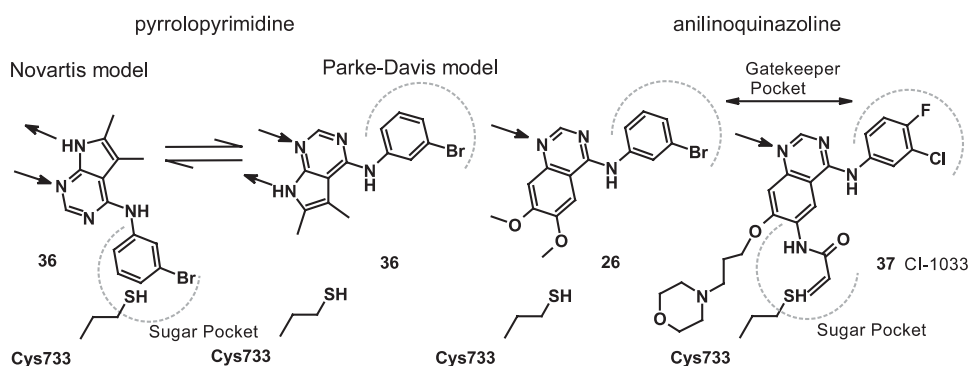


Fig. 2.8 The pyrrolopyrimidine- and anilinoquinazoline–EGFR binding models proposed by researchers at Novartis and Parke-Davis.

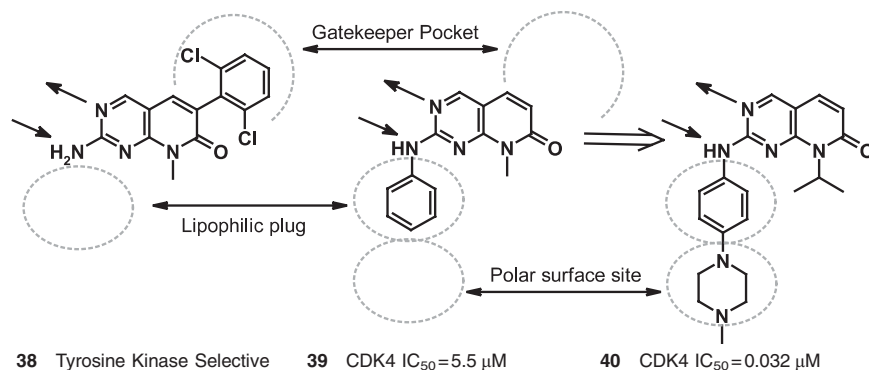


Fig. 2.9 The transformation of a tyrosine kinase-selective inhibitor **38** into a CDK-selective template (**39** and **40**) was achieved by adjusting the template substitution pattern.

and rationale used to develop the model. Because these new series potentially inhibited FGFR and c-Src but also displayed significant activity against PDGFR and EGFR, homology models were built and docking studies performed for all four kinases. The analysis of docking studies were guided by the SAR and rationalize the selectivity for inhibition of the these four tyrosine kinases. Their conclusion of a 4-anilinoquinazoline-like binding mode for **38** (and related series) was verified in a crystal structure of **14** in FGFR1 [72]. Furthermore, the more general conclusion on the importance of the gatekeeper (Val-561 of FGFR1) as a key specificity residue for all kinase classes is now widely recognized.

An excellent example illustrating how structural knowledge guides design is the transformation of the 6-aryl-pyridopyrimidones **38** from a tyrosine kinase inhibitor-specific template to favor inhibition of the cyclin-dependent kinases (Fig. 2.9) [114]. Docking of the **38** into CDK2, using the same binding mode developed for the tyrosine kinases, was not possible as access to the gatekeeper-dependent binding pocket was blocked by the large phenylalanine residue which is present in all the cyclin kinases. The authors concluded this to be the primary reason these compounds were not cyclin-dependent kinase inhibitors, and if true, then pyridopyrimidones lacking this 6-aryl group should gain affinity for the cyclin kinases while losing affinity for tyrosine kinases. Pyridopyrimidines, **39** and **40**, having hydrogen in place of aryl at the 6-position were synthesized and proved to be potent cyclin kinase inhibitors.

Concurrent with these modeling efforts, several X-ray crystallographic studies were published for inhibitor-kinase complexes containing selective inhibitors. These included p38 [18, 43, 71], CDK2 [61, 82, 90, 115], several of the Src kinases [31, 32, 73, 74] and FGFR [30, 72]. The consistent theme of inhibitor-protein interactions (outlined in Sections 2.4.1/2.4.2) elucidated by these studies have provided the foundation for subsequent inhibitor design. Much of this design has been the reworking and optimization of templates discovered in screening, such as quinazolines, purines, diaryl imidazoles. Exemplary of this work are a structure-guided

iterative design cycle described by Davies et al. for the optimization of **24**, a purinyl inhibitor of CDK2 [116], and the design of CI-1033 **37**, an anilinoquinazoline containing a warhead acrylamide, which confers irreversible inhibition to the EGFR family of kinases [98].

2.4.3.3 Mechanism-Based and Ligand Mimetic Design

The skeptics that predicted the impossibility of a highly kinase-selective ATP-competitive inhibitor have now been silenced. However, the reasoning from which this skepticism arose remains to frustrate attempts to employ the rationale substrate-based design algorithms that have proven fruitful for other protein classes. Structure-based design efforts which rely on close analogues of either ATP or a peptide or a combination of both in a multi-substrate analogue have met with limited success, and their pursuit till now has been largely as an academic exercise to probe mechanistic details of the kinase reaction. Exemplary of this work is an ATP-peptide bi-substrate analogue inhibitor for which the kinetics and structure have been determined with the insulin receptor kinase [117]. For further information on the status of these approaches see the reviews by Lawrence and Niu [88] and Parang and Cole [118].

2.4.3.4 Computational Chemistry and Virtual Screening

Honma et al. from Banyu Pharmaceuticals have published one of the few examples of kinase inhibitor design which begins with a de novo ligand-building program to generate novel ligands [119]. Realizing the limitations of the de novo lead generation program (LEGEND), Honma chose to filter the structures generated by the de novo software by developing a program (SEEDS) to extract the essential information from the de novo structures and search it against a database of commercially available and/or synthetically accessible ligands. The protein residues chosen as starting points for design of a CDK4-specific inhibitor were those that make the backbone donor-acceptor hydrogen bonds to adenine and the conversed lysine (Lys-33) and glutamic acid (Glu-51) which were known to be involved in the binding of several CDK2 inhibitors. The coordinates used were from a CDK4 homology model based upon CDK2. Following further visual filtering to eliminate undesired skeletons and functional groups, a set of 350 compounds were purchased and screened. Hits were ranked based upon their structural tractability as leads. Optimization of this lead was performed in an iterative manner, making small arrays followed by evaluation of SAR using docking back into the homology model to guide the selection of reagents for subsequent arrays. The end result was **41**, a 42 nM inhibitor of CDK4, which was shown in an X-ray with CDK2 to mimic the main features of the proposed binding mode [119].

With the exception of the de novo ligand building, the design process used by Banyu to discover the CDK4 inhibitor provides a good introduction to the types of computational approaches that are routinely being employed with increasing success. Given the success of the de novo ligand design, it is worth considering why

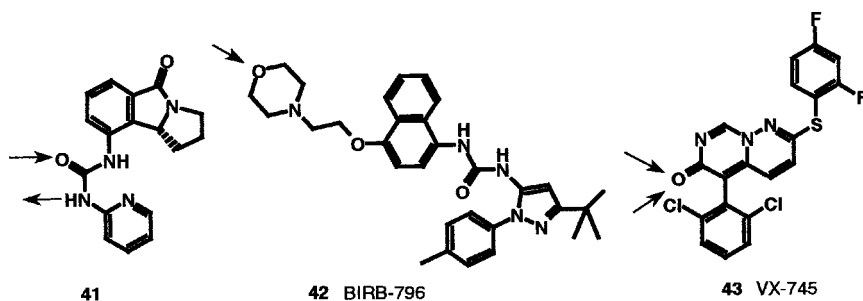
this has not been more widely used. Based upon our own experience, the most likely answer is that it has been tried but has yet to offer any advantage over the other available methods. In the Banyu implementation of de novo design, no “wet” chemistry was performed prior to screening of commercially available ligands. Given this restriction, one could have directly employed an alternate strategy of 3D pharmacophore searching of commercial databases. Such a process of pharmacophore generation and virtual screening has been described by Kinetix Pharmaceuticals [60]. Also Furet et al. from Novartis have reported on two examples where structure-based design was combined with database searching to discover leads for CDK1 and CDK2 [120, 121]. The development of kinase inhibitor libraries and the assembly of kinase screening sets is another way the generic pharmacophore elements described above and in the Traxler/Bower model have been applied to advance lead discovery. To run a complete HTS screen of one million compounds is time consuming and expensive. As such there is considerable value in developing strategies to assemble smaller kinase inhibitor-enriched screening sets. For example, purvalanol B **22**, a potent CDK2 inhibitor, was discovered in a purine-based combinatorial library [122]. Finally, it should be noted that the chances of getting a lead from HTS are increasing as the screening collections are populated with compounds prepared for other kinase inhibitor efforts. For example, compounds from the pyridinylimidazole class of p38 inhibitors have served as leads for c-Raf [123] and Alk5 [124]. For further details on the role of computational chemistry in kinase inhibitor structure-based design strategies and the range of computational tools being applied in this area see a recent overview by Woolfrey and Weston [89].

2.4.4

Selectivity

A realistic goal for inhibitors binding in the ATP site would be to realize that in most cases very high selectivity is not obtainable, is unnecessary and undeterminable. That absolute selectivity is not required is demonstrated by **1**, the first kinase inhibitor approved for use in humans, which in addition to inhibiting the Bcr-Abl target, also blocks c-Kit and PDGFR [87]. Such selectivity being undeterminable relates to the technical challenge of developing the methodology to assay all 518 putative human kinases. Setting aside further discussion of what level of selectivity is required for a specific target, the fact remains that if inhibitor selectivity is insufficient the undesired side-effects are expected to be intolerable. Hence, developing strategies to achieve the required level of selectivity for the given target will be a top priority for a kinase drug discovery effort.

As noted earlier, unlike the phosphate-binding region, where the majority of residues are identical across all kinases, the adenine and ribose-binding pockets are lined by a less rigidly conserved set of residues. We have already covered examples illustrating the effect the gatekeeper has on the selectivity of inhibitors that access the gatekeeper-dependent binding pocket. The exploitation of Cys-773 in the EGFR receptor to develop irreversible inhibitors (**37**) is another example



Scheme 2.6

where a specific residue has been exploited to enhance potency and selectivity. The interpretation of X-ray structural data in light of inhibitor SAR led to the identification of these examples and promises to identify more differences that can be exploited to enhance selectivity (for additional examples see [30, 91]). However, in other cases where selective inhibition has been observed, elucidating the features responsible for selectivity has proven more difficult [64]. Looking further beyond the ATP binding site, several groups have discovered inhibitor binding sites at the solvent interface that impart selectivity. Two examples are the interaction of the sulfonamide of inhibitors **10** [67] and **24** [116] with D86 of CDK2 and the exploitation of the change from lysine (CDK2) to threonine (CDK4) at position of 89 to prepare a CDK4-selective inhibitor, **29** [55].

The conformational mobility of kinases, which is key for controlling the catalytic activity and substrate selectivity of kinases, may also prove to be key for obtaining inhibitor selectivity. Gleevec **1** [27], which inhibits Abl, and BIRB-796 **42** [29], which inhibits p38 α , both bind to catalytically inactive kinase conformations in which the phenylalanine of the conserved DFG motif in the activation loop is displaced by the binding of the inhibitor to the aspartate NH (Scheme 2.6). As noted in Section 2.2.3.1, this behavior has been observed for additional kinases and is likely to have applicability for the design of selective inhibitors of many kinases. Structural studies on the p38-VX-745, **43**, complex have revealed an inhibitor-induced flip of the Gly-110 peptide geometry which allows the inhibitor carbonyl oxygen to form two hydrogen bonds to amide NH groups in the linker region [125]. Researchers at Merck, who first reported this observation, speculate that the Gly-110 of p38 is uniquely suited for this conformational flip relative to kinases which do not have a glycine at this position and that this could explain the enhanced selectivity of **43** for p38. Several new classes of p38 inhibitors have been designed to take advantage of this feature and demonstrate improved kinase selectivity [125, 126].

2.5

Conclusion

Structural insights gained from crystallographic studies with individual members of the kinase family have resulted in the elucidation of a single, universally applicable kinase motif. Similarly, crystallographic studies of kinase-inhibitor complexes are providing an understanding of inhibitor affinity and selectivity that can be broadly applied. As reviewed in this chapter, when coupled with computational techniques, researchers have been able to design new inhibitor scaffolds and improve inhibitor selectivity. Recognition of the high strategic value of having crystallographic information at the beginning of a lead optimization effort is a driving force for improving the techniques to obtain crystalline kinases and the rapid solution of their structure, and if this fails, for the use of a surrogate kinase. This recognition and the elucidation of the general principles of kinase inhibitor binding is also driving the synthesis of kinase inhibitor libraries and the assembly of focus screening sets comprised of compounds which fit these kinase inhibitor descriptors.

It is important to note that current methodologies for the generation of homology models are inadequate for the task of generating the novel conformations that have been revealed by protein crystallography. Consequently, the best starting point for a kinase structure-based lead optimization effort based upon a novel inhibitor scaffold is to solve the structure of the ligand-kinase complex. NMR offers an alternative experimental technique for studying the solution dynamics of protein-ligand interaction and this technique has been successfully used in the design of novel inhibitors of several protein classes [127]. At present, the utility of NMR for the study of kinases is limited by the low resolution of structural information which may be gained and the general difficulty of the experiments. Another promising approach is the application of novel computational techniques to predict kinase-inhibitor structures. However, for this to be realized new approaches must be devised to account for the surprising conformational flexibility of kinases. Perhaps, by supplying a learning set of allowed motions, crystallography can provide the information required for the implementation of new conformational searching and inhibitor-docking algorithms.

Protein crystallography is anticipated to continue its dominant role as a spectroscopic technique to drive the optimization of ATP-competitive kinase inhibitors. Moreover, structural insights revealed by crystallography are likely to be essential for maximizing the potential of alternative non-ATP-competitive inhibitors. Examples of non-ATP-competitive inhibitors that have been explored are antagonists of SH2-binding domains [128] and the MEK inhibitors first discovered by scientists at Parke-Davis [129]. In the case the SH2-binding domain, protein crystallography has provided a detailed understanding of the structural features required for high affinity binding. However, the identified binding motif presented significant challenges for the development of small molecular weight cell-permeable inhibitors, and thus far chemists have been unable to develop drug-like inhibitors of these protein-protein interaction domains. In contrast, while the binding site of the

MEK inhibitors remains unknown, these compounds have been developed into a compound appropriate for clinical development [130]. In their role as proteins that integrate and transduce cell-signalling stimuli, kinases often have multiple activators and multiple substrates plus additional interactions which modulate their activity. Thus there exist many potential sites for intervention beyond the substrates involved in catalysis (ATP and phosphoacceptor). Finding leads which bind at these sites will require new, non-ATP-centric screening methodologies. Once identified, protein crystallography will again be expected to lead the way towards fully exploiting these new opportunities.

2.6

References

- MANNING, G., WHYTE, D. B., MARTINEZ, R., HUNTER, T., SUDARSANAM, S., *Science* **2002**, 298, 1912–1934.
- ENGH, R. A., BOSSEMEYER, D., *Pharmacol. Ther.* **2002**, 93, 99–111.
- HUBBARD, S. R., *Prog. Biophys. Mol. Biol.* **1999**, 71, 343–358.
- HUBBARD, S. R., *Front Biosci.* **2002**, 7, d330–d340.
- HUSE, M., KURIYAN, J., *Cell* **2002**, 109, 275–282.
- SICHERI, F., KURIYAN, J., *Curr. Opin. Struct. Biol.* **1997**, 7, 777–785.
- SOWADSKI, J. M., EPSTEIN, L. F., LANKIEWICZ, L., KARLSSON, R., *Pharmacol. Ther.* **1999**, 82, 157–164.
- TAYLOR, S. S., KNIGHTON, D. R., ZHENG, J., TEN EYCK, L. F., SOWADSKI, J. M., *Annu. Rev. Cell Biol.* **1992**, 8, 429–462.
- TAYLOR, S. S., RADZIO-ANDZELM, E., *Structure* **1994**, 2, 345–355.
- DE BOND, H. L., ROSENBLATT, J., JANCARIK, J., JONES, H. D., MORGAN, D. O., KIM, S. H., *Nature* **1993**, 363, 595–602.
- JEFFREY, P. D., RUSSO, A. A., POLYAK, K., GIBBS, E., HURWITZ, J., MASSAGUE, J., PAVLETICH, N. P., *Nature* **1995**, 376, 313–320.
- KNIGHTON, D. R., ZHENG, J. H., TEN EYCK, L. F., ASHFORD, V. A., XUONG, N. H., TAYLOR, S. S., SOWADSKI, J. M., *Science* **1991**, 253, 407–414.
- HUBBARD, S. R., *EMBO J.* **1997**, 16, 5572–5581.
- HUBBARD, S. R., WEI, L., ELLIS, L., HENDRICKSON, W. A., *Nature* **1994**, 372, 746–754.
- MCTIGUE, M. A., WICKERSHAM, J. A., PINKO, C., SHOWALTER, R. E., PARAST, C. V., TEMPczyk-RUSSELL, A., GEHRING, M. R., MROCKOWSKI, B., KAN, C. C., VILLAFRANCA, J. E., APPELT, K., *Structure Fold. Des.* **1999**, 7, 319–330.
- WANG, Z., HARKINS, P. C., ULEVITCH, R. J., HAN, J., COBB, M. H., GOLDSMITH, E. J., *Proc. Natl. Acad. Sci. USA* **1997**, 94, 2327–2332.
- WILSON, K. P., FITZGIBBON, M. J., CARON, P. R., GRIFFITH, J. P., CHEN, W., MCCAFFREY, P. G., CHAMBERS, S. P., SU, M. S., *J. Biol. Chem.* **1996**, 271, 27696–27700.
- WILSON, K. P., MCCAFFREY, P. G., HSIAO, K., PAZHANISAMY, S., GALULLO, V., BEMIS, G. W., FITZGIBBON, M. J., CARON, P. R., MURCKO, M. A., SU, M. S., *Chem. Biol.* **1997**, 4, 423–431.
- ZHANG, F., STRAND, A., ROBBINS, D., COBB, M. H., GOLDSMITH, E. J., *Nature* **1994**, 367, 704–711.
- ABLOOGLU, A. J., TILL, J. H., KIM, K., PARANG, K., COLE, P. A., HUBBARD, S. R., KOHANSKI, R. A., *J. Biol. Chem.* **2000**, 275, 30394–30398.
- BOSSEMEYER, D., ENGH, R. A., KINZEL, V., PONTING, H., HUBER, R., *EMBO J.* **1993**, 12, 849–859.
- COOK, A., LOWE, E. D., CHRYSINA, E. D., SKAMNAKI, V. T., OIKONOMAKOS, N. G., JOHNSON, L. N., *Biochemistry* **2002**, 41, 7301–7311.

- 23 MADHUSUDAN, TRAFNY, E.A., XUONG, N.H., ADAMS, J.A., TEN EYCK, L.F., TAYLOR, S.S., SOWADSKI, J.M., *Protein Sci.* **1994**, 3, 176–187.
- 24 TILL, J.H., ABLOOGLU, A.J., FRANKEL, M., BISHOP, S.M., KOHANSKI, R.A., HUBBARD, S.R., *J. Biol. Chem.* **2001**, 276, 10049–10055.
- 25 JOHNSON, L.N., NOBLE, M.E., OWEN, D.J., *Cell* **1996**, 85, 149–158.
- 26 NAGAR, B., BORNMAN, W.G., PELLICENA, P., SCHINDLER, T., VEACH, D.R., MILLER, W.T., CLARKSON, B., KURIYAN, J., *Cancer Res.* **2002**, 62, 4236–4243.
- 27 SCHINDLER, T., BORNMAN, W., PELLICENA, P., MILLER, W.T., CLARKSON, B., KURIYAN, J., *Science* **2000**, 289, 1938–1942.
- 28 KOBE, B., HEIERHORST, J., FEIL, S.C., PARKER, M.W., BENIAN, G.M., WEISS, K.R., KEMP, B.E., *EMBO J.* **1996**, 15, 6810–6821.
- 29 PARGELLIS, C., TONG, L., CHURCHILL, L., CIRILLO, P.F., GILMORE, T., GRAHAM, A.G., GROB, P.M., HICKEY, E.R., MOSS, N., PAV, S., REGAN, J., *Nat. Struct. Biol.* **2002**, 9, 268–272.
- 30 MOHAMMADI, M., MCMAHON, G., SUN, L., TANG, C., HIRTH, P., YEH, B.K., HUBBARD, S.R., SCHLESSINGER, J., *Science* **1997**, 276, 955–960.
- 31 LAMERS, M.B., ANTON, A.A., HUBBARD, R.E., SCOTT, R.K., WILLIAMS, D.H., *J. Mol. Biol.* **1999**, 285, 713–725.
- 32 SCHINDLER, T., SICHERI, F., PICO, A., GAZIT, A., LEVITZKI, A., KURIYAN, J., *Mol. Cell* **1999**, 3, 639–648.
- 33 SICHERI, F., MOAREFI, I., KURIYAN, J., *Nature* **1997**, 385, 602–609.
- 34 XU, W., HARRISON, S.C., ECK, M.J., *Nature* **1997**, 385, 595–602.
- 35 XU, W., DOSHI, A., LEI, M., ECK, M.J., HARRISON, S.C., *Mol. Cell* **1999**, 3, 629–638.
- 36 ENGH, R.A., GIROD, A., KINZEL, V., HUBER, R., BOSSEMEYER, D., *J. Biol. Chem.* **1996**, 271, 26157–26164.
- 37 FRODIN, M., ANTAL, T.L., DUMMLER, B.A., JENSEN, C.J., DEAK, M., GAMMELTOFT, S., BIONDI, R.M., *EMBO J.* **2002**, 21, 5396–5407.
- 38 YANG, J., CRON, P., GOOD, V.M., THOMPSON, V., HEMMINGS, B.A., BARFORD, D., *Nat. Struct. Biol.* **2002**, 9, 940–944.
- 39 YANG, J., CRON, P., THOMPSON, V., GOOD, V.M., HESS, D., HEMMINGS, B.A., BARFORD, D., *Mol. Cell* **2002**, 9, 1227–1240.
- 40 ZHENG, J., KNIGHTON, D.R., TEN EYCK, L.F., KARLSSON, R., XUONG, N., TAYLOR, S.S., SOWADSKI, J.M., *Biochemistry* **1993**, 32, 2154–2161.
- 41 KNIGHTON, D.R., ZHENG, J.H., TEN EYCK, L.F., XUONG, N.H., TAYLOR, S.S., SOWADSKI, J.M., *Science* **1991**, 253, 414–420.
- 42 SHEWCHUK, L., HASSELL, A., WISELY, B., ROCQUE, W., HOLMES, W., VEAL, J., KUYPER, L.F., *J. Med. Chem.* **2000**, 43, 133–138.
- 43 WANG, Z., CANAGARAJAH, B.J., BOEHM, J.C., KASSISA, S., COBB, M.H., YOUNG, P.R., ABDEL-MEGUID, S., ADAMS, J.L., GOLDSMITH, E.J., *Structure* **1998**, 6, 1117–1128.
- 44 HUSE, M., CHEN, Y.G., MASSAGUE, J., KURIYAN, J., *Cell* **1999**, 96, 425–436.
- 45 WYBENGA-GROOT, L.E., BASKIN, B., ONG, S.H., TONG, J., PAWSON, T., SICHERI, F., *Cell* **2001**, 106, 745–757.
- 46 GOLDBERG, J., NAIRN, A.C., KURIYAN, J., *Cell* **1996**, 84, 875–887.
- 47 HU, S.H., PARKER, M.W., LEI, J.Y., WILCE, M.C., BENIAN, G.M., KEMP, B.E., *Nature* **1994**, 369, 581–584.
- 48 LEI, M., LU, W., MENG, W., PARRINI, M.C., ECK, M.J., MAYER, B.J., HARRISON, S.C., *Cell* **2000**, 102, 387–397.
- 49 PAUTSCH, A., ZOEPHEL, A., AHORN, H., SPEVAK, W., HAUPTMANN, R., NAR, H., *Structure (Camb.)* **2001**, 9, 955–965.
- 50 SULZENBACHER, G., GRUEZ, A., ROIG-ZAMBONI, V., SPINELLI, S., VALENCIA, C., PAGOT, F., VINCENTELLI, R., BIGNON, C., SALOMONI, A., GRISEL, S., MAURIN, D., HUYGHE, C., JOHANSSON, K., GRASSICK, A., ROUSSEL, A., BOURNE, Y., PERRIER, S., MIALLAU, L., CANTAU, P., BLANC, E., GENEVOIS, M., GROSSI, A., ZENATTI, A., CAMPANACCI, V., CAMBILLAU, C., *Acta Crystallogr. D. Biol. Crystallogr.* **2002**, 58, 2109–2115.

- 51 TARRICONE, C., DHAVAN, R., PENG, J., ARECES, L.B., TSAI, L.H., MUSACCHIO, A., *Mol. Cell* **2001**, 8, 657–669.
- 52 JEFFREY, P.D., TONG, L., PAVLETICH, N.P., *Genes Dev.* **2000**, 14, 3115–3125.
- 53 BAX, B., CARTER, P.S., LEWIS, C., GUY, A.R., BRIDGES, A., TANNER, R., PETTMAN, G., MANNIX, C., CULBERT, A.A., BROWN, M.J., SMITH, D.G., REITH, A.D., *Structure (Camb.)* **2001**, 9, 1143–1152.
- 54 MOHAMMADI, M., SCHLESSINGER, J., HUBBARD, S.R., *Cell* **1996**, 86, 577–587.
- 55 IKUTA, M., KAMATA, K., FUKASAWA, K., HONMA, T., MACHIDA, T., HIRAI, H., SUZUKI-TAKAHASHI, I., HAYAMA, T., NISHIMURA, S., *J. Biol. Chem.* **2001**, 276, 27548–27554.
- 56 BINNS, K.L., TAYLOR, P.P., SICHERI, F., PAWSON, T., HOLLAND, S.J., *Mol. Cell Biol.* **2000**, 20, 4791–4805.
- 57 ARNOLD, L.D., DIXON, R., TALANIAN, R., GAZA-BULSECO, G., BUMP, N., SESHADRI, T., BELLAMACINA, C., ALLEN, K., HOFFKEN, W., RIEBEL, A., XU, Y., abstract for *Proceedings of the American Association for Cancer Research* **2002**, 43, 848.
- 58 BELLON, S., FITZGIBBON, M.J., FOX, T., HSIAO, H.M., WILSON, K.P., *Structure Fold. Des.* **1999**, 7, 1057–1065.
- 59 TRAXLER, P., FURET, P., *Pharmacol. Ther.* **1999**, 82, 195–206.
- 60 TOLEDO, L.M., LYDON, N.B., *Structure* **1997**, 5, 1551–1556.
- 61 DE AZEVEDO, W.F., LECLERC, S., MEIJER, L., HAVLICEK, L., STRNAD, M., KIM, S.H., *Eur. J. Biochem.* **1997**, 243, 518–526.
- 62 TRAXLER, P., GREEN, J., METT, H., SEQUIN, U., FURET, P., *J. Med. Chem.* **1999**, 42, 1018–1026.
- 63 NIEFIND, K., PUTTER, M., GUERRA, B., ISINGER, O.G., SCHOMBURG, D., *Nat. Struct. Biol.* **1999**, 6, 1100–1103.
- 64 MEIJER, L., THUNNISSEN, A.M., WHITE, A.W., GARNIER, M., NIKOLIC, M., TSAI, L.H., WALTER, J., CLEVERLEY, K.E., SALINAS, P.C., WU, Y.Z., BIERNAT, J., MANDULKOW, E.M., KIM, S.H., PETTIT, G.R., *Chem. Biol.* **2000**, 7, 51–63.
- 65 HOESSEL, R., LECLERC, S., ENDICOTT, J.A., NOBEL, M.E., LAWRIE, A., TUNNAH, P., LEOST, M., DAMIENS, E., MARIE, D., MARKO, D., NIEDERBERGER, E., TANG, W., EISENBRAND, G., MEIJER, L., *Nat. Cell Biol.* **1999**, 1, 60–67.
- 66 DAVIES, T.G., BENTLEY, J., ARRIS, C.E., BOYLE, F.T., CURTIN, N.J., ENDICOTT, J.A., GIBSON, A.E., GOLDING, B.T., GRIFFIN, R.J., HARDCASTLE, I.R., JEWSEBURY, P., JOHNSON, L.N., MESGUICHE, V., NEWELL, D.R., NOBLE, M.E., TUCKER, J.A., WANG, L., WHITFIELD, H.J., *Nat. Struct. Biol.* **2002**, 9, 745–749.
- 67 BRAMSON, H.N., CORONA, J., DAVIS, S.T., DICKERSON, S.H., EDELSTEIN, M., FRYE, S.V., GAMPE, R.T., JR., HARRIS, P.A., HASSELL, A., HOLMES, W.D., HUNTER, R.N., LACKEY, K.E., LOVEJOY, B., LUZZIO, M.J., MONTANA, V., ROCQUE, W.J., RUSNAK, D., SHEWCHUK, L., VEAL, J.M., WALKER, D.H., KUYPER, L.F., *J. Med. Chem.* **2001**, 44, 4339–4358.
- 68 NARAYANA, N., DILLER, T.C., KOIDE, K., BUNNAGE, M.E., NICOLAOU, K.C., BRUNTON, L.L., XUONG, N.H., TEN EYCK, L.F., TAYLOR, S.S., *Biochemistry* **1999**, 38, 2367–2376.
- 69 KOIDE, K., BUNNAGE, M.E., GOMEZ, P.L., KANTER, J.R., TAYLOR, S.S., BRUNTON, L.L., NICOLAOU, K.C., *Chem. Biol.* **1995**, 2, 601–608.
- 70 HUNENBERGER, P.H., HELMS, V., NARAYANA, N., TAYLOR, S.S., MCCAMMON, J.A., *Biochemistry* **1999**, 38, 2358–2366.
- 71 TONG, L., PAV, S., WHITE, D.M., ROGERS, S., CRANE, K.M., CYWIN, C.L., BROWN, M.L., PARGELLIS, C.A., *Nat. Struct. Biol.* **1997**, 4, 311–316.
- 72 MOHAMMADI, M., FROUM, S., HAMBY, J.M., SCHROEDER, M.C., PANEK, R.L., LU, G.H., ELISEENKOVA, A.V., GREEN, D., SCHLESSINGER, J., HUBBARD, S.R., *EMBO J.* **1998**, 17, 5896–5904.
- 73 ZHU, X., KIM, J.L., NEWCOMB, J.R., ROSE, P.E., STOVER, D.R., TOLEDO, L.M., ZHAO, H., MORGENSTERN, K.A., *Structure Fold. Des.* **1999**, 7, 651–661.
- 74 LIU, Y., BISHOP, A., WITUCKI, L., KRAYBILL, B., SHIMIZU, E., TSIEH, J., UBERSAX, J., BLETHROW, J., MORGAN, D.O., SHOKAT, K.M., *Chem. Biol.* **1999**, 6, 671–678.
- 75 EYERS, P.A., CRAXTON, M., MORRICE, N., COHEN, P., GOEDERT, M., *Chem. Biol.* **1998**, 5, 321–328.
- 76 GUM, R.J., MCLAUGHLIN, M.M., KUMAR, S., WANG, Z., BOWER, M.J., LEE, J.C.,

- ADAMS, J. L., LIVI, G. P., GOLDSMITH, E. J., YOUNG, P. R., *J. Biol. Chem.* **1998**, 273, 15605–15610.
- 77 ADAMS, J. L., BOEHM, J. C., GALLAGHER, T. F., KASSIS, S., WEBB, E. F., HALL, R., SORENSON, M., GARIGIPATI, R., GRISWOLD, D. E., LEE, J. C., *Bioorg. Med. Chem. Lett.* **2001**, 11, 2867–2870.
- 78 BARKER, A. J., GIBSON, K. H., GRUNDY, W., GODFREY, A. A., BARLOW, J. J., HEALY, M. P., WOODBURN, J. R., ASHTON, S. E., CURRY, B. J., SCARLETT, L., HENTHORN, L., RICHARDS, L., *Bioorg. Med. Chem. Lett.* **2001**, 11, 1911–1914.
- 79 HENNEQUIN, L. F., STOKES, E. S., THOMAS, A. P., JOHNSTONE, C., PLE, P. A., OGILVIE, D. J., DUKES, M., WEDGE, S. R., KENDREW, J., CURWEN, J. O., *J. Med. Chem.* **2002**, 45, 1300–1312.
- 80 STAMOS, J., SLIWKOWSKI, M. X., EIGENBROT, C., *J. Biol. Chem.* **2002**, 277, 46265–46272.
- 81 HUBBARD, S. R., TILL, J. H., *Annu. Rev. Biochem.* **2000**, 69, 373–398.
- 82 KIM, S. H., SCHULZE-GAHMEN, U., BRANDSEN, J., AZEVEDO JR., W. F., *Prog. Cell Cycle Res.* **1996**, 2, 137–145.
- 83 JOHNSON, L. N., LEWIS, R. J., *Chem. Rev.* **2001**, 101, 2209–2242.
- 84 PRADE, L., ENGH, R. A., GIROD, A., KINZEL, V., HUBER, R., BOSSEMAYER, D., *Structure* **1997**, 5, 1627–1637.
- 85 KLEBE, G., BOHM, H. J., *J. Recept. Signal. Transduct. Res.* **1997**, 17, 459–473.
- 86 DAVIES, T. G., TUNNAH, P., MEIJER, L., MARKO, D., EISENBRAND, G., ENDICOTT, J. A., NOBLE, M. E., *Structure (Camb.)* **2001**, 9, 389–397.
- 87 DRUKER, B. J., *Oncogene* **2002**, 21, 8541–8546.
- 88 LAWRENCE, D. S., NIU, J., *Pharmacol. Ther.* **1998**, 77, 81–114.
- 89 WOOLFREY, J. R., WESTON, G. S., *Curr. Pharm. Des.* **2002**, 8, 1527–1545.
- 90 LAWRIE, A. M., NOBLE, M. E., TUNNAH, P., BROWN, N. R., JOHNSON, L. N., ENDICOTT, J. A., *Nat. Struct. Biol.* **1997**, 4, 796–801.
- 91 ZHAO, B., BOWER, M. J., McDEVITT, P. J., ZHAO, H., DAVIS, S. T., JOHANSON, K. O., GREEN, S. M., CONCHA, N. O., ZHOU, B. B., *J. Biol. Chem.* **2002**, 277, 46609–46615.
- 92 MENG, W., SWENSON, L. L., FITZGIBBON, M. J., HAYAKAWA, K., TER HAAR, E., BEHRENS, A. E., FULGHUM, J. R., LIPPKE, J. A., *J. Biol. Chem.* **2002**, 277, 37401–37405.
- 93 CHEETHAM, G. M., KNEGTEL, R. M., COLL, J. T., RENWICK, S. B., SWENSON, L., WEBER, P., LIPPKE, J. A., AUSTEN, D. A., *J. Biol. Chem.* **2002**, 277, 42419–42422.
- 94 NOWAKOWSKI, J., CRONIN, C. N., McREE, D. E., KNUTH, M. W., NELSON, C. G., PAVLETICH, N. P., ROGERS, J., SANG, B. C., SCHEIBE, D. N., SWANSON, R. V., THOMPSON, D. A., *Structure (Camb.)* **2002**, 10, 1659–1667.
- 95 SENDEROWICZ, A. M., *Invest New Drugs* **1999**, 17, 313–320.
- 96 JIROUSEK, M. R., GILLIG, J. R., GONZALEZ, C. M., HEATH, W. F., McDONALD, J. H., III, NEEL, D. A., RITO, C. J., SINGH, U., STRAMM, L. E., MELIKIAN-BADALIAN, A., BAEVSKY, M., BALLAS, L. M., HALL, S. E., WINNEROSKI, L. L., FAUL, M. M., *J. Med. Chem.* **1996**, 39, 2664–2671.
- 97 BOLD, G., ALTMANN, K. H., FREI, J., LANG, M., MANLEY, P. W., TRAXLER, P., WIETFIELD, B., BRUGGEN, J., BUCHDUNGER, E., COZENS, R., FERRARI, S., FURET, P., HOFMANN, F., MARTINY-BARON, G., MESTAN, J., ROSEL, J., SILLS, M., STOVER, D., ACEMOGLU, F., BOSS, E., EMMENEGGER, R., LASSER, L., MASSO, E., ROTH, R., SCHLACHTER, C., VETTERLI, W., *J. Med. Chem.* **2000**, 43, 2310–2323.
- 98 BRIDGES, A. J., *Chem. Rev.* **2001**, 101, 2541–2572.
- 99 SUN, L., TRAN, N., LIANG, C., HUBBARD, S., TANG, F., LIPSON, K., SCHRECK, R., ZHOU, Y., McMAHON, G., TANG, C., *J. Med. Chem.* **2000**, 43, 2655–2663.
- 100 ZIMMERMANN, J., CARAVATTI, G., METT, H., MEYER, T., MULLER, M., LYDON, N. B., FABBRO, D., *Arch. Pharm. (Weinheim)* **1996**, 329, 371–376.
- 101 FURET, P., CARAVATTI, G., LYDON, N., PRIESTLE, J. P., SOWADSKI, J. M., TRINKS, U., TRAXLER, P., *J. Comput. Aided Mol. Des.* **1995**, 9, 465–472.
- 102 SCAPIN, G., *Drug Discov. Today* **2002**, 7, 601–611.
- 103 GARCIA-ECHEVERRIA, C., TRAXLER, P., EVANS, D. B., *Med. Res. Rev.* **2000**, 20, 28–57.

- 104 BIT, R.A., DAVIS, P.D., ELLIOTT, L.H., HARRIS, W., HILL, C.H., KEECH, E., KUMAR, H., LAWTON, G., MAW, A., NIXON, J.S., *J. Med. Chem.* **1993**, 36, 21–29.
- 105 TRINKS, U., BUCHDUNGER, E., FURET, P., KUMP, W., METT, H., MEYER, T., MULLER, M., REGENASS, U., RIHS, G., LYDON, N., *J. Med. Chem.* **1994**, 37, 1015–1027.
- 106 TRAXLER, P., BOLD, G., FREI, J., LANG, M., LYDON, N., METT, H., BUCHDUNGER, E., MEYER, T., MUELLER, M., FURET, P., *J. Med. Chem.* **1997**, 40, 3601–3616.
- 107 TRAXLER, P.M., FURET, P., METT, H., BUCHDUNGER, E., MEYER, T., LYDON, N., *J. Med. Chem.* **1996**, 39, 2285–2292.
- 108 FRY, D.W., KRAKER, A.J., MCMICHAEL, A., AMBROSO, L.A., NELSON, J.M., LEOPOLD, W.R., CONNORS, R.W., BRIDGES, A.J., *Science* **1994**, 265, 1093–1095.
- 109 GIBSON, K.H., GRUNDY, W., GODFREY, A.A., WOODBURN, J.R., ASHTON, S.E., CURRY, B.J., SCARLETT, L., BARKER, A.J., BROWN, D.S., *Bioorg. Med. Chem. Lett.* **1997**, 7, 2723–2728.
- 110 PALMER, B.D., TRUMPP-KALLMEYER, S., FRY, D.W., NELSON, J.M., SHOWALTER, H.D., DENNY, W.A., *J. Med. Chem.* **1997**, 40, 1519–1529.
- 111 FRY, D.W., BRIDGES, A.J., DENNY, W.A., DOHERTY, A., GREIS, K.D., HICKS, J.L., HOOK, K.E., KELLER, P.R., LEOPOLD, W.R., LOO, J.A., MCNAMARA, D.J., NELSON, J.M., SHERWOOD, V., SMAILL, J.B., TRUMPP-KALLMEYER, S., DOBRUSIN, E.M., *Proc. Natl. Acad. Sci. USA* **1998**, 95, 12022–12027.
- 112 SCHULZE-GAHMEN, U., BRANDSEN, J., JONES, H.D., MORGAN, D.O., MEIJER, L., VESELY, J., KIM, S.H., *Proteins* **1995**, 22, 378–391.
- 113 TRUMPP-KALLMEYER, S., RUBIN, J.R., HUMBLET, C., HAMBY, J.M., HOLLIS SHOWALTER, H.D., *J. Med. Chem.* **1998**, 41, 1752–1763.
- 114 BARVIAN, M., BOSCHELLI, D.H., COSSROW, J., DOBRUSIN, E., FATTAEY, A., FRITSCH, A., FRY, D., HARVEY, P., KELLER, P., GARRETT, M., LA, F., LEOPOLD, W., MCNAMARA, D., QUIN, M., TRUMPP-KALLMEYER, S., TOOGOOD, P., WU, Z., ZHANG, E., *J. Med. Chem.* **2000**, 43, 4606–4616.
- 115 DE, A.W., JR., MUELLER-DIECKMANN, H.J., SCHULZE-GAHMEN, U., WORLAND, P.J., SAUSVILLE, E., KIM, S.H., *Proc. Natl. Acad. Sci. USA* **1996**, 93, 2735–2740.
- 116 DAVIES, T.G., PRATT, D.J., ENDICOTT, J.A., JOHNSON, L.N., NOBLE, M.E., *Pharmacol. Ther.* **2002**, 93, 125–133.
- 117 PARANG, K., TILL, J.H., ABLOOGLU, A.J., KOHANSKI, R.A., HUBBARD, S.R., COLE, P.A., *Nat. Struct. Biol.* **2001**, 8, 37–41.
- 118 PARANG, K., COLE, P.A., *Pharmacol. Ther.* **2002**, 93, 145–157.
- 119 HONMA, T., HAYASHI, K., AOYAMA, T., HASHIMOTO, N., MACHIDA, T., FUKASAWA, K., IWAMA, T., IKEURA, C., IKUTA, M., SUZUKI-TAKAHASHI, I., IWASAWA, Y., HAYAMA, T., NISHIMURA, S., MORISHIMA, H., *J. Med. Chem.* **2001**, 44, 4615–4627.
- 120 FURET, P., MEYER, T., STRAUSS, A., RACCUGLIA, S., RONDEAU, J.M., *Bioorg. Med. Chem. Lett.* **2002**, 12, 221–224.
- 121 FURET, P., MEYER, T., MITTL, P., FRETZ, H., *J. Comput. Aided Mol. Des.* **2001**, 15, 489–495.
- 122 GRAY, N.S., WODICKA, L., THUNNISSEN, A.M., NORMAN, T.C., KWON, S., ESPINOZA, F.H., MORGAN, D.O., BARNES, G., LECLERC, S., MEIJER, L., KIM, S.H., LOCKHART, D.J., SCHULTZ, P.G., *Science* **1998**, 281, 533–538.
- 123 ADAMS, J.L., LEE, D., *Curr. Opin. Drug Discov. Dev.* **1999**, 2, 96–109.
- 124 CALLAHAN, J.F., BURGESS, J.L., FORNWALD, J.A., GASTER, L.M., HARLING, J.D., HARRINGTON, F.P., HEER, J., KWON, C., LEHR, R., MATHUR, A., OLSON, B.A., WEINSTOCK, J., LAPING, N.J., *J. Med. Chem.* **2002**, 45, 999–1001.
- 125 COLLETTI, S.L., FRIE, J.L., DIXON, E.C., SINGH, S.B., CHOI, B.K., SCAPIN, G., FITZGERALD, C.E., KUMAR, S., NICHOLS, E.A., O'KEEFE, S.J., O'NEILL, E.A., PORTER, G., SAMUEL, K., SCHMATZ, D.M., SCHWARTZ, C.D., SHOOP, W.L., THOMPSON, C.M., THOMPSON, J.E., WANG, R., WOODS, A., ZALLER, D.M., DOHERTY, J.B., *J. Med. Chem.* **2003**, 46, 349–352.
- 126 STELMACH, J.E., LIU, L., PATEL, S.B., PIVNICHNY, J.V., SCAPIN, G., SINGH, S., HOP, C.E., WANG, Z., STRAUSS, J.R., CAMERON, P.M., NICHOLS, E.A., O'KEEFE, S.J., O'NEILL, E.A., SCHMATZ, D.M., SCHWARTZ, C.D., THOMPSON, C.M., ZALLER, D.M., DOHERTY, J.B., *Bioorg. Med. Chem. Lett.* **2003**, 13, 277–280.

- 127 SHUKER, S. B., HAJDUK, P. J., MEADOWS, R. P., FESIK, S. W., *Science* **1996**, 274, 1531–1534.
- 128 CODY, W. L., LIN, Z., PANEK, R. L., ROSE, D. W., RUBIN, J. R., *Curr. Pharm. Des.* **2000**, 6, 59–98.
- 129 PANG, L., SAWADA, T., DECKER, S. J., SALTIEL, A. R., *J. Biol. Chem.* **1995**, 270, 13585–13588.
- 130 HERRERA, R., SEBOLT-LEOPOLD, J. S., *Trends Mol. Med.* **2002**, 8, S27–S31.