

Contents

<i>Foreword, Professor Sir Alan Fersht</i>	xv
<i>Preface</i>	xvii
<i>About the Author</i>	xix
<i>Acknowledgments</i>	xxi
<i>Abbreviations and Acronyms</i>	xxiii
1 Principles of Protein Structure and Function	1
1.1 Physical Forces That Shape Protein Structure	1
1.2 Primary Structure: Amino Acid Sequence	3
1.3 Protein-Coding Genes	4
1.4 Post-Translational Modifications of Amino Acids	5
1.5 Hierarchical Description of Structure	6
1.5.1 Secondary Structure	6
1.5.2 Tertiary Structure	9
1.5.3 Quaternary Structure	11
1.6 Folding of a Protein	12
1.6.1 Thermodynamic Aspects of Protein Folding	12
1.6.2 Kinetic Aspects of Protein Folding	13
1.6.3 Mechanism of Protein Folding	14
1.6.4 Folding and Chaperones	14
1.7 Unfolding of a Protein: Lessons from Polymer Theory	15
1.8 The Limits of Global Descriptions of the Unfolded State	17
1.9 Databases of Proteins and Protein Structures	17
1.10 DisProt: The Database of Disordered Proteins	18
1.11 The Classical Structure-Function Paradigm	18
2 A Brief History of Protein Disorder	21
2.1 Can We Define Disorder?	21
2.2 The History of Disorder	22
2.2.1 The Legacy of the Lock-and-Key Hypothesis	22
2.2.2 Structural Adaptability of Binding Sites	23
2.2.3 Polymer Theory and Protein Folding	23
2.2.4 Caseins Are Different	24
2.2.5 If a Protein Does Not Crystallize	24
2.2.6 The Advent of NMR	25
2.3 So We Have Disordered Proteins	27

3 Indirect Techniques for Recognizing and Characterizing Protein Disorder	31
3.1 Resistance to Heat	31
3.2 Resistance to Chemical Denaturation	33
3.3 Unusual SDS-PAGE Mobility	33
3.4 Enhanced Proteolytic Sensitivity	34
3.5 Limited Proteolysis and Local Structure	35
3.6 Differential Scanning Calorimetry	35
3.6.1 Transition to a More Ordered State	37
3.6.2 Residual Structure in Calpastatin	37
3.7 Isothermal Titration Calorimetry	37
3.7.1 The Energetics of Binding of a PPII Helix to Its Cognate SH3 Domain	38
3.7.2 Binding of the KID Domain of p27 ^{Kip1} to Cyclin A-Cdk2	38
3.8 Chemical Cross-Linking	40
3.9 H/D Exchange	41
4 Hydrodynamic Techniques	43
4.1 Gel Filtration (Size-Exclusion) Chromatography	43
4.2 Dynamic Light Scattering	45
4.3 Analytical Ultracentrifugation	46
4.4 Small-Angle X-Ray Scattering	47
4.4.1 Measles Virus Nucleoprotein	50
4.4.2 Bacterial Cellulase	51
4.4.3 p53	52
4.5 Pulsed-Field Gradient NMR	53
5 Spectroscopic Techniques for Characterizing Disorder	55
5.1 X-Ray Crystallography	55
5.2 Fluorescence Spectroscopy	57
5.2.1 UV Fluorescence	58
5.2.2 Fluorescence Quenching	58
5.2.3 ANS Binding	60
5.2.4 Fluorescence Resonance Energy Transfer	60
5.2.5 Fluorescence Correlation Spectroscopy	61
5.2.5.1 Dimensions of an IDP and the Effect of Crowding	61
5.2.5.2 Internal Protein Dynamics	62
5.3 Fourier-Transform Infrared Resonance Spectroscopy	62
5.4 Circular Dichroism	63
5.5 Raman Optical Activity Spectroscopy	65
5.6 Electron Paramagnetic Resonance Spectroscopy	66
5.7 Electron Microscopy	69
5.8 Atomic Force Microscopy	71
5.8.1 Matrix Metalloproteinase 9	72
5.8.2 α -Synuclein	72

6	Nuclear Magnetic Resonance	73
6.1	Basic Principles	73
6.2	Global Characterization by NMR	74
6.2.1	1-D ^1H NMR	74
6.2.2	Wide-Line NMR	75
6.2.3	Pulsed-Field Gradient NMR	76
6.2.4	HSQC	76
6.3	Sequence-Specific Structural Information	78
6.3.1	Chemical Shifts	79
6.3.2	Dynamic Information from Relaxation Data	79
6.3.3	Distance Information from NOE	81
6.3.4	Coupling Constants	82
6.4	Special Applications	83
6.4.1	Combinations with MD	83
6.4.2	Amide Proton Exchange Rate	83
6.4.3	In-Cell NMR	84
7	Proteomic Approaches for the Identification of IDPs	85
7.1	Expectations and Limitations of Proteomic Studies	85
7.2	2DE-MS Identification of Proteins in Extracts Enriched for Disorder	86
7.3	Native/Urea 2DE Provides Direct Information on Disorder	88
8	IDPs under Conditions Approaching <i>In Vivo</i>	91
8.1	Macromolecular Crowding in the Cell	91
8.2	<i>In Vitro</i> Approaches to Mimicking Crowding Conditions	92
8.3	The State of IDPs <i>In Vivo</i>	95
8.3.1	Proteasomal Degradation	95
8.3.2	In-Cell NMR	97
8.4	Physiological Half-Life of IDPs: No Signs of Rapid Degradation	99
8.5	Indirect Considerations Underscoring Disorder of IDPs <i>In Vivo</i>	100
9	Prediction of Disorder	103
9.1	General Points	103
9.2	Propensity-Based Predictors	103
9.2.1	Prediction of Low-Complexity Regions	106
9.2.2	Charge-Hydrophathy Plot	106
9.2.3	Prediction of Globularity and Disorder	107
9.2.4	Composition and Hydrophobic Cluster Analysis	108
9.3	Machine-Learning Algorithms	109
9.3.1	Neural Networks	109
9.3.2	Support-Vector Machines	110
9.4	Prediction Based on Interresidue Contacts	111
9.4.1	Contact Numbers of Amino Acids	111
9.4.2	Estimating Pair-Wise Interresidue Interaction Energies	112
9.4.3	Predictor of Contact Potentials	112

9.5	Prediction of Short and Long Regions of Disorder Separately	112
9.6	Combination of Predictors: Meta-Servers	114
9.7	Prediction of Functional Motifs In IDPs	115
9.8	Comparison of the Accuracy of Predictors: The CASP Experiment	116
9.9	A Better Target Prioritization in Structural Genomics	119
10	Structure of IDPs	121
10.1	Primary Structure of Disordered Proteins	121
10.1.1	Amino Acid Composition	121
10.1.2	Sequence Features Characterizing Disorder	123
10.1.3	Flavors of Disorder?	124
10.2	Secondary Structure of Disordered Proteins	125
10.2.1	Secondary Structure in Solution State: Signs of Transient Order	125
10.2.2	A Lot of PPII Helix Conformation	126
10.2.3	Secondary Structure in Solution State: Sequence-Specific Information	127
10.2.3.1	p27 ^{Kip1}	127
10.2.3.2	CREB KID	130
10.2.3.3	Tau Protein	131
10.2.3.4	Fibronectin-Binding Protein A	131
10.2.3.5	α -Synuclein	132
10.2.3.6	p53	132
10.2.3.7	Calpastatin	132
10.2.4	Secondary Structure in the Bound State	133
10.3	Ambiguity in Structure	134
10.3.1	Chameleon Sequences	134
10.3.2	Dual-Personality Sequences	135
10.3.3	The Twilight Zone between Order and Disorder	135
10.4	Tertiary Structure: Global Features of IDP Structures	136
10.4.1	Hydrodynamic Description	136
10.4.2	Spectroscopic Approaches	137
10.4.3	Global Structure: Is It Related to the Structure in the Bound State?	139
10.5	Dynamics of IDP Structure: The Time-Course of Fluctuations within the Ensemble	140
10.5.1	The Importance of Dynamics in Structural Descriptions	140
10.5.1.1	Local/Segmental Motions	140
10.5.1.2	Restricted Segmental Motions	141
10.5.1.3	Reduced Local Motion Signals Transient Structural Elements	141
10.5.2	A Reduction in Motility Signals Disorder-to-Order Transition	142
10.6	A Readout of Structure: The Hydrate Layer of IDPs	142

11 Biological Processes Enriched in Disorder	143
11.1 Biological Functions Enriched in Disorder	143
11.2 Disorder in Transcription/Transcription Regulation	145
11.2.1 Transcription Factors	145
11.2.2 Transcription Co-Activators	146
11.2.3 Disorder in the Core Apparatus	148
11.3 Disorder in Signaling Proteins	149
11.3.1 Receptors and Membrane Proteins	149
11.3.2 Scaffold Proteins and Hub Proteins	151
11.3.3 Regulation of the Cell Cycle	152
11.4 Nucleic Acid-Containing Organelles	152
11.4.1 Ribosome	152
11.4.2 Disorder in Chromatin Organization	153
11.4.2.1 Histones	153
11.4.2.2 Other Chromatin Organizing Proteins	154
11.5 Disorder in RNA-Binding Proteins: Transcription and RNA Folding	155
11.6 Cytoskeletal Proteins	157
11.6.1 Microfilaments	157
11.6.2 Intermediate Filaments	158
11.6.3 Microtubules	159
11.7 Disorder in Stress Proteins	159
11.8 Disorder and Metal Binding	160
11.9 Disorder and Enzyme Activity	161
11.10 Is There a Link between the Pattern of Disorder and Function?	162
12 Molecular Functions of Disordered Proteins	163
12.1 Entropic Chain Functions	163
12.1.1 Linkers and Spacers	163
12.1.2 Entropic Clocks	165
12.1.3 Entropic Springs	166
12.1.4 Entropic Bristles/Brushes	166
12.2 Display Site Functions	168
12.2.1 Phosphorylation Sites	168
12.2.2 Sites of Proteolytic Processing	170
12.2.3 Ubiquitination Sites	171
12.2.4 Acetylation Sites	172
12.3 Chaperone Functions	172
12.3.1 Disorder in Protein Chaperones	174
12.3.2 Disorder in RNA Chaperones	174
12.4 Effector Functions	176
12.4.1 Inhibitors	177
12.4.2 Activators	177
12.5 Scavenger Functions	178
12.5.1 Salivary Proline-Rich Glycoproteins	178
12.5.2 Caseins	178
12.5.3 Calsequestrin	179

12.6	Assembler Functions	179
12.6.1	Targeting Activity	179
12.6.2	Assembling Complexes	180
12.6.2.1	HMGA, a Fully Disordered Hub Protein	183
12.6.2.2	MDM2, a Partially Disordered Hub Protein	183
12.6.2.3	Calmodulin, an Ordered Hub Protein	184
12.6.2.4	Disorder and Complex Size	185
12.6.2.5	Scaffold Proteins	185
12.7	Prion Functions	187
12.7.1	Sup35	187
12.7.2	Cytoplasmic Polyadenylation Element Binding Protein	188
13	Evolution and Prevalence of Disorder	189
13.1	Phylogenetic Distribution of Disorder	189
13.1.1	Predicted Disorder in Genomes and Proteomes	189
13.1.2	The Origin of Disordered Proteins in Eukaryotes	191
13.1.3	The Generation of Disordered Domains by Gene Duplication and Module Exchange	192
13.2	Fast Evolution of IDPs by Point Mutations	193
13.2.1	Neutrality in the Evolution of IDPs	194
13.2.2	Disordered Regions May Also Be Conserved	195
13.3	Fast Evolution of IDPs by Repeat Expansion	195
13.3.1	Micro- and Minisatellites in Protein Evolution	196
13.3.1.1	Mechanisms of Repeat Expansion	197
13.3.1.2	Tandem Repeats in the CTD of RNA Polymerase II	198
13.3.1.3	Tandem Repeats in the PEVK Region of Titin	198
13.3.1.4	Tandem Repeats in Prion Protein	199
13.3.2	A Functional Model of Repeat Expansion in IDPs	199
13.4	Fast Evolution and Functionality of Disordered Proteins	200
13.4.1	Retention of Entropic-Chain Functions and Recognition Functions	200
13.4.2	Recognition Another Way: The Lessons from Fuzziness	202
13.4.3	Co-Evolution of IDPs and Their Partners	202
13.5	Structural Variability and Evolvability of New Functions	203
14	Extension of the Structure-Function Paradigm	205
14.1	Functions That Stem Directly from the Disordered State	205
14.2	Recognition Functions: Recognition by Short Motifs	206
14.2.1	Preformed Structural Elements	206
14.2.2	Linear Motifs	208
14.2.3	Molecular Recognition Elements/Features	210
14.2.4	Recognition by Domain-Sized Motifs and Mutual Folding	210

14.2.5	Recognition Interfaces	212
14.2.6	Unification of Concepts?	214
14.3	Disorder-to-Order Transition in Recognition: Mechanistic and Thermodynamic Aspects	214
14.3.1	Site-Directed Mutagenesis Studies of Induced Folding	215
14.3.2	Molecular Dynamics Simulations of Induced Folding	216
14.3.3	NMR Studies of the Mechanism of Induced Folding	217
14.3.4	The Analogy of Folding and Induced Folding	218
14.4	Recognition Functions: Uncoupling Specificity from Binding Strength	219
14.4.1	Disorder May Contribute to Recognition of Specific Sites	219
14.4.2	Disorder May Make Interactions Weaker	220
14.4.3	Strong Multivalent Binding and Weak Aspecific Binding	220
14.5	Implications of Disorder for the Kinetics of Interactions	221
14.5.1	Primary Contact Sites	222
14.5.2	Fly-Casting in Recognition	222
14.6	Adaptability and Moonlighting	223
14.7	Nested Interfaces	225
14.8	Disorder in the Bound State: Fuzziness	226
14.8.1	Structural Polymorphism in the Bound State	226
14.8.2	Clamp-Type of Fuzziness	227
14.8.3	Flanking-Type of Fuzziness	228
14.8.4	Random-Type of Fuzziness	228
14.9	Processivity of Binding	229
14.10	Sequence Independence In Recognition	230
14.11	Ultrasensitivity of Recognition	230
14.11.1	Recognition of Sic1 by Cdc4	231
14.11.2	Regulation of CFTR by Its Disordered R Domain	231
14.11.3	Electrostatics in Ultrasensitivity	232
14.12	Signal Propagation in the Structural Ensemble of IDPs	232
14.12.1	The Signaling Conduit P27 ^{Kip1}	232
14.12.2	Tailored Auto-Activation of WASP	233
14.12.3	Allostery Mediated by Order–Disorder Transitions	234
14.13	Disorder and Alternative Splicing	234
14.14	Molecular Mimicry by a Disordered Region	235
14.15	Entropy Transfer in Chaperone Action	235
15	Structural Disorder and Disease	237
15.1	Structural Disorder and Cancer	237
15.1.1	Disorder in Cancer-Associated Proteins	237
15.1.2	P53	238
15.1.3	Cip/Kip Cdk Inhibitors	240
15.1.4	Breast-Cancer 1	242

15.1.5	Securin (PTTG)	243
15.1.6	Disorder in Proteins Generated by Chromosomal Translocations	244
15.2	Structural Disorder in Proteins Involved in Cardiovascular Diseases, Diabetes, and Autoimmune Diseases	245
15.3	Structural Disorder and Neurodegenerative Diseases	246
15.3.1	Alzheimer's Disease	248
15.3.1.1	A β Peptide	248
15.3.1.2	Tau Protein	248
15.3.2	Parkinson's Disease	249
15.3.2.1	α -Synuclein (NACP)	250
15.3.3	Glutamine-Repeat Diseases	251
15.3.3.1	Huntington's Disease	252
15.3.3.2	Huntingtin	253
15.3.4	Prion Diseases	253
15.3.4.1	Prion Protein	254
15.4	Systemic Amyloidoses	255
15.5	Common Themes in Amyloid Formation	255
15.5.1	Kinetics of Amyloid Formation	256
15.5.2	Disorder in Amyloidogenic Proteins	256
15.5.3	The Structure of the Amyloid	257
15.5.4	Molecular Mechanism of Transition to the Amyloid State	258
15.6	Does Structural Disorder Pose a Danger?	259
15.7	Disorder in Pathogenic Organisms	260
15.8	Rational Drug Design Based on Protein Disorder	262
<i>References</i>		265
<i>Index</i>		313