

# TABLE OF CONTENTS

<b>Chapter 1 Proteins Are Polymers That Fold into Specific Structures</b>		
PROTEINS ARE THE MACHINES THAT PERFORM CELLULAR FUNCTIONS	1	1
PROTEINS HAVE SEQUENCE–STRUCTURE–FUNCTION RELATIONSHIPS	2	2
AMINO ACIDS ARE THE REPEAT UNITS OF PROTEINS	4	4
Amino Acids Are Chiral Molecules	5	5
The 20 Amino Acids Have Different Physical Properties	5	5
In Proteins, Amino Acids Are Linked Together through <i>Peptide Bonds</i>	7	7
Peptide Bonds Are Planar	7	7
The Rotational Freedom around the Backbone Peptide Bond Is Described Using Ramachandran Maps	8	8
Side Chains Adopt Preferred Conformations	9	9
NATIVE PROTEINS HAVE COMPACT WELL-DEFINED 3D STRUCTURES	10	10
Proteins Come in Different Sizes and Shapes	10	10
Native Protein Chains Are Balled Up and Tightly Packed	11	11
Proteins Have Hydrophobic Cores	12	12
The Amino Acids in Native Proteins Are Hydrogen-Bonded to Each Other	13	13
Cysteines Can Form Disulfide Bonds	13	13
PROTEINS HAVE HIERARCHIES OF STRUCTURE	13	13
The Secondary Structures of Proteins Are Helices and Sheets	14	14
<i>Supersecondary Structures</i> , Also Called <i>Structural Motifs</i> , Are Common Combinations of Secondary Structures	18	18
Some Substructures of Proteins Are Compact Functional <i>Domains</i>	18	18
Native Protein Topologies Are Described Using <i>Contact Maps</i>	19	19
How Can You Classify Protein Tertiary Structures?	20	20
Proteins Are Classified by Structural and Evolutionary Properties in the CATH Database	22	22
Quaternary Structures: Higher-Order Structures Result from Noncovalent Assemblies of Multiple Chains	22	22
Domain Swapping Is Another Way That Proteins Can Form Quaternary Structures	23	23
SOME PROTEINS ARE STABLE AND FUNCTION IN THE MEMBRANE ENVIRONMENT	23	23
SOME PROTEINS HAVE FIBROUS STRUCTURES	24	24
Fibrous Proteins Include Coiled Coils and $\beta$ -Helices	25	25
NATIVE PROTEINS ARE CONFORMATIONAL ENSEMBLES	26	26
Proteins Fluctuate around Their Native Structures	26	26
A Protein Can Sample Multiple Substates under Native Conditions	26	26
Some Proteins Are Intrinsically Disordered	27	27
SUMMARY	27	27
REFERENCES	27	27
SUGGESTED READING	28	28
<b>Chapter 2 Proteins Perform Cellular Functions</b>	<b>29</b>	<b>29</b>
PROTEINS CARRY OUT MANY ACTIVITIES IN THE CELL	29	29
A PROTEIN'S FUNCTIONALITY IS ENCODED IN ITS STRUCTURE AND DYNAMICS	30	30
PROTEINS ARE BORN	31	31
Ribosomes Manufacture Proteins	31	31
Molecular Chaperones Are Proteins That Help Other Proteins to Fold Correctly	32	32
PROTEINS WORK FOR A LIVING	33	33
Some Proteins Are Biochemical Catalysts, Called <i>Enzymes</i>	33	33
Some Proteins, Called <i>Motors</i> , Convert Energy into Motion	36	36
Some Proteins Help Move Molecules or Ions across Membranes	40	40
Some Proteins Have Signaling and Regulatory Actions	43	43
Some Proteins Are Structural and Protective Materials	48	48
PROTEINS ARE HEALTHY OR SICK OR DIE	49	49
Proteins Can Misfold, Sometimes in Association with Disease	49	49
Proteasomes Digest Proteins	49	49
SUMMARY	50	50
REFERENCES	50	50
SUGGESTED READING	51	51
<b>Chapter 3 Proteins Have Stable Equilibrium Conformations</b>	<b>53</b>	<b>53</b>
NATIVE AND DENATURED STATES ARE STABLE STATES OF PROTEINS	53	53
Anfinsen's Hypothesis: Native States Are Thermodynamically Stable	54	54
The Basic Experiment of Protein Stability Is Equilibrium Denaturation	55	55
Stabilities and Structures Give Insights about the Driving Forces of Folding	57	57
STATISTICAL MECHANICS IS THE LANGUAGE FOR DESCRIBING PROTEIN STABILITIES	61	61
Why Do Proteins Fold and Unfold? The HP Model	62	62
Proteins Have a Folding Code	64	64
The Collapse of the Chain Helps Stabilize Secondary Structures	65	65
Protein Folding Energy Landscapes Are Funnel-Shaped	66	66
SIMPLE PROTEINS DENATURE WITH TWO-STATE THERMODYNAMICS	67	67
Protein Stability Depends Linearly on Denaturant Concentration	69	69
Protein Stability Is a Nonlinear Function of Temperature	70	70
PROTEINS TEND TO UNFOLD IN ACIDIC OR BASIC SOLUTIONS	71	71
A DENATURED STATE IS A DISTRIBUTION OF CONFORMATIONS	73	73
The Radius of the Denatured Chain Grows as $N^{0.6}$	74	74
Confinement or Crowding Can Increase a Protein's Folding Stability	75	75
SUMMARY	76	76
APPENDIX 3A: A SIMPLE ELECTROSTATIC MODEL OF DENATURATION BY ACIDS AND BASES	76	76
REFERENCES	80	80
SUGGESTED READING	80	80

<b>Chapter 4 Protein Binding Leads to Biological Actions</b>			
INTRODUCTION	81		
BINDING CAN BE MODELED USING BINDING POLYNOMIALS	82		
Binding Polynomials Are Used to Compute Binding Curves	82		
One Ligand Can Bind to, and Saturate, One Site	83		
Binding Polynomials Can Be Constructed Using the Rules of Probability	84		
Michaelis–Menten Kinetics Arises from an Underlying Binding Step	85		
Ligand Binding Rates May Follow Two-State Kinetics, or Be More Complex	87		
IN ALLOSTERY, BINDING IS COUPLED TO CONFORMATIONAL CHANGE	88		
Binding Can Be Cooperative between Two Binding Sites	89		
The Hill Model Describes a Type of Cooperative Binding	89		
Oxygen Binding to Hemoglobin Is a Cooperative Process	90		
The MWC Model Describes Allostery	90		
Binding Affinities Have Their Molecular Basis in Free Energies	94		
INHIBITORS AND ACTIVATORS CAN MODULATE OTHER BINDING ACTIONS	95		
Some Effector Molecules Are Competitive	95		
Prebinding: Ligand Y Binds Only If X Also Binds	97		
Some Effectors Are Noncompetitive	97		
COUPLED BINDING IS KEY TO REGULATION, SIGNALING, AND ENERGY TRANSDUCTION	98		
Biochemical Engines Harness Energetically Downhill Processes to Drive Uphill Processes	98		
BROWNIAN RATCHETS PRODUCE DIRECTED MOTION FROM COUPLED BINDING EVENTS	100		
SUMMARY	104		
APPENDIX 4A: TYPICAL DISSOCIATION CONSTANTS FOR PROTEINS	104		
REFERENCES	105		
SUGGESTED READING	105		
<b>Chapter 5 Folding and Aggregation Are Cooperative Transitions</b>	<b>107</b>		
PROTEINS CAN UNDERGO SHARP TRANSITIONS IN THEIR STRUCTURES OR PROPERTIES	107		
Stable and Unstable States Are Represented on Energy Landscapes	109		
PROTEINS AND PEPTIDES CAN UNDERGO A COOPERATIVE HELIX–COIL TRANSITION	110		
The Schellman Model Describes the Helix–Coil Transition as Nucleation Followed by Growth	113		
Helices Can Be Denatured by Heating	115		
PROTEIN FOLDING COOPERATIVITY ARISES FROM SECONDARY AND TERTIARY INTERACTIONS	116		
Helix–Helix Interactions Contribute to Folding Cooperativity in Helix–Bundle Proteins	117		
PROTEINS CAN ASSEMBLE COOPERATIVELY INTO AGGREGATES, FIBRILS, OR CRYSTALS	120		
Attractive Interactions Can Drive Proteins to Aggregate	121		
Amyloid Peptides Can Assemble to Form Fibrils	121		
SUMMARY	124		
APPENDIX 5A: ADVANCED HELIX–COIL THEORIES	124		
APPENDIX 5B: AMYLOID AGGREGATION THEORY	125		
Most Fibrils Are Relatively Short	125		
Fibril <i>Lengths</i> Increase Sharply with Peptide Concentration	126		
Increasing the Concentration of Peptides Increases the Fibril Concentration	126		
REFERENCES	127		
SUGGESTED READING	127		
<b>Chapter 6 The Principles of Protein Folding Kinetics</b>	<b>129</b>		
THE LEVINTHAL PARADOX MOTIVATED THE SEARCH FOR A PROTEIN FOLDING MECHANISM	129		
FOLDING RATE EXPERIMENTS ARE CAPTURED BY MASS-ACTION MODELS	130		
Small Proteins Fold Rapidly through Two-State Kinetics	131		
Single-Exponential Kinetics Is Characterized through the Concept of a <i>Transition State</i>	133		
Large Proteins Typically Fold via Multi-exponential Kinetics	135		
What Is the Difference between a Kinetic Intermediate State and a Transition State?	136		
RATE MEASUREMENTS GIVE INSIGHTS INTO THE PATHWAYS OF PROTEIN FOLDING	137		
Mutational Studies Can Probe Folding Pathways	139		
Denaturants Can Change Folding Rates: The Chevron Plot	140		
Whole Secondary Structures Often Fold as a Unit	142		
Ultrafast Folders Shed Light on the Speed Limits to Protein Folding	142		
HOW DO PROTEINS FOLD SO FAST? THEY FOLD ON FUNNEL-SHAPED ENERGY LANDSCAPES	143		
What Do You Learn from Folding Funnel?	144		
DIFFERENT PROTEINS CAN FOLD AT VERY DIFFERENT RATES	146		
SUMMARY	148		
APPENDIX 6A: MASTER EQUATIONS DESCRIBE DYNAMICS	148		
APPENDIX 6B: THE ZWANZIG–SZABO–BAGCHI MODEL SHOWS HOW FUNNELS ACCELERATE FOLDING	153		
The Equilibrium Properties of the ZSB Funnel Landscape Model	153		
The ZSB Model Relates Landscape Shape to Folding Speed	154		
The ZSB Model Explains the Unusual Temperature Dependences of Ultrafast Folders	156		
The Foldon Assembly Model Is a Folding Mechanism Variant of the ZSB Model	157		
APPENDIX 6C: PROTEIN FOLDING FUNNELS CAN BE BUMPY: THE SPIN-CLASS MODEL	157		
REFERENCES	159		
SUGGESTED READING	159		
<b>Chapter 7 Proteins Evolve</b>	<b>161</b>		
PROTEINS CHANGE THROUGH EVOLUTIONARY PROCESSES	161		
What Are the Mechanisms of Evolutionary Change?	162		
<i>Homologs</i> Are Proteins with Similar Functions and a Common Ancestor	163		
When Two Protein Sequences Are Similar, Their Structures and Functions Are Usually Similar Too	165		
There Are Three Types of Homolog Relationships: <i>Orthologs</i> , <i>Paralogs</i> , and <i>Xenologs</i>	165		
Gene Duplication Can Explain the High Symmetries in Protein Structures	166		
Evolutionary Processes Can Be <i>Convergent</i> or <i>Divergent</i>	167		
Phylogenetic Trees Show Evolutionary Relationships between Organisms or Sequences	169		
MANY DIFFERENT SEQUENCES FOLD INTO THE SAME NATIVE STRUCTURE	170		
The HP Model Gives Insight into Sequence–Structure Mapping	171		

Protein Structures Are Tolerant of Single Amino Acid Changes	172	TO SIMULATE PROTEIN PHYSICS ON A COMPUTER, YOU NEED A MODEL OF INTERATOMIC ENERGIES	207
Sequence Space Is Filled with Sequences That Collapse to Compact Protein-Like Folds	173	Molecular Energetics Can Be Described by Atomistic Force Fields	207
Proteins and Other Polymers Can Be Designed to Fold to Stable Structures	174	Bond Lengths and Bond Angles Are Treated Using Spring-Like Forces	209
EVOLUTION IS NOT AN ABSTRACTION. IT'S REAL. IT'S HAPPENING NOW	175	Van der Waals Interactions Are Short-Range Attractions and Repulsions	210
Drug Resistance Is an Example of Evolution in Action	175	Charge Interactions Are Modeled Using Coulomb's Law	210
Molecular Clocks: Some Evolutionary Changes Proceed at a Constant Rate	176	Solvent Interactions Are a Major Determinant of Protein Conformations	211
<i>Directed Evolution</i> Is a Way to Improve Proteins in the Laboratory	177	Explicit-water models Represent Waters as Individual Molecules	212
SUMMARY	179	Implicit-water models Treat Water as a Continuous Medium	212
REFERENCES	179	SUMMARY	215
SUGGESTED READING	180	APPENDIX 9A: HOW TO COMPUTE CARTESIAN COORDINATES FROM INTERNAL COORDINATES	215
<b>Chapter 8 Bioinformatics: Insights from Protein Sequences</b>	<b>181</b>	APPENDIX 9B: HOW TO OPTIMALLY SUPERIMPOSE TWO STRUCTURES	216
COMPARING AMINO ACID SEQUENCES GIVES INSIGHT INTO PROTEIN STRUCTURE AND FUNCTION	181	APPENDIX 9C: THE POISSON-BOLTZMANN EQUATION TREATS ELECTROSTATIC INTERACTIONS	217
Sequences Change through Evolutionary Mutations	182	REFERENCES	219
Proteins Having Similar Sequences Usually Have Similar Structures and Functions	182	SUGGESTED READING	220
HOW DO YOU DETERMINE THE RELATEDNESS BETWEEN SEQUENCES?	183	<b>Chapter 10 Molecular Simulations and Conformational Sampling</b>	<b>221</b>
Some Amino Acids Swap More Often Than Others	184	YOU CAN FIND STATES OF LOW ENERGY BY ENERGY MINIMIZATION	222
TO COMPARE SEQUENCES, YOU START WITH GOOD ALIGNMENTS	186	To Compute Forces, Take the Derivative of the Potential Energy	222
How Do You Align One Sequence with Another?	186	Following Gradients Downhill Leads to States of Low Energy	225
BLAST Uses a Query Sequence to Search a Database for Related Sequences	189	MOLECULAR DYNAMICS SIMULATIONS SOLVE NEWTON'S EQUATIONS OF MOTION ITERATIVELY	225
Aligning Multiple Sequences Gives More Insight Than Aligning Two Sequences	190	How Do You Compute a Molecular Dynamics Trajectory?	226
You Can Improve Sequence Alignments by Structure Matching	191	What Time Step $\Delta t$ Should You Use for MD Simulations?	227
HOW DO YOU CONSTRUCT A PHYLOGENETIC TREE?	191	What Is the Computational Time for MD Simulations?	228
EVOLUTION CONSERVES SOME AMINO ACIDS AND CHANGES OTHERS	193	How Do You Analyze Trajectories?	228
You Can Express the Degree of Residue Conservation Using <i>Sequence Entropy</i>	193	Accelerated Sampling Methods Can Explore Larger Actions and Longer Timescales	231
Physical and Biological Factors Affect the Evolutionary Conservation of Amino Acids	195	Stochastic Dynamics Entails Averaging over Solvent Fluctuations	233
Evolutionary Variations Are Sometimes Correlated in the Sequence	196	METROPOLIS MONTE CARLO SIMULATION IS A STOCHASTIC METHOD OF SAMPLING CONFORMATIONS	234
Mutual Information (MI) Measures the Tendencies of Pairs of Amino Acids to Coevolve	197	Here's How to Estimate Averages by Uniform Sampling	234
SUMMARY	197	MMC Is an Efficient Method of Sampling Populated States	235
APPENDIX 8A: EXAMPLE OF A BLAST RUN	198	ADDITIONAL PRINCIPLES LEAD TO IMPROVED COMPUTATIONAL SAMPLING	239
APPENDIX 8B: ESTIMATING EVOLUTIONARY RATES USING A MARKOV MODEL FOR RESIDUE SUBSTITUTIONS	198	Sampling at Higher Temperatures Allows Broader Exploration of Configuration Space	240
REFERENCES	200	The Replica-Exchange Method (REM) Is an Efficient Sampling Method	240
SUGGESTED READING	200	Using Chemical Alchemy, You Can Compute Changes in Free Energy	242
<b>Chapter 9 Protein Geometries and Energetics</b>	<b>201</b>	SUMMARY	243
YOU CAN REPRESENT A PROTEIN STRUCTURE BY ITS ATOMIC COORDINATES	201	APPENDIX 10A: THE VERLET AND LEAPFROG ALGORITHMS GENERATE MD TRAJECTORIES	244
From the Coordinates, You Can Compute the Radius of Gyration	205	The Verlet Algorithm	244
How Similar Are Two Protein Structures? Compute the RMSD between Them	207	The Leapfrog Algorithm	244
		APPENDIX 10B: PERIODIC BOUNDARY CONDITIONS ARE USED IN MD SIMULATIONS	245
		APPENDIX 10C: SOME METHODS FOR ENHANCED SAMPLING	245
		Histogram-Reweighting Methods Let You Predict a Distribution at One Temperature from Another	245

<i>Changes</i> in Free Energy Can Be Computed by Umbrella Sampling, Free-Energy Perturbation, or Thermodynamic Integration Methods	246	MULTIPROTEIN ASSEMBLIES CAN BE STUDIED BY ELASTIC NETWORK MODELS	284
REFERENCES	248	SUMMARY	287
SUGGESTED READING	249	APPENDIX 12A: HERE'S HOW TO EXPRESS THE ELASTIC FREE ENERGY IN TERMS OF THE ADJACENCY MATRIX	287
<b>Chapter 11 Predicting Protein Structures from Sequences</b>	<b>251</b>	APPENDIX 12B: HOW IS $T$ RELATED TO LOCAL PACKING DENSITIES?	288
SOME PROTEINS HAVE COMPUTABLE NATIVE STRUCTURES	251	APPENDIX 12C: HOW DO YOU DETERMINE THE GNM MODES?	289
COMPARATIVE MODELING IS A MAIN TOOL FOR STRUCTURE PREDICTION	252	APPENDIX 12D: NORMAL MODE ANALYSIS	290
Here Is How Homology Modeling Works	252	Normal Mode Analysis Describes the Motions Near Equilibrium	290
A First Step in Modeling a Protein Is Often Determining Its Secondary Structures	254	There Is an Inverse Relationship between Covariance and Stiffness	292
You Can Assemble Protein Structures from <i>Fragments</i> Instead of Secondary Structures	255	You Can Find the Collective Modes Using NMA	292
<i>Threading</i> or <i>Fold Recognition</i> Can Help You Find Local Structures within Your Target Protein	256	APPENDIX 12E: MEAN-SQUARE FLUCTUATIONS IN INTERNAL DISTANCES DEPEND ON THE NETWORK CONNECTIVITY	294
Sequence Co-evolution Analysis Can Help You Predict Protein Structures	256	APPENDIX 12F: HOW CAN YOU COMPARE ONE MOTION WITH ANOTHER?	294
		REFERENCES	295
		SUGGESTED READING	296
<i>STATISTICAL POTENTIALS</i> ARE "ENERGY-LIKE" SCORING FUNCTIONS FOR SELECTING NATIVE-LIKE PROTEIN STRUCTURES	257	<b>Chapter 13 Molecular Modeling for Drug Discovery</b>	
<i>Contact Potentials</i> Express the Noncovalent Pairing Preferences of Amino Acids in Native Structures	259	DRUGS OFTEN ACT BY BINDING TO PROTEINS	297
OTHER COMPUTATIONAL TOOLS CAN ALSO HELP YOU PREDICT NATIVE STRUCTURES	261	PHARMACEUTICAL DISCOVERY IS A MULTISTAGE PIPELINE PROCESS	298
Clustering Algorithms Can Separate Similar Conformations from Different Ones	261	DESIGNING A DRUG REQUIRES OPTIMIZING MULTIPLE PROPERTIES	300
Databases of <i>Decoys</i> Can Help You Develop Energy-Like Scoring Functions	262	What Properties of a Molecule Are Drug-Like?	300
To Compute Protein Structures, You Need Accurate Conformations of Side Chains and Loops	262	What Are the Properties of a "Druggable" Protein?	301
Full Protocols Are Available for Predicting Structures	263	LIGAND-BASED DISCOVERY USES KNOWN LIGANDS TO DESIGN NEW ONES	303
CASP: A COMMUNITY-WIDE EVENT EVALUATES STRUCTURE-PREDICTION METHODS	263	If You Know the Biological Activities of Some Ligands, You Can Estimate the Activities of Others Using QSAR Methods	303
ATOMISTIC PHYSICAL SIMULATIONS CAN PREDICT THE STRUCTURES OF SOME SMALL PROTEINS	264	Similarity Searching Seeks Compounds Similar to the Good Ones You Know Already	304
METHODS ARE AVAILABLE FOR PREDICTING THE STRUCTURES OF PROTEIN COMPLEXES, MULTIMERS, AND ASSEMBLIES	265	A Pharmacophore Is a 3D Arrangement of Properties of Some Atoms	305
SUMMARY	265	Some Drugs Are Developed by Linking Fragments Together	305
APPENDIX 11A: THE MIYAZAWA–JERNIGAN CONTACT-POTENTIAL MATRIX	267	TARGET-BASED DISCOVERY DESIGNS DRUGS BY USING THE STRUCTURE OF A TARGET PROTEIN	305
REFERENCES	267	Docking Is a Fast Way to Find Ligands That Bind to a Given Protein Structure	306
SUGGESTED READING	268	Including Protein Flexibility Can Improve the Modeling of Ligand Binding	308
<b>Chapter 12 Biological Actions Arise from Protein Motions</b>	<b>269</b>	Molecular Dynamics Simulations Can Calculate Binding Free Energies	308
NATIVE PROTEINS HAVE CORRELATED MOTIONS	269	Structure-Based Methods Have Helped to Discover New Drugs	309
ELASTIC NETWORK MODELS USE BEADS AND SPRINGS TO DESCRIBE PROTEIN MOTIONS	270	A MAJOR CLASS OF DRUGS IS THE BIOLOGICS	310
Some Motions of Different Parts of a Structure Are Correlated with Each Other	272	CHALLENGES AND RECENT DEVELOPMENTS IN DRUG DISCOVERY	311
PROTEIN MOTIONS CAN BE OBSERVED IN EXPERIMENTS AND PREDICTED BY THE GNM	273	Drug Resistance Can Be Caused by a Protein Mutation Near a Drug-Binding Site	311
Protein Fluctuations Can Be Observed in NMR Experiments	274	Quantitative Systems Pharmacology Goes beyond "One Target–One Mechanism"	312
Protein Breathing Motions Can Be Observed in Hydrogen Exchange (HX) Experiments	275	Peptides and Macrocycles Can Interfere with Protein–Protein Interactions	313
PROTEIN MOTIONS ARE RELEVANT TO MECHANISMS OF ACTION	278	Sometimes Biological Activity Correlates Better with Ligand–Protein Off-Rates Than with Binding Affinities	314
Protein Motions Can Be Decomposed into a Spectrum of Normal Modes	278	You Can Sometimes <i>Repurpose</i> an Old Drug for a New Medical Indication	314
Directions of Motion Can Be Found Using the Anisotropic Network Model	280	SUMMARY	315
		REFERENCES	316
		SUGGESTED READING	316
		INDEX	317