

## Contents

### Preface xv

<b>1</b>	<b>Protein Structure and Conformational Dynamics 1</b>
	<i>Volkhard Helms</i>
1.1	Structural and Hierarchical Aspects 1
1.1.1	Size of Proteins 1
1.1.2	Protein Domains 1
1.1.3	Protein Composition 2
1.1.4	Secondary Structure Elements 3
1.1.5	Active Sites 3
1.1.6	Membrane Proteins 6
1.1.7	Folding of Proteins 7
1.2	Conformational Dynamics 8
1.2.1	Large-Scale Domain Motions 8
1.2.2	Dynamics of N-Terminal and C-Terminal Tails 9
1.2.3	Surface Dynamics 9
1.2.4	Disordered Proteins 9
1.3	From Structure to Function 10
1.3.1	Evolutionary Conservation 10
1.3.2	Binding Interfaces 10
1.3.3	Surface Loops 11
1.3.4	Posttranslational Modifications 11
1.4	Summary 11
	References 12
<b>2</b>	<b>Protein-Protein-Binding Interfaces 15</b>
	<i>Zeynep Abali, Damla Ovez, Simege Senyuz, Ozlem Keskin, and Attila Gursoy</i>
2.1	Definition and Properties of Protein-Protein Interfaces 15
2.2	Growing Number of Known Protein-Protein Interface Structures 18
2.3	Surface Areas of Protein-Protein Interfaces 21
2.4	Gap Volume of Protein-Protein Interfaces 22
2.5	Amino Acid Composition of Interfaces 22
2.6	Secondary Structure of Interfaces 23

2.7	Protein–Protein-Binding Energy	24
2.8	Interfaces of Homo- and Hetero-Dimeric Complexes	24
2.9	Interfaces of Non-obligate and Obligate Complexes	25
2.10	Interfaces of Transient and Permanent Complexes	25
2.11	Biological vs. Crystal Interfaces	26
2.12	Type I, Type II, and Type III Interfaces	27
2.13	Conserved Residues and Hot Spots in Interfaces	28
2.14	Conclusion and Future Implications	29
	References	30
<b>3</b>	<b>Correlated Coevolving Mutations at Protein–Protein Interfaces</b>	<b>39</b>
	<i>Alexander Schug</i>	
3.1	Introduction	39
3.2	A Short Introduction into Biomolecular Modeling	41
3.3	Statistical Inference of Coevolution	41
3.3.1	Limitations of Local Statistical Inference	41
3.3.2	Direct-Coupling Analysis – A Potts Model Based on Multiple Sequence Alignments	42
3.4	Solving the Inverse Potts Model	43
3.5	Contact Guided Protein and RNA Structure Prediction	45
3.6	Inter-Monomer Interaction and Signaling	45
3.7	Summary	46
	References	47
<b>4</b>	<b>Computational Protein–Protein Docking</b>	<b>53</b>
	<i>Martin Zacharias</i>	
4.1	Introduction	53
4.2	Rigid Body Protein–Protein Docking Approaches	56
4.3	Accounting for Conformational Changes during Docking	59
4.4	Integration of Bioinformatics and Experimental Data for Protein–Protein Docking	61
4.5	Template-Based Protein–Protein Docking	62
4.6	Flexible Refinement of Docked Complexes	64
4.7	Scoring of Docked Complexes	66
4.8	Conclusions and Future Developments	67
	Acknowledgments	68
	References	68
<b>5</b>	<b>Identification of Putative Protein Complexes in Protein–Protein Interaction Networks</b>	<b>77</b>
	<i>Sudharshini Thangamurugan, Markus Hollander, and Volkhard Helms</i>	
5.1	Protein–Protein Interaction Networks	77
5.2	Integration of Various PPI Resources in Public Data Repositories	79
5.3	Protein–Protein Interaction Networks of Model Organisms	80

5.3.1	PPIN of <i>Saccharomyces cerevisiae</i>	80
5.3.2	PPIN of Human	83
5.4	Algorithms to Identify Protein Complexes in PPI Networks	84
5.4.1	Molecular Complex Detection (MCODE)	84
5.4.1.1	Definitions	85
5.4.1.2	Algorithm	86
5.4.1.3	Examples	88
5.4.2	Clustering with Overlapping Neighborhood Expansion (ClusterONE)	89
5.4.2.1	Definitions	89
5.4.2.2	Algorithm	90
5.4.3	Domain-Aware Cohesiveness Optimization (DACO)	92
5.5	Summary	94
	References	95

## **6 Structure, Composition, and Modeling of Protein Complexes** 101

*Olga V. Kalinina*

6.1	Protein Complex Structure	101
6.1.1	Protein Quaternary Structure	101
6.1.2	Classification of Protein–Protein Interaction Interfaces	102
6.1.3	Classification and Evolution of Protein Complexes	105
6.2	Methods for Automated Assignment of Biological Assemblies	106
6.2.1	Assignment from Crystallographic Data	107
6.2.2	Employing Machine-Learning Methods	108
6.2.3	Leveraging Evolutionary Information	109
6.3	Computational Approaches to Predicting 3D Structure of Protein Complexes	110
6.3.1	Combinatorial Docking	110
6.3.2	Homology-Based Complex Reconstruction	114
6.3.3	Prediction from Sequence	115
6.3.4	Assisted Docking	116
6.4	Conclusion and Outlook	117
	Acknowledgments	118
	References	118

## **7 Live-Cell Structural Biology to Solve Molecular Mechanisms: Structural Dynamics in the Exocyst Function** 127

*Altair C. Hernandez, Baldo Oliva, Damien P. Devos, and Oriol Gallego*

7.1	Introduction	127
7.2	Structural Biology Using Light Microscopy Methods	129
7.3	Hybrid Methods: Integrative Structural Biology	131
7.4	Integrative Modeling: The Case of the Exocyst Complex	133
7.5	Comparing the <i>In Situ</i> Architecture of the Exocyst with a High-Resolution Cryo-EM Model	136

7.6	Discussion and Future Perspectives	138
	Acknowledgements	139
	References	140
<b>8</b>	<b>Kinetics and Thermodynamics of Protein–Protein Encounter</b>	<b>143</b>
	<i>Nicolas Künzel and Volkhard Helms</i>	
8.1	Introduction	143
8.2	Thermodynamic Ensembles and Free Energy	143
8.2.1	The Isothermal–Isobaric Ensemble and the Gibbs Free Energy	144
8.3	Overview of Computational Methods to Determine Binding Free Energies	146
8.3.1	Coarse Graining	147
8.3.1.1	Brownian Dynamics	147
8.3.2	Endpoint Methods	149
8.3.2.1	MM/PBSA/MM/GBSA	149
8.3.3	Potential of Mean Force/Pathway Methods	150
8.3.3.1	Thermodynamic Integration	151
8.3.3.2	Umbrella Sampling (US)	151
8.3.3.3	Steered MD (SMD)	153
8.3.3.4	Metadynamics	153
8.3.3.5	Adaptive Biasing Force (ABF)	155
8.3.4	Replica-Exchange Methods	155
8.3.4.1	Parallel Tempering	155
8.3.4.2	Generalized/Hamiltonian Replica-Exchange Methods	156
8.3.5	Additional Pathway Methods	156
8.3.6	Relative Binding Free Energies	156
	References	157
<b>9</b>	<b>Markov State Models of Protein–Protein Encounters</b>	<b>163</b>
	<i>Simon Olsson</i>	
9.1	Notation	163
9.1	Introduction	163
9.2	Molecular Dynamics and Markov State Models	164
9.2.1	Markov State Models: Theory and Properties	165
9.3	Strategies for MSM Estimation, Validation, and Analysis	169
9.3.1	Variational Approach for Conformational Dynamics and Markov Processes (VAC and VAMP)	169
9.3.2	Feature Selection	170
9.3.3	Dimensionality Reduction	171
9.3.4	Clustering	172
9.3.5	Model Estimation and Validation	173
9.3.6	Spectral Gaps and Coarse-Graining	174
9.3.7	Adaptive and Enhanced Sampling Strategies	175
9.3.8	Practical Consideration for Studying Protein–Protein Encounters	176

9.3.9	Analysis of the Association–Dissociation Path Ensemble	177
9.4	The Connection to Experiments	178
9.4.1	Experimental Observability, Forward Models, and Errors	178
9.4.1.1	Sources of Errors and Uncertainty	179
9.4.2	Predicting Experimental Observables Using MSMs	180
9.4.3	Integrating Experimental and Simulation Data into Augmented Markov Models	181
9.5	Protein–Protein and Protein–Peptide Encounters	182
9.6	Emerging Technologies	184
	Acknowledgments	186
	References	186
<b>10</b>	<b>Transcription Factor – DNA Complexes</b>	<b>195</b>
	<i>Volkhard Helms</i>	
10.1	Introduction	195
10.2	Principles of Sequence Recognition	197
10.3	Dimerization of Eukaryotic TFs	198
10.4	Detection of Epigenetic Modifications	199
10.5	Detection of DNA Curvature/Bending	200
10.6	Modifications of Transcription Factors	200
10.7	Transcription Factor Binding Sites	201
10.8	Experimental Detection of TFBS	201
10.8.1	Protein-Binding Microarrays	202
10.8.2	Chromatin Immunoprecipitation Assays	203
10.8.3	DamID Profiling of Protein–DNA Interactions	204
10.9	Position-Specific Scoring Matrices	204
10.10	Molecular Modeling of TF–DNA Complexes	204
10.11	Cis-Regulatory Modules	205
10.12	Relating Gene Expression to Binding of Transcription Factors	207
10.13	Summary	208
	References	208
<b>11</b>	<b>The Chromatin Interaction System</b>	<b>213</b>
	<i>Sarah Kreuz, Stefan-Sebastian David, Lorena Viridiana Cortes Medina, and Wolfgang Fischle</i>	
11.1	Chromatin Is a Special Interaction Platform	213
11.2	Interaction of Proteins with Histone Posttranslational Modifications	215
11.2.1	The History of Histone Posttranslational Modifications and the Histone Code	215
11.2.2	Peptides and Nucleosomal Templates for Studying Histone PTMs	222
11.2.3	Qualitative Analysis of Histone PTM Readout	224
11.2.3.1	Characterizing Binding Specificities of Known Readers	224
11.2.3.2	Identification of New Reader Proteins	225
11.2.4	Molecular Parameters of Histone PTM–Reader Interaction	226

11.2.5	Cellular Assays to Characterize Histone PTM–Reader Interactions	227
11.2.5.1	Visualizing Histone–Reader Interactions	227
11.2.5.2	Chromatin Immunoprecipitation	229
11.2.5.3	Cellular Labeling and Affinity Enrichment	231
11.3	Interaction of Proteins with Modified Nucleic Acids	231
11.3.1	Discovery of DNA Methylation and the First Reader Proteins	231
11.3.2	RNA Modifications	234
11.3.3	Modified DNA and RNA Templates	234
11.3.4	<i>In Vitro</i> Assays for Identifying Readers of Nucleic Acid Methylation	235
11.3.4.1	Affinity Purification to Identify Novel Modification Readers	235
11.3.4.2	Characterizing Binding Specificities of Known Readers	235
11.3.5	Cellular Assays for Identifying Readers of Nucleic Acid Modifications	236
11.4	UHRF1 as an Example of a Multidomain Reader/Writer Protein of Histone and DNA Modifications	239
11.5	Histone Chaperones and Chromatin Remodeling Complexes	241
11.5.1	Chromatin Assembly and Remodeling	241
11.5.2	Discovery of Histone Chaperones and Chromatin Remodelers	242
11.5.3	Methods for Identifying Histone Chaperones and Remodeling Factors	244
11.5.3.1	Immunoprecipitation Assays	244
11.5.3.2	Computational Methods	244
11.5.4	Assays to Study Chaperone and Remodeler Activities	245
11.5.5	Cellular Assays	245
11.6	Challenges in Chromatin Interactomics	247
	References	248
<b>12</b>	<b>RNA–Protein Interactomics</b>	<b>271</b>
	<i>Cornelia Kilchert</i>	
12.1	Introduction	271
12.2	Interactions of Proteins with mRNA and ncRNA	272
12.3	The Basic Toolbox	273
12.3.1	Metabolic RNA Labeling with Modified Nucleobases	273
12.3.2	RNA–Protein Crosslinking	274
12.4	RNA–Protein Interactomics	276
12.4.1	What Proteins Are Bound to my RNA (or RNA in General)?	276
12.4.1.1	Cataloging the RBPome	276
12.4.1.2	Interactomes of Specific RNAs	278
12.4.2	Which RNA Species Are Bound by my RBP?	280
12.4.2.1	Copurification Methods: CLIP and Derivatives	280
12.4.2.2	Proximity-Dependent Labeling Methods	280
12.5	Outlook	282
	Notes	283
	References	283

<b>13</b>	<b>Interaction Between Proteins and Biological Membranes</b>	<b>293</b>
	<i>Lorant Janosi and Alemayehu A. Gorfe</i>	
13.1	Introduction	293
13.2	The Plasma Membrane: Overview of Its Structure, Composition, and Function	294
13.3	Lipid-Based and Protein-Based Sorting of Plasma Membrane Components	295
13.3.1	Lipid-Based Sorting and Domain Formation	295
13.3.2	Protein-Based Sorting and Membrane Curvature	296
13.3.3	Proteolipid Sorting and Membrane Domain Stabilization	297
13.4	Interaction of Peripheral Membrane Proteins with Membrane Lipids	297
13.4.1	Protein-Based Membrane-Targeting Motifs	298
13.4.2	Lipid-Based Membrane-Targeting Motifs	301
13.5	Interactions and Conformations of Transmembrane Proteins in Lipid Membranes	303
13.5.1	Glycophorin A and EGFR as Examples of Single-Pass Transmembrane Proteins	303
13.5.2	GPCR as an Example of Multi-Pass TM Helical Proteins	306
13.5.3	Aquaporin as an Example of Oligomeric Multi-Pass TM Proteins	306
13.5.4	Antimicrobial Peptides: Peripheral or Integral?	307
13.6	Summary	308
	Acknowledgment	308
	References	309
<b>14</b>	<b>Interactions of Proteins with Small Molecules, Allosteric Effects</b>	<b>315</b>
	<i>Michael C. Hutter</i>	
	Abbreviations	315
14.1	Introduction	315
14.2	Modes of Binding to Proteins	316
14.3	Types of Interaction Between Protein and Ligand	317
14.3.1	Salt Bridges	317
14.3.2	Coordination of Ions via Lone Pairs	318
14.3.3	Hydrogen Bonds	319
14.3.3.1	Definition	319
14.3.3.2	Occurrence and Functionality of Hydrogen Bonds in Biological Systems	320
14.3.3.3	Classification of Hydrogen Bonds	321
14.3.3.4	Weak Hydrogen Bonds	321
14.3.3.5	Hydrogen Bonds to Fluorine	322
14.3.3.6	Nitrogen vs. Oxygen as Competing Hydrogen Bond Acceptors	322
14.3.3.7	Bifurcated Hydrogen Bonds	322
14.3.4	Halogen Bonds	323

14.3.5	van der Waals Interactions	324
14.3.6	Mutual Interactions of Delocalized $\pi$ -Electron Systems	325
14.3.7	Cation- $\pi$ Interaction	325
14.3.8	Anion- $\pi$ Interaction	325
14.3.9	Unusual Protein-Ligand Contacts	326
14.4	Modeling Intermolecular Interactions by Force Fields and Docking Simulations	326
14.5	Entropic Aspects	327
14.6	Allosteric Effects: Conformational Changes Upon Ligand Binding	327
14.7	Aspects of Ligand Design Beyond Protein-Ligand Interactions	329
14.8	Conclusions	330
	References	330
<b>15</b>	<b>Effects of Mutations in Proteins on Their Interactions</b>	<b>333</b>
	<i>Alexander Gress and Olga V. Kalinina</i>	
15.1	Introduction	333
15.2	Structural Annotation of Mutations in Proteins	334
15.2.1	Databases for Structural Annotation of Mutations	335
15.2.2	Dynamic Structural Annotation Pipelines	340
15.3	Methods for Predicting Effect of Protein Mutations	342
15.3.1	Prediction of Phenotypic Effect	343
15.3.2	Estimation of Mutation Effects by Modeling Biophysical Properties of Proteins	344
15.3.3	Prediction of Mechanistic Effects of Mutations on Interactions of Proteins	345
15.4	Conclusion	348
	Acknowledgments	349
	References	349
<b>16</b>	<b>Not Quite the Same: How Alternative Splicing Affects Protein Interactions</b>	<b>359</b>
	<i>Zakaria Louadi, Olga Tsoy, Jan Baumbach, Tim Kacprowski, and Markus List</i>	
	List of Abbreviations	359
16.1	Introduction	359
16.2	Effects of Alternative Splicing on Individual Proteins	362
16.2.1	Alternative Splicing and Protein Structure	362
16.2.2	Alternative Splicing and Intrinsically Disordered Regions	362
16.3	Effects of Alternative Splicing on Protein-Protein Interaction Networks	367
16.3.1	Alternative Splicing Rewires Protein-Protein Interactions	367
16.3.2	Alternative Splicing in Diseases	368
16.3.3	Resources for Studying the Effect of Alternative Splicing on Protein-Protein Interactions	369
16.4	Conclusion and Future Work	373
	References	374

<b>17</b>	<b>Phosphorylation-Based Molecular Switches</b>	<b>381</b>
	<i>Attila Reményi</i>	
17.1	Introduction	381
17.1.1	Structural and Functional Effects of Protein Phosphorylation	383
17.2	Reversible Protein Phosphorylation in Cellular Signaling: Writers, Readers, and Erasers	386
17.3	Protein Kinases as Molecular Switches and as Components of Signaling Cascades	388
17.4	Mechanisms of Phosphorylation Specificity: the Importance of Short Linear Motifs	390
17.5	Examples of Phospho-Switch-Based Biological Regulation	392
17.6	Conclusion	395
	Acknowledgments	397
	References	397
<b>18</b>	<b>Summary and Outlook</b>	<b>401</b>
	<i>Volkhard Helms and Olga V. Kalinina</i>	
18.1	Technical State of the Art	401
18.2	Role of Machine Learning	401
18.3	Challenges	402
18.4	What Picture(s) May Evolve?	403
	References	404
	<b>Index</b>	<b>405</b>