

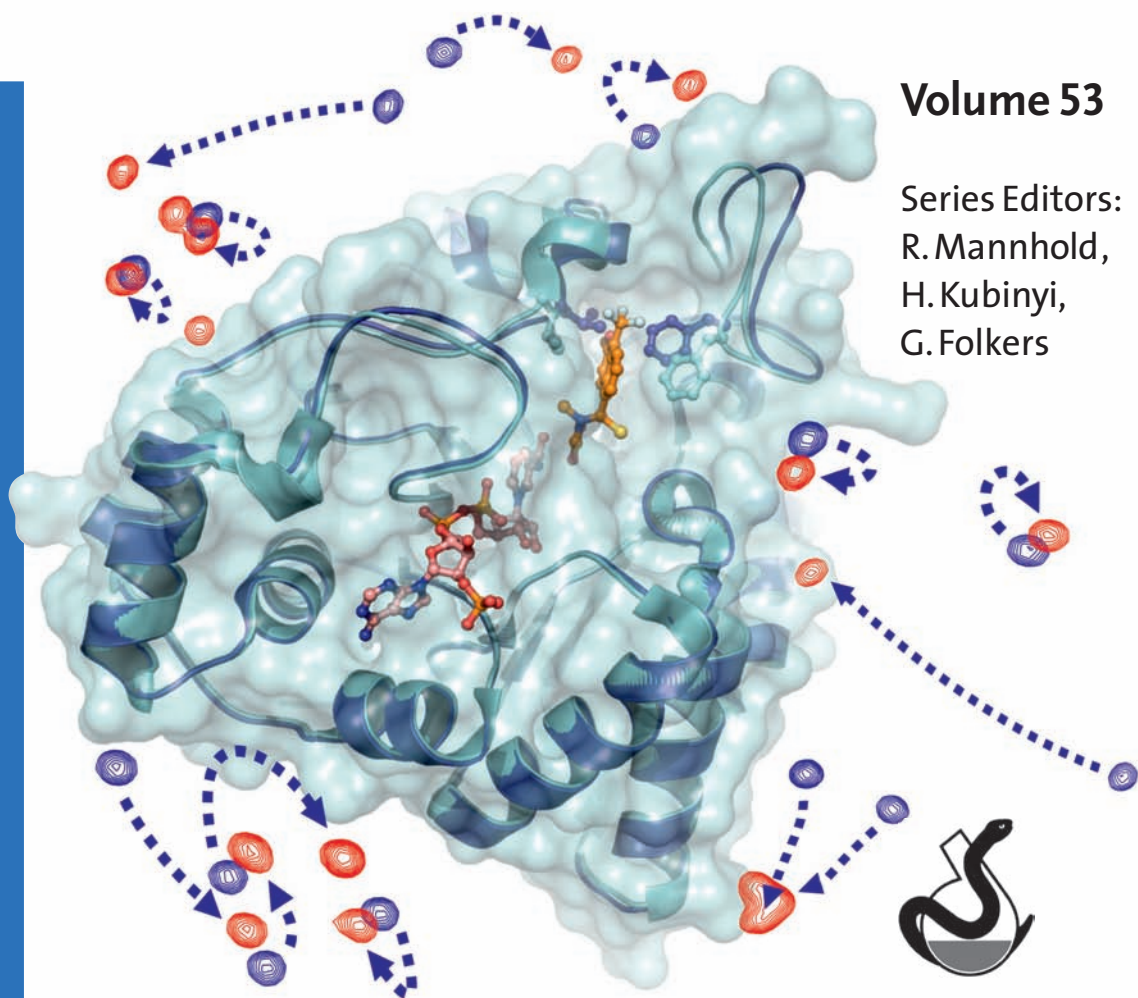
Edited by Holger Gohlke

 WILEY-VCH

# Protein-Ligand Interactions

**Volume 53**

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





*Edited by*  
*Holger Gohlke*

**Protein-Ligand Interactions**

# ***Methods and Principles in Medicinal Chemistry***

Edited by R. Mannhold, H. Kubinyi, G. Folkers

Editorial Board

H. Buschmann, H. Timmerman, H. van  
de Waterbeemd, T. Wieland

## ***Previous Volumes of this Series:***

Kappe, C. Oliver / Stadler, Alexander /  
Dallinger, Doris

### **Microwaves in Organic and Medicinal Chemistry Second, Completely Revised and Enlarged Edition**

2012

ISBN: 978-3-527-33185-7

Vol. 52

Smith, Dennis A. / Allerton, Charlotte /  
Kalgutkar, Amit S. / van de Waterbeemd, Han /  
Walker, Don K.

### **Pharmacokinetics and Metabolism in Drug Design Third, Revised and Updated Edition**

2012

ISBN: 978-3-527-32954-0

Vol. 51

De Clercq, Erik (Ed.)

### **Antiviral Drug Strategies**

2011

ISBN: 978-3-527-32696-9

Vol. 50

Klebl, Bert / Müller, Gerhard / Hamacher, Michael  
(Eds.)

### **Protein Kinases as Drug Targets**

2011

ISBN: 978-3-527-31790-5

Vol. 49

Sotriffer, Christoph (Ed.)

### **Virtual Screening Principles, Challenges, and Practical Guidelines**

2011

ISBN: 978-3-527-32636-5

Vol. 48

Rautio, Jarkko (Ed.)

### **Prodrugs and Targeted Delivery Towards Better ADME Properties**

2011

ISBN: 978-3-527-32603-7

Vol. 47

Smit, Martine J. / Lira, Sergio A. / Leurs, Rob (Eds.)

### **Chemokine Receptors as Drug Targets**

2011

ISBN: 978-3-527-32118-6

Vol. 46

Ghosh, Arun K. (Ed.)

### **Aspartic Acid Proteases as Therapeutic Targets**

2010

ISBN: 978-3-527-31811-7

Vol. 45

Ecker, Gerhard F. / Chiba, Peter (Eds.)

### **Transporters as Drug Carriers Structure, Function, Substrates**

2009

ISBN: 978-3-527-31661-8

Vol. 44

Faller, Bernhard / Urban, Laszlo (Eds.)

### **Hit and Lead Profiling**

### **Identification and Optimization of Drug-like Molecules**

2009

ISBN: 978-3-527-32331-9

Vol. 43

*Edited by Holger Gohlke*

## **Protein-Ligand Interactions**



**WILEY-  
VCH**

WILEY-VCH Verlag GmbH & Co. KGaA

#### Series Editors

**Prof. Dr. Raimund Mannhold**

Molecular Drug Research Group  
Heinrich-Heine-Universität  
Universitätsstrasse 1  
40225 Düsseldorf  
Germany  
mannhold@uni-duesseldorf.de

**Prof. Dr. Hugo Kubinyi**

Donnersbergstrasse 9  
67256 Weisenheim am Sand  
Germany  
kubinyi@t-online.de

**Prof. Dr. Gerd Folkers**

Collegium Helveticum  
STW/ETH Zurich  
8092 Zurich  
Switzerland  
folkers@collegium.ethz.ch

#### Volume Editor

**Prof. Dr. Holger Gohlke**

Heinrich-Heine-Universität  
Inst.für Pharmaz. + Med.Chemie  
Universitätsstr. 1  
40225 Düsseldorf  
Germany

#### Cover Description

Aldose reductase bound to NADP (PDB code: 1ads; dark blue) as well as bound to NADP and the inhibitor tolrestat (PDB code: 2fzd; light blue). The protein shows pronounced movements of Phe122 and Leu300 to accommodate hydrophobic parts of the inhibitor. Such movements can be detected by perturbations of NMR chemical shifts, as schematically shown around the protein.

All books published by **Wiley-VCH** are carefully produced. Nevertheless, authors, editors, and publisher do not warrant the information contained in these books, including this book, 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.

**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 the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at <http://dnb.d-nb.de>.

© 2012 Wiley-VCH Verlag & Co. KGaA,  
Boschstr. 12, 69469 Weinheim, Germany

All rights reserved (including those of translation into 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 a 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.

**Print ISBN:** 978-3-527-32966-3

**ePDF ISBN:** 978-3-527-64597-8

**ePub ISBN:** 978-3-527-64596-1

**mobi ISBN:** 978-3-527-64595-4

**oBook ISBN:** 978-3-527-64594-7

**Cover Design** Grafik-Design Schulz, Fußgönheim

**Typesetting** Thomson Digital, Noida, India

## Contents

List of Contributors XIII

Preface XVII

A Personal Foreword XIX

### Part I Binding Thermodynamics 1

#### 1 Statistical Thermodynamics of Binding and Molecular Recognition Models 3

*Kim A. Sharp*

- 1.1 Introductory Remarks 3
- 1.2 The Binding Constant and Free Energy 3
- 1.3 A Statistical Mechanical Treatment of Binding 4
  - 1.3.1 Binding in a Square Well Potential 6
  - 1.3.2 Binding in a Harmonic Potential 7
- 1.4 Strategies for Calculating Binding Free Energies 9
  - 1.4.1 Direct Association Simulations 9
  - 1.4.2 The Quasi-Harmonic Approximation 10
  - 1.4.3 Estimation of Entropy Contributions to Binding 11
  - 1.4.4 The Molecule Mechanics Poisson–Boltzmann Surface Area Method 13
  - 1.4.5 Thermodynamic Work Methods 14
  - 1.4.6 Ligand Decoupling 15
  - 1.4.7 Linear Interaction Methods 15
  - 1.4.8 Salt Effects on Binding 16
  - 1.4.9 Statistical Potentials 17
  - 1.4.10 Empirical Potentials 18
- References 19

#### 2 Some Practical Rules for the Thermodynamic Optimization of Drug Candidates 23

*Ernesto Freire*

- 2.1 Engineering Binding Contributions 25
- 2.2 Eliminating Unfavorable Enthalpy 25
- 2.3 Improving Binding Enthalpy 26

2.4	Improving Binding Affinity	27
2.5	Improving Selectivity	28
2.6	Thermodynamic Optimization Plot	28
	Acknowledgments	30
	References	31
<b>3</b>	<b>Enthalpy–Entropy Compensation as Deduced from Measurements of Temperature Dependence</b>	<b>33</b>
	<i>Athel Cornish-Bowden</i>	
3.1	Introduction	33
3.2	The Current Status of Enthalpy–Entropy Compensation	34
3.3	Measurement of the Entropy and Enthalpy of Activation	34
3.4	An Example	35
3.5	The Compensation Temperature	38
3.6	Effect of High Correlation on Estimates of Entropy and Enthalpy	39
3.7	Evolutionary Considerations	40
3.8	Textbooks	40
	References	42
<b>Part II</b>	<b>Learning from Biophysical Experiments</b>	<b>45</b>
<b>4</b>	<b>Interaction Kinetic Data Generated by Surface Plasmon Resonance Biosensors and the Use of Kinetic Rate Constants in Lead Generation and Optimization</b>	<b>47</b>
	<i>U. Helena Danielson</i>	
4.1	Background	47
4.2	SPR Biosensor Technology	48
4.2.1	Principles	48
4.2.2	Sensitivity	49
4.2.3	Kinetic Resolution	50
4.2.4	Performance for Drug Discovery	51
4.3	From Interaction Models to Kinetic Rate Constants and Affinity	53
4.3.1	Determination of Interaction Kinetic Rate Constants	53
4.3.2	Determination of Affinities	54
4.3.3	Steady-State Analysis versus Analysis of Complete Sensorgrams	54
4.4	Affinity versus Kinetic Rate Constants for Evaluation of Interactions	55
4.5	From Models to Mechanisms	56
4.5.1	Irreversible Interactions	57
4.5.2	Induced Fit	57
4.5.3	Conformational Selection	58
4.5.4	Unified Model for Dynamic Targets	58
4.5.5	Heterogeneous Systems/Parallel Reactions	59
4.5.6	Mechanism-Based Inhibitors	60
4.5.7	Multiple Binding Sites and Influence of Cofactors	61
4.6	Structural Information	61



4.7	The Use of Kinetic Rate Constants in Lead Generation and Optimization	62
4.7.1	Structure–Kinetic Relationships	62
4.7.2	Selectivity/Specificity and Resistance	63
4.7.3	Chemodynamics	63
4.7.4	Thermodynamics	64
4.8	Designing Compounds with Optimal Properties	65
4.8.1	Correlation between Kinetic and Thermodynamic Parameters and Pharmacological Efficacy	65
4.8.2	Structural Modeling	66
4.9	Conclusions	67
	Acknowledgments	67
	References	67
<b>5</b>	<b>NMR Methods for the Determination of Protein–Ligand Interactions</b>	<b>71</b>
	<i>Bernd W. Koenig, Sven Schünke, Matthias Stoldt, and Dieter Willbold</i>	
5.1	Experimental Parameters from NMR	72
5.2	Aspects of Protein–Ligand Interactions That Can Be Addressed by NMR	77
5.2.1	Detection and Verification of Ligand Binding	77
5.2.2	Interaction Site Mapping	78
5.2.3	Interaction Models and Binding Affinity	80
5.2.4	Molecular Recognition	81
5.2.5	Structure of Protein–Ligand Complexes	82
5.3	Ligand-Induced Conformational Changes of a Cyclic Nucleotide Binding Domain	84
5.4	Ligand Binding to GABARAP Binding Site and Affinity Mapping	86
5.5	Transient Binding of Peptide Ligands to Membrane Proteins	88
	References	90
<b>Part III</b>	<b>Modeling Protein–Ligand Interactions</b>	<b>99</b>
<b>6</b>	<b>Polarizable Force Fields for Scoring Protein–Ligand Interactions</b>	<b>101</b>
	<i>Jiajing Zhang, Yue Shi, and Pengyu Ren</i>	
6.1	Introduction and Overview	101
6.2	AMOEBA Polarizable Potential Energy Model	102
6.2.1	Bond, Angle, and Cross-Energy Terms	102
6.2.2	Torsional Energy Term	103
6.2.3	Van der Waals Interactions	103
6.2.4	Permanent Electrostatic Interactions	103
6.2.5	Electronic Polarization	104
6.2.6	Polarization Energy	105
6.3	AMOEBA Explicit Water Simulation Applications	106
6.3.1	Small-Molecule Hydration Free Energy Calculations	106
6.3.2	Ion Solvation Thermodynamics	108

6.3.3	Binding Free Energy of Trypsin and Benzamidine Analogs	110
6.4	Implicit Solvent Calculation Using AMOEBA Polarizable Force Field	113
6.5	Conclusions and Future Directions	115
	References	116
<b>7</b>	<b>Quantum Mechanics in Structure-Based Ligand Design</b>	<b>121</b>
	<i>Pär Söderhjelm, Samuel Genheden, and Ulf Ryde</i>	
7.1	Introduction	121
7.2	Three MM-Based Methods	122
7.3	QM-Based Force Fields	123
7.4	QM Calculations of Ligand Binding Sites	125
7.5	QM/MM Calculations	126
7.6	QM Calculations of Entire Proteins	127
7.6.1	Linear Scaling Methods	128
7.6.2	Fragmentation Methods	129
7.7	Concluding Remarks	133
	Acknowledgments	134
	References	134
<b>8</b>	<b>Hydrophobic Association and Volume-Confined Water Molecules</b>	<b>145</b>
	<i>Riccardo Baron, Piotr Setny, and J. Andrew McCammon</i>	
8.1	Introduction	145
8.2	Water as a Whole in Hydrophobic Association	146
8.2.1	Background	146
8.2.2	Computational Modeling of Hydrophobic Association	150
8.2.2.1	Explicit versus Implicit Solvent: Is the Computational Cost Motivated?	152
8.3	Confined Water Molecules in Protein–Ligand Binding	153
8.3.1	Protein Hydration Sites	153
8.3.2	Thermodynamics of Volume-Confined Water Localization	154
8.3.3	Computational Modeling of Volume-Confined Water Molecules	156
8.3.4	Identifying Hydration Sites	158
8.3.5	Water in Protein–Ligand Docking	160
	Acknowledgments	161
	References	161
<b>9</b>	<b>Implicit Solvent Models and Electrostatics in Molecular Recognition</b>	<b>171</b>
	<i>Tyler Luchko and David A. Case</i>	
9.1	Introduction	171
9.2	Poisson–Boltzmann Methods	173
9.3	The Generalized Born Model	175
9.4	Reference Interaction Site Model of Molecular Solvation	176
9.5	Applications	179

9.5.1	The “MM-PBSA” Model	180
9.5.2	Rescoring Docking Poses	182
9.5.3	MM/3D-RISM	182
	Acknowledgments	185
	References	185
<b>10</b>	<b>Ligand and Receptor Conformational Energies</b>	<b>191</b>
	<i>Themis Lazaridis</i>	
10.1	The Treatment of Ligand and Receptor Conformational Energy in Various Theoretical Formulations of Binding	191
10.1.1	Double Decoupling Free Energy Calculations	192
10.1.2	MM-PB(GB)SA	192
10.1.3	Mining Minima	193
10.1.4	Free Energy Functional Approach	194
10.1.5	Linear Interaction Energy Methods	195
10.1.6	Scoring Functions	196
10.2	Computational Results on Ligand Conformational Energy	196
10.3	Computational Results on Receptor Conformational Energy	198
10.4	Concluding Remarks	199
	Acknowledgments	199
	References	199
<b>11</b>	<b>Free Energy Calculations in Drug Lead Optimization</b>	<b>207</b>
	<i>Thomas Steinbrecher</i>	
11.1	Modern Drug Design	207
11.1.1	<i>In Silico</i> Drug Design	210
11.2	Free Energy Calculations	212
11.2.1	Considerations for Accurate and Precise Results	215
11.3	Example Protocols and Applications	217
11.3.1	Example 1: Disappearing an Ion	219
11.3.2	Example 2: Relative Ligand Binding Strengths	221
11.3.3	Applications	223
11.4	Discussion	226
	References	227
<b>12</b>	<b>Scoring Functions for Protein–Ligand Interactions</b>	<b>237</b>
	<i>Christoph Sotriffer</i>	
12.1	Introduction	237
12.2	Scoring Protein–Ligand Interactions: What for and How to?	237
12.2.1	Knowledge-Based Scoring Functions	238
12.2.2	Force Field-Based Methods	240
12.2.3	Empirical Scoring Functions	242
12.2.4	Further Approaches	244
12.3	Application of Scoring Functions: What Is Possible and What Is Not?	246

12.4	Thermodynamic Contributions and Intermolecular Interactions: Which Are Accounted for and Which Are Not?	248
12.5	Conclusions or What Remains to be Done and What Can be Expected?	254
	Acknowledgments	255
	References	255

**Part IV Challenges in Molecular Recognition 265**

<b>13</b>	<b>Druggability Prediction</b>	267
	<i>Daniel Alvarez-Garcia, Jesus Seco, Peter Schmidtke, and Xavier Barril</i>	
13.1	Introduction	267
13.2	Druggability: Ligand Properties	267
13.3	Druggability: Ligand Binding	268
13.4	Druggability Prediction by Protein Class	270
13.5	Druggability Predictions: Experimental Methods	270
13.5.1	High-Throughput Screening	270
13.5.2	Fragment Screening	271
13.5.3	Multiple Solvent Crystallographic Screening	272
13.6	Druggability Predictions: Computational Methods	272
13.6.1	Cavity Detection Algorithms	272
13.6.2	Empirical Models	273
13.6.2.1	Training Sets	273
13.6.2.2	Applicability and Prediction Performance	274
13.6.3	Physical Chemistry Predictions	275
13.7	A Test Case: PTP1B	276
13.8	Outlook and Concluding Remarks	278
	References	278
<b>14</b>	<b>Embracing Protein Plasticity in Ligand Docking</b>	283
	<i>Manuel Rueda and Ruben Abagyan</i>	
14.1	Introduction	283
14.2	Docking by Sampling Internal Coordinates	284
14.3	Fast Docking to Multiple Receptor Conformations	285
14.4	Single Receptor Conformation	285
14.5	Multiple Receptor Conformations	286
14.5.1	Exploiting Existing Experimental Conformational Diversity	286
14.5.2	Selecting “Important” Conformations	288
14.5.3	Generating <i>In Silico</i> Models	288
14.6	Improving Poor Homology Models of the Binding Pocket	289
14.7	State of the Art: GPCR Dock 2010 Modeling and Docking Assessment	290
14.8	Conclusions and Outlook	290
	Acknowledgments	292
	References	292

<b>15</b>	<b>Prospects of Modulating Protein–Protein Interactions</b>	<b>295</b>
	<i>Shijun Zhong, Taiji Oashi, Wenbo Yu, Paul Shapiro, and Alexander D. MacKerell Jr.</i>	
15.1	Introduction	295
15.2	Thermodynamics of Protein–Protein Interactions	297
15.3	CADD Methods for the Identification and Optimization of Small-Molecule Inhibitors of PPIs	298
15.3.1	Identifying Inhibitors of PPIs Using SBDD	299
15.3.1.1	Protein Structure Preparation	299
15.3.1.2	Binding Site Identification	300
15.3.1.3	Virtual Chemical Database	302
15.3.1.4	Virtual Screening of Compound Database	302
15.3.1.5	Rescoring	304
15.3.1.6	Final Selection of Ligands for Experimental Assay	306
15.3.2	Lead Optimization	307
15.3.2.1	Ligand-Based Optimization	307
15.3.2.2	Computation of Binding Free Energy	308
15.4	Examples of CADD Applied to PPIs	308
15.4.1	ERK	309
15.4.2	BCL6	311
15.4.3	S100B	313
15.4.4	p56Lck Kinase SH2 Domain	313
15.5	Summary	315
	Acknowledgments	315
	References	315
	<b>Index</b>	<b>331</b>



## List of Contributors

### **Ruben Abagyan**

University of California, San Diego  
Skaggs School of Pharmacy and  
Pharmaceutical Sciences  
9500 Gilman Drive  
La Jolla, CA 92093  
USA

and

University of California, San Diego  
San Diego Supercomputer Center  
La Jolla, CA 92093  
USA

and

Molsoft LLC  
11199 Sorrento Valley Road, S209  
San Diego, CA 92121  
USA

### **Daniel Alvarez-Garcia**

Universitat de Barcelona  
Facultat de Farmàcia  
Departament de Físicoquímica  
Av. Joan XXIII s/n  
08028 Barcelona  
Spain

and

Institut de Biomedicina de la Universitat  
de Barcelona (IBUB)  
Barcelona  
Spain

### **Riccardo Baron**

The University of Utah  
Department of Medicinal Chemistry  
College of Pharmacy

and

The Henry Eyring Center for Theoretical  
Chemistry  
Salt Lake City  
UT 84112-5820  
USA  
Email: r.baron@utah.edu

### **Xavier Barril**

Universitat de Barcelona  
Facultat de Farmàcia  
Departament de Físicoquímica  
Av. Joan XXIII s/n  
08028 Barcelona  
Spain

and

Institut de Biomedicina de la Universitat  
de Barcelona (IBUB)  
Barcelona  
Spain

and

Catalan Institution for Research and  
Advanced Studies (ICREA)  
Passeig Lluís Companys 23  
08010 Barcelona  
Spain

**David A. Case**

Rutgers University  
Department of Chemistry and Chemical  
Biology and BioMaPS Institute  
610 Taylor Road  
Piscataway, NJ 08854-8087  
USA

**Athel Cornish-Bowden**

CNRS  
Bioénergétique et Ingénierie des  
Protéines  
B.P. 71, 31 chemin Joseph-Aiguier  
13402 Marseille Cedex 20  
France

**U. Helena Danielson**

Uppsala University  
Department of Biochemistry and  
Organic Chemistry  
BMC, Box 576  
75123 Uppsala  
Sweden

**Ernesto Freire**

Johns Hopkins University  
Department of Biology  
114A Mudd Hall, 3400 North Charles  
Street  
Baltimore  
MD 21218-2685  
USA

**Samuel Genheden**

Lund University  
Department of Theoretical Chemistry  
Chemical Centre  
P.O. Box 124  
221 00 Lund  
Sweden

**Bernd W. Koenig**

Forschungszentrum Jülich  
IBI-2  
52425 Jülich  
Germany

**Themis Lazaridis**

City University of New York  
City College of New York  
Department of Chemistry  
138th Street & Convent Avenue  
New York, NY 10031  
USA

**Tyler Luchko**

Rutgers University  
Department of Chemistry and Chemical  
Biology and BioMaPS Institute  
610 Taylor Road  
Piscataway, NJ 08854-8087  
USA

**Alexander D. MacKerell Jr.**

University of Maryland Baltimore  
School of Pharmacy  
Department of Pharmaceutical Sciences  
20 Penn Street  
Baltimore, MD 21201  
USA

**J. Andrew McCammon**

University of California  
Department of Chemistry and  
Biochemistry  
Center for Theoretical Biological  
Physics  
La Jolla  
CA 92093-0365  
USA



**Tajji Oashi**

University of Maryland Baltimore  
 School of Pharmacy  
 Department of Pharmaceutical Sciences  
 20 Penn Street  
 Baltimore, MD 21201  
 USA

**Pengyu Ren**

The University of Texas at Austin  
 Department of Biomedical Engineering  
 BME 5.202M, 1 University Station,  
 C0800  
 Austin  
 TX 78712-1062  
 USA

**Manuel Rueda**

University of California, San Diego  
 Skaggs School of Pharmacy and  
 Pharmaceutical Sciences  
 9500 Gilman Drive  
 La Jolla, CA 92093  
 USA

**Ulf Ryde**

Lund University  
 Department of Theoretical Chemistry  
 Chemical Centre  
 P.O. Box 124  
 221 00 Lund  
 Sweden

**Peter Schmidtke**

Universitat de Barcelona  
 Facultat de Farmàcia  
 Departament de Físicoquímica  
 Av. Joan XXIII s/n  
 08028 Barcelona  
 Spain

and

Institut de Biomedicina de la Universitat  
 de Barcelona (IBUB)  
 Barcelona  
 Spain

**Sven Schünke**

Forschungszentrum Jülich  
 IBI-2  
 52425 Jülich  
 Germany

**Jesus Seco**

Universitat de Barcelona  
 Facultat de Farmàcia  
 Departament de Físicoquímica  
 Av. Joan XXIII s/n  
 08028 Barcelona  
 Spain

and

Institut de Biomedicina de la Universitat  
 de Barcelona (IBUB)  
 Barcelona  
 Spain

**Piotr Setny**

Technical University Munich  
 Department of Physics  
 Munich  
 Germany

**Paul Shapiro**

University of Maryland Baltimore  
 School of Pharmacy  
 Department of Pharmaceutical Sciences  
 20 Penn Street  
 Baltimore, MD 21201  
 USA

**Kim A. Sharp**

University of Pennsylvania  
Department of Biochemistry and  
Biophysics  
37th and Hamilton Walk  
Philadelphia  
PA 19104-6059  
USA

**Yue Shi**

The University of Texas at Austin  
Department of Biomedical Engineering  
BME 5.202M, 1 University Station,  
C0800  
Austin  
TX 78712-1062  
USA

**Pär Söderhjelm**

ETH Zürich  
Department of Chemistry and Applied  
Biosciences  
Computational Science  
Via Giuseppe Buffi 13  
6900 Lugano  
Switzerland

**Christoph Sotriffer**

University of Würzburg  
Institute of Pharmacy and Food  
Chemistry  
Department of Pharmaceutical  
Chemistry  
Am Hubland  
97074 Würzburg  
Germany

**Thomas Steinbrecher**

Karlsruher Institut für Technologie  
Institut für Physikalische Chemie  
Gebäude 30.45, Kaiserstr. 12  
76131 Karlsruhe  
Germany

**Matthias Stoldt**

Forschungszentrum Jülich  
IBI-2  
52425 Jülich  
Germany

**Dieter Willbold**

Forschungszentrum Jülich  
IBI-2  
52425 Jülich  
Germany

**Wenbo Yu**

University of Maryland Baltimore  
School of Pharmacy  
Department of Pharmaceutical Sciences  
20 Penn Street  
Baltimore, MD 21201  
USA

**Jiajing Zhang**

The University of Texas at Austin  
Department of Biomedical Engineering  
BME 5.202M, 1 University Station,  
C0800  
Austin  
TX 78712-1062  
USA

**Shijun Zhong**

School of Life Science and Technology  
Dalian University of Technology  
2 Linggong Road, Dalian  
Liaoning Province, 116024, China  
Phone: (086)0411-84707913  
Email: sjzhong@dlut.edu.cn,  
sjzhong@gmail.com

## Preface

*“Um ein Bild zu gebrauchen, will ich sagen, dass Enzym und Glucosid wie Schloss und Schlüssel zueinander passen müssen, um eine chemische Wirkung aufeinander ausüben zu können”* [To use a picture, I would like to say that enzyme and glucoside have to fit like a lock and a key, in order to exert a chemical action on each other] wrote Emil Fischer in 1894, to illustrate his concept on protein-ligand interactions. Well, our picture of the protein-ligand interaction has developed further. Instead of a rigid fit, the concepts of induced fit and, later, flexible fit were formulated. Indeed, we have to understand the interaction like a handshake, where the one partner adapts to the other, in a mutual fit. Of course, this accommodation should not waste too much conformational energy, otherwise the affinity of the ligand would be significantly reduced. Correspondingly experience shows that compounds where the bioactive conformation is fixed or at least stabilized are most often high-affinity ligands. On the other hand, ligands that cannot achieve such a conformation will have no affinity at all. However, these geometric requirements are only part of the story. In addition, there must be complementary properties - *similia similibus*. Lipophilic groups should find their counterpart, charges or partial charges should have opposite signs, and hydrogen bond donor and acceptor groups should find together. The better this complementarity, the higher will be the affinity, provided that there are no steric clashes. Now, whereas this is a correct description of the requirements of a protein-ligand interaction, it is a very simple one.

Holger Gohlke assembled a team of leading experts in this field to describe not only the thermodynamics of binding but also the underlying biophysical approaches. The major part of the book is devoted to the discussion of factors that are responsible for the intermolecular interactions. Finally some challenges in molecular recognition are discussed. In a logical and didactic way, this volume is organized in four sections. The three introductory chapters review statistical thermodynamics of binding and molecular recognition models, practical rules for the thermodynamic optimization of drug candidates, and the puzzling concept of enthalpy-entropy compensation, as deduced from measurements of temperature dependence. A section on the impact of biophysical experiments focuses, in particular, on interaction kinetic data generated by surface plasmon resonance biosensors as well as NMR methods for the determination of protein-ligand interactions. The central and most comprehensive section is dedicated to aspects of modeling

protein-ligand interactions including polarizable force fields, quantum mechanics in structure-based ligand design, the role of water in hydrophobic association, implicit solvation models and electrostatics in molecular recognition, conformational aspects, free energy calculations in drug lead optimization, as well as scoring functions for protein-ligand interactions. The final section on challenges for protein-ligand interaction modelling considers druggability prediction, protein plasticity, and protein-protein interactions.

The series editors are grateful to Holger Gohlke for his enthusiasm to organize this volume and to work with such a selection of excellent authors. We believe that this book adds a fascinating new facet to our book series on “Methods and Principles in Medicinal Chemistry”. Last, but not least we thank the publisher Wiley-VCH, in particular Frank Weinreich and Heike Nöthe, for their valuable contributions to this project and the entire series.

February 2012  
Düsseldorf  
Weisenheim am Sand  
Zürich

*Raimund Mannhold*  
*Hugo Kubinyi*  
*Gerd Folkers*

## A Personal Foreword

Ever since I started my scientific work I have been fascinated by the questions what makes two molecules bind to each other and how can one make use of this knowledge to modulate biological processes: After all, “*corpora non agunt nisi fixata*”, as Paul Ehrlich put it almost 100 years ago. Ehrlich’s statement that “bodies do not act if they are not bound” is strikingly exemplified by about 130,000 binary interactions in the human protein-protein “interactome” [1]. Along the same lines and more oriented towards the topic of this book, the famous wall chart “Biochemical Pathways” introduced by Gerhard Michal [2] is a vivid picture in the eyes of a life science scientist, with its comprehensive view on metabolic pathways and cellular and molecular processes, particularly involving interactions between proteins and endogenous small-molecules.

For a xenobiotic ligand to exert an influence on an organism, it must bind to a biological target, too. While this statement seems unspectacular nowadays, it still has far-reaching consequences because it provides a valuable handle to explain and predict *biological activity*, both beneficial and detrimental, in terms of *affinity*, a well-defined thermodynamic property, of a ligand towards a (or multiple) targets. In fact, the fields of medicinal chemistry and drug design have made use of different variations of this theme over time depending on which type of information about binding was available in each case.

From an *inductive point of view*, already since the very early days of modern chemistry [3] structures of ligands and, hence, their binding properties, have been correlated to activities. Further milestones on this route were the establishment of quantitative structure-activity relationships and the comparison of ligands based on concepts of similarity or dissimilarity of shape and chemical properties. From a *deductive point of view*, the above theme summons medicinal chemists to look at biological activity from the perspective of processes and contributions that lead to binding, with structural information of the binding partners being of invaluable help. As such, the event of (reversible) binding is a consequence of association and dissociation processes and involves enthalpic and entropic components. If and how these separate processes and components can be modulated by modification of a ligand’s structure for the sake of overall optimized binding properties is a “hot topic” at present in drug research. Furthermore, the deductive point of view allows one to

apply a divide-and-conquer strategy when it comes to understanding and predicting binding from a theoretical perspective. As a first approximation, binding can be attributed to direct interactions between the binding partners mediated by an aqueous environment, to which contributions due to changes in the conformation and configuration of the binding partners add. While this approximation is valuable in that it allows optimizing the description of each of these terms separately, it also provokes the question to what extent does it hold or, phrased differently, when do cooperativity or compensation effects prevail over additivity?

The majority of topics for this book were selected following this deductive point of view, with an emphasis on rigorous approaches because I believe that these will be more successful in the long term than ad hoc ones. The selection also focused on topics that, at that time, had most quickened interests, had seen considerable progress, or had still been major stumbling blocks in the description and prediction of binding. Unsurprisingly, while much has been achieved in all of the covered areas as undoubtedly laid out in each of the chapters, not in all cases have methods or approaches lived up to one's expectations so far. I am grateful to the authors for pointing this out clearly – such insights will drive further developments that aim at improving our understanding of protein-ligand interactions. Finally, I also tried to balance topics related to biophysical experiments against theoretical and computational approaches, because I have learnt from my own work how well both sides can complement and enhance each other, and what joy this gives.

Last but not least, I express my gratitude to all contributors for providing insightful accounts on the topic of protein-ligand interactions, to the series editors Raimund Mannhold, Hugo Kubinyi, and Gerd Folkers for giving me the opportunity to address this topic, to my current working group for providing fruitful comments about the chapters, and to Frank Weinreich and Heike Nöthe from the publisher Wiley-VCH for their continuous support, great help, and even greater patience.

Düsseldorf, Germany

*Holger Gohlke*

## References

- 1 Venkatesan, K., Rual, J.F., Vazquez, A., Stelzl, U., Lemmens, I., Hirozane-Kishikawa, T., Hao, T., Zenkner, M., Xin, X.F., Goh, K.I., Yildirim, M.A., Simonis, N., Heinzmann, K., Gebreab, F., Sahalie, J.M., Cevik, S., Simon, C., de Smet, A.S., Dann, E., Smolyar, A., Vinayagam, A., Yu, H.Y., Szeto, D., Borick, H., Dricot, A., Klitgord, N., Murray, R.R., Lin, C., Lalowski, M., Timm, J., Rau, K., Boone, C., Braun, P., Cusick, M.E., Roth, F.P., Hill, D.E., Tavernier, J., Wanker, E.E., Barabasi, A.L. and Vidal, M., An empirical framework for binary interactome mapping. *Nature Methods* 2009. 6:83–90.
- 2 Michal, G., *Biochemical Pathways: Biochemie-Atlas*. Spektrum Akademischer Verlag, Heidelberg: 1999.
- 3 Borman, S., New QSAR techniques eyed for environmental assessments. *Chem Eng News* 1990. 68:20–23.

## Part I

### Binding Thermodynamics





# 1

## Statistical Thermodynamics of Binding and Molecular Recognition Models

Kim A. Sharp

### 1.1

#### Introductory Remarks

Equilibrium binding or association of two molecules to form a bimolecular complex,  $A + B \rightleftharpoons AB$ , is a thermodynamic event. This chapter will cover some of the fundamental thermodynamics and statistical mechanics aspects of this event. The aim is to introduce general principles and broad theoretical approaches to the calculation of binding constants, while later chapters will provide examples. Only the noncovalent, bimolecular association under ambient pressure conditions will be considered. However, extension to higher order association involves no additional principles, and extension to high pressure by inclusion of the appropriate pressure–volume work term is straightforward. In terms of the binding reaction above, the association and dissociation constants are defined as  $K = [AB]/[A][B]$  and  $K_D = [A][B]/[AB]$  respectively, where  $[\ ]$  indicates concentration. Either  $K$  or  $K_D$  is the primary experimental observable measured in binding reactions.  $K_D$  is sometimes obtained indirectly by inhibition of binding of a different ligand as a  $K_i$ . From a thermodynamic perspective, the information content from  $K$ ,  $K_D$ , and  $K_i$  is the same.

### 1.2

#### The Binding Constant and Free Energy

To connect the experimental observable  $K$  to thermodynamics, one often finds in the literature the relationship

$$\Delta G_{\text{bind}} = -kT \ln K, \quad (1.1)$$

where  $k$  is the Boltzmann constant,  $T$  is the absolute temperature, and  $\Delta G_{\text{bind}}$  is the “absolute” or “standard” binding free energy. Several comments are given to avoid misuse of this expression. First, one cannot properly take the logarithm of a quantity with units such as  $K$ , so Eq. (1.1) is implicitly

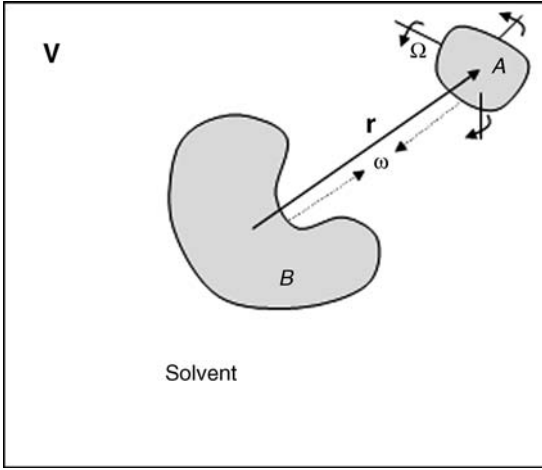
$$\Delta G_{\text{bind}} = -kT \ln \frac{K}{V_{\text{ref}}}, \quad (1.2)$$

where  $V_{\text{ref}}$  is the reference volume in units consistent with the units of concentration in  $K$ , that is, 1 l/mol or about  $1660 \text{ \AA}^3$ /molecule for molarity units. The choice of  $V_{\text{ref}}$  is often referred to as the “standard state” problem. Equivalently, one says that  $\Delta G_{\text{bind}}$  is the free energy change when reactants A and B and the product AB are all at the reference concentration. Second, although the units of concentration used in  $K$  are almost always moles/liter, this is entirely a *convention*, so the actual numerical value for  $\Delta G_{\text{bind}}$  obtained from Eq. (1.2) is arbitrary. Put another way, any method for calculating the free energy of binding must explicitly account for a particular choice of  $V_{\text{ref}}$  before it can meaningfully be compared with experimental values of  $\Delta G_{\text{bind}}$  obtained using Eq. (1.2). Furthermore, ligand efficiency-type measures, such as  $\Delta G_{\text{bind}}/n$  where  $n$  is the number of heavy atoms in a ligand or the molecular weight of a ligand [1], can change radically with (arbitrary) choice of concentration units. Of course, differences in  $\Delta G_{\text{bind}}$  can be sensibly compared provided the same reference state concentration is used. Finally, in Eq. (1.2), the free energy actually depends on the ratio of activities of reactants and products, not on concentrations. For neutral ligands and molecules of low charge density at less than micromolar concentrations, the activity and concentration are nearly equal and little error is introduced. However, this is not true for high charge density molecules such as nucleic acids and many of the ligands and proteins that bind to nucleic acids. Here, the activity coefficient can be substantially different from unity even at infinitely low concentration. Indeed, much of the salt dependence of ligand–DNA binding can be treated as an activity coefficient effect [2–4]. The issue of standard state concentrations, the formal relationship between the binding constant and the free energy, and the effect of activity coefficients are all treatable by a consistent statistical mechanical treatment of binding, as described in Section 1.3.

### 1.3

#### A Statistical Mechanical Treatment of Binding

Derivation of a general expression for the binding constant follows closely the approach of Luo and Sharp [5], although somewhat different treatments using chemical potentials, which provide the same final result, are given elsewhere [6–8]. It is a statistical mechanical principle that any equilibrium observable can be obtained as an ensemble, or Boltzmann weighted average, of the appropriate quantity. Here, the binding constant  $K = [AB]/[A][B]$  is the required observable. Consider a single molecule each of A and B in some volume  $V$  (Figure 1.1) and for convenience define a coordinate system centered on B (the target) in a fixed orientation. Over time, the ligand (A) will explore different positions and orientations (poses) relative to B, where  $\mathbf{r}$  and  $\Omega$  represent the three position and three orientation coordinates of A with respect to B. Now A and B interact with each other with an energy that depends not only on their relative position ( $\mathbf{r}$ ,  $\Omega$ ) but in general also on the conformations of A, B, and the surrounding solvent. If  $n_a$ ,  $n_b$ ,  $n_s$  are the number of atoms in A, B, and solvent, then the energy is a function of  $3n_a + 3n_b + 3n_s - 6$  coordinates. In principle, one could keep all these degrees of freedom explicit. From a



**Figure 1.1** Schematic illustration of two molecules A and B interacting through solvent with a potential of mean force  $\omega$  as a function of their relative position  $\mathbf{r}$  and orientation  $\Omega$ .

practical standpoint, this would be a complicated and expensive function to evaluate. However, one may integrate over the solvent coordinates and the  $(3n_a - 6) + (3n_b - 6)$  internal coordinates so that the interaction between A and B for a given  $(\mathbf{r}, \Omega)$  is described by an interaction potential of mean force (pmf)  $\omega(\mathbf{r}, \Omega)$ . If one defines the pmf between A and B at infinite separation in their equilibrium conformations to be 0, then  $\omega(\mathbf{r}, \Omega)$  is the thermodynamic work of bringing A and B from far apart to some mutual pose  $(\mathbf{r}, \Omega)$ , accounting for both solvent effects and internal degrees of freedom of A and B. A will sample each pose  $(\mathbf{r}, \Omega)$  with a probability given by the Boltzmann factor of the pmf:

$$p(\mathbf{r}, \Omega) \propto e^{-\beta\omega(\mathbf{r}, \Omega)}, \quad (1.3)$$

where  $\beta = 1/kT$ . Indeed, one may consider the pmf to be defined by this equation. The binding constant will then be given by the fraction of time A is in the bound state,  $f_{ab}$ , relative to that in the free state,  $f_f$ :

$$K = \frac{[AB]}{[A][B]} = \frac{f_{ab}/V}{(f_f/V)(f_f/V)} \xrightarrow{V \rightarrow \infty} f_{ab}V, \quad (1.4)$$

where in the dilute limit  $f_f \rightarrow 1$ . It is convenient to introduce a function  $H(\mathbf{r}, \Omega)$  that takes a value of 1 for poses where A is bound and a value of 0 when it is free. Then, the fraction of the time A is bound is given by the ensemble average of  $H$ :

$$f_{ab} = \int d\mathbf{r} d\Omega H(\mathbf{r}, \Omega) e^{-\beta\omega(\mathbf{r}, \Omega)} / \int d\mathbf{r} d\Omega e^{-\beta\omega(\mathbf{r}, \Omega)}. \quad (1.5)$$

The integrals are taken over all orientations and over the entire volume of the solution, so the denominator gives  $8\pi^2V$ . Substituting into Eq. (1.4), the final expression for the association constant is

$$K = \frac{1}{8\pi^2} \int d\mathbf{r} d\Omega H(\mathbf{r}, \Omega) e^{-\beta\omega(\mathbf{r}, \Omega)}. \quad (1.6)$$

One may then convert this to an “absolute” binding free energy using Eq. (1.2):

$$\Delta G_{\text{bind}} = kT \ln(8\pi^2 V_{\text{ref}}) - kT \ln \int d\mathbf{r} d\Omega H(\mathbf{r}, \Omega) e^{-\beta\omega(\mathbf{r}, \Omega)}. \quad (1.7)$$

- Equation 1.6 is a general and exact expression for the association constant. The integral depends explicitly on just six variables describing the pose of A with respect to B. The other degrees of freedom are included implicitly, but exactly through the thermodynamic quantity  $\omega(\mathbf{r}, \Omega)$ , the potential of mean force.
- The different treatment of coordinates for translation/orientation versus the others is a formal one: Any subset of coordinates may in principle be kept explicit, with the appropriate pmf being used for the rest. For example, one may keep the internal coordinates of A and B explicit, making the solvent coordinates implicit. The choice here is designed to highlight the translation/rotation contribution to binding that has been widely discussed, with some disagreement, in the literature [5, 6, 9–13]. It also reflects the practical fact that in many docking and screening applications, a particular pose is generated explicitly, that is,  $(\mathbf{r}, \Omega)$  is specified, and then the pose is “scored” in some way. The pmf also provides a natural way to introduce approximations necessary for any practical calculation of  $K$  in biological systems, for example, in the treatment of solvent.
- The integral has the correct units of volume, with the length scale for the translation coordinates being determined by the units of concentration used in  $K$ . The first term in Eq. (1.7) is the contribution of the rotation/translation (R/T) entropy in the unbound state, which depends on the reference concentration. The integral term in Eq. (1.6) is the Boltzmann phase volume of the bound state.
- Through  $H(\mathbf{r}, \Omega)$ , there is explicit consideration of what constitutes the bound complex, in terms of the relative position and orientation of A with respect to B. For example, if B has more than one binding site for A, this would be taken into account in the specification of where  $H = 1$ .
- Either Cartesian coordinates or the bond length, bond angle, and dihedral angle coordinates may be used. The trend now is toward the latter, as they lend themselves more naturally to the analysis of different internal motions of the molecules and their contribution to binding.

The meaning of Eq. (1.6) is illustrated by two simple examples.

### 1.3.1

#### Binding in a Square Well Potential

Let the pmf be approximated by a simple, three-dimensional square well potential of depth  $\varepsilon$  and width  $b$  in each of the  $x$ ,  $y$ ,  $z$  directions and the bound complex be the region in the well only. From Eq. (1.6), the association constant is

$$K = b^3 e^{\beta\varepsilon} \quad (1.8)$$

and Eq. (1.2) yields

$$\Delta G_{\text{bind}} = -\varepsilon + kT \ln(V_{\text{ref}}/b^3). \quad (1.9)$$

The first term, the well depth, makes a direct, linear contribution to the binding free energy. The second term is positive and comes from the restriction of the ligand to the square well. It is the translation entropy penalty for binding, and it depends on the ratio of the volumes available to the ligand in the free state at say 1 M (the entire volume  $V_{\text{ref}}$ ) versus that in the bound state. In this simple example, there is no rotational entropy penalty because in the bound state the ligand can rotate freely in  $8\pi^2$  of orientation phase volume, just as in the free state. However, restriction in rotation in the bound state will add another positive term to  $\Delta G_{\text{bind}}$ , the rotation entropy penalty, with a similar form:  $kT \ln(8\pi^2/V_{\Omega})$ , where  $V_{\Omega} < 8\pi^2$  is the orientation phase volume in the bound state. We can see even from this simple example that for any meaningful degree of binding, the translational and rotational phase volumes available to a ligand in the bound state must be less than  $V_{\text{ref}}$  and  $8\pi^2$ , respectively, so there is *always* a R/T entropy penalty to be overcome for binding to occur. The question is how much is it in specific cases. A related point is that even though the depth of the well may be known, for example, from some calculation (in the parlance of the field, from a single point energy determination), this cannot be directly compared with  $\Delta G_{\text{bind}}$  because the second term is not included. The numerical value of the binding free energy depends on the reference concentration, which is nowhere in the single point calculation. One way or another, the residual R/T entropy of A in the bound state must be accounted for.

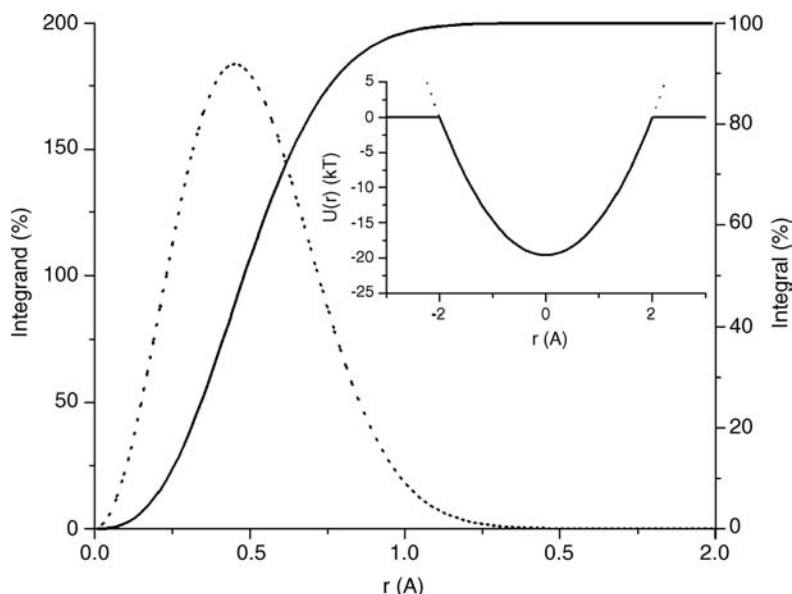
### 1.3.2

#### Binding in a Harmonic Potential

If one is starting from a known complex structure derived from, for example, X-ray, NMR, or molecular mechanics minimization, one is presumably close to the minimum energy (pmf) configuration. The pmf in this region may be close to harmonic or at least expandable in a Taylor expansion, which to second order is harmonic. It is, therefore, instructive to consider binding in a harmonic potential, although this is a simplified model of the real situation. Let the potential well be a three-dimensional harmonic potential of the form

$$\omega(r) = \varepsilon \left( \left( \frac{r}{b} \right)^2 - 1 \right) (r < b), \quad \omega(r) = 0 (r \geq b), \quad (1.10)$$

where  $\varepsilon$  is the depth of the well at the minimum,  $r$  is the radial distance from the minimum, and  $b$  defines the width so that for  $r \geq b$ ,  $\omega = 0$  (Figure 1.2). Again, the bound complex is defined to be the region in the well only. Substituting Eq. (1.10) into Eq. (1.6) and integrating, the association constant for this truncated harmonic potential is



**Figure 1.2** Contributions to the binding-phase integral. *Dotted line*: Value of the integrand of Eq. (1.6) at  $r$ . *Solid line*: Value of the resulting integral from 0 to  $r$ . Both are expressed as a percentage of the total association constant.

Contributions were calculated for a truncated three-dimensional harmonic well potential, half-width  $2 \text{ \AA}$ , and depth  $19.6kT$  (inset) that has  $K_d = 10 \mu\text{M}$ .

$$K = b^3 e^{\beta\varepsilon} \left( 4\pi \sum_{n=0}^{\infty} \frac{-1^n (\beta\varepsilon)^n}{(2n+3)n!} \right) \approx b^3 e^{\beta\varepsilon} \left( \frac{\pi}{\beta\varepsilon} \right)^{3/2}. \quad (1.11)$$

The approximate equality comes from using an untruncated harmonic potential (i.e., the potential goes to infinity as the complex is dissociated), which for this case gives a binding free energy of

$$\Delta G_{\text{bind}} = -\varepsilon + kT \ln(V_{\text{ref}}/b^3) - 3/2 kT \ln(\pi/\beta\varepsilon). \quad (1.12)$$

Comparing the square well and harmonic potential models, one sees that the “depth” and “volume” factors,  $e^{\beta\varepsilon}$  and  $b^3$ , contribute in the same way to the binding constant, the difference being a “well shape” factor. We see from the form of the expression for the association constant that the lower the pmf, the more the contribution to the integral by that region, so most of the contribution to binding should come from the near minimum energy configuration. This is illustrated in Figure 1.2, using a well half-width of  $2 \text{ \AA}$  and a depth of  $19.6kT$ . These parameters are chosen to give a moderate affinity of  $10 \mu\text{M}$  – typical of the compounds studied by virtual screening and docking calculations in early lead identification – with a reasonable degree of motion in the binding pocket. It can be seen that almost all the contribution to the binding constant comes significantly before reaching the well boundary. Thus, the problem of giving the exact definition of the complex in