

CONTENTS

List of Figures	xi
List of Tables	xvii
Foreword	xix
Acknowledgments	xxi
1 Geometric models of protein structure and function prediction	1
1.1 Introduction	1
1.2 Theory and model	2
1.2.1 Idealized ball model	2
1.2.2 Surface models of proteins	3
1.2.3 Geometric constructs	4
1.2.4 Topological structures	6
1.2.5 Metric measurements	9
1.3 Algorithm and computation	12
1.4 Applications	15
1.4.1 Protein packing	15
1.4.2 Predicting protein functions from structures	16
1.5 Discussion and summary	18
References	23
Exercises	26

2	Scoring functions for predicting structure and binding of proteins	29
2.1	Introduction	29
2.2	General framework of scoring function and potential function	31
2.2.1	Protein representation and descriptors	31
2.2.2	Functional form	31
2.2.3	Deriving parameters of potential functions	32
2.3	Statistical method	32
2.3.1	Background	32
2.3.2	Theoretical model	32
2.3.3	Miyazawa-Jernigan contact potential	34
2.3.4	Distance dependent potential function	41
2.3.5	Geometric potential functions	44
2.4	Optimization method	47
2.4.1	Geometric nature of discrimination	48
2.4.2	Optimal linear potential function	50
2.4.3	Optimal nonlinear potential function	51
2.4.4	Deriving optimal nonlinear scoring function	52
2.4.5	Optimization techniques	53
2.5	Applications	53
2.5.1	Protein structure prediction	53
2.5.2	Protein-protein docking prediction	54
2.5.3	Protein design	56
2.5.4	Protein stability and binding affinity	57
2.6	Discussion and summary	58
2.6.1	Knowledge-based statistical potential functions	58
2.6.2	Relationship of knowledge-based energy functions and further development	61
2.6.3	Optimized potential function	63
2.6.4	Data dependency of knowledge-based potentials	64
	References	67
	Exercises	75
3	Sampling techniques: Estimating evolutionary rates and generating molecular structures	77
3.1	Introduction	77
3.2	Principles of Monte Carlo sampling	79
3.2.1	Estimation through sampling from target distribution	79
3.2.2	Rejection sampling	80
3.3	Markov chains and Metropolis Monte Carlo sampling	81
3.3.1	Properties of Markov chains	81
3.3.2	Markov chain Monte Carlo sampling	83

3.4	Sequential Monte Carlo sampling	84
3.4.1	Importance sampling	85
3.4.2	sequential importance sampling	85
3.4.3	Resampling	89
3.5	Applicatoins	90
3.5.1	Markov Chain Monte Carlo for evolutionary rate estimation	90
3.5.2	Sequentail chain growth Monte Carlo for estimating conformational entropy of RNA loops	93
3.6	Discussion and summary	94
References		97
	Exercises	99
4	Stochastic molecular networks	103
4.1	Introduction	103
4.2	Reaction system and discrete chemical master equation	104
4.3	Direct solution of chemical master equation	106
4.3.1	State enumeration with finite buffer	106
4.3.2	Generalization and Multi-Buffer dCME method.	107
4.3.3	Calculation of steady state probability landscape	108
4.3.4	Calculation of dynamically evolving probability landscape	108
4.3.5	Methods for state space truncation for simplification	108
4.4	Quantifying and controlling errors from state space truncation	111
4.5	Approximating discrete chemical master equation	113
4.5.1	Continuous chemical master equation	113
4.5.2	Stochastic differential equation: Fokker-Planck approach	113
4.5.3	Stochastic differential equation: Langevin approach	115
4.5.4	Other approximations	116
4.6	Stochastic simulation	117
4.6.1	Reaction probability	117
4.6.2	Reaction trajectory	117
4.6.3	Probability of reaction trajectory	118
4.6.4	Stochastic simulation algorithm	118
4.7	Applications	120
4.7.1	Probability landscape of a stochastic toggle switch	121
4.7.2	Epigenetic decision network of cellular fate in phage lambda	121
4.8	Discussions and summary	125
References		127
	Exercises	130
5	Cellular interaction networks	133

5.1	Basic Definitions and Graph-theoretic Notions	134
5.1.1	Topological representation	134
5.1.2	Dynamical representation	135
5.1.3	Topological representation of dynamical models	137
5.2	Boolean interaction networks	137
5.3	Signal transduction networks	139
5.3.1	Synthesizing signal transduction networks	140
5.3.2	Collecting data for network synthesis	142
5.3.3	Transitive reduction and pseudo-node collapse	144
5.3.4	Redundancy and degeneracy of networks	149
5.3.5	Random interaction networks and statistical evaluations	153
5.4	Reverse engineering of biological networks	155
5.4.1	Modular response analysis approach	156
5.4.2	Parsimonious combinatorial approaches	163
5.4.3	Evaluation of quality of the reconstructed network	166
	References	168
	Exercises	173
6	Dynamical systems and interaction networks	177
6.1	Some basic control-theoretic concepts	179
6.2	Discrete-time Boolean network models	180
6.3	Artificial neural network models	181
6.3.1	Computational powers of ANNs	183
6.3.2	Reverse engineering of ANNs	184
6.3.3	Applications of ANN models in studying biological networks	185
6.4	Piecewise linear models	186
6.4.1	Dynamics of PL models	187
6.4.2	Biological application of PL models	188
6.5	Monotone systems	193
6.5.1	Definition of monotonicity	194
6.5.2	Combinatorial characterizations and measure of monotonicity	196
6.5.3	Algorithmic issues in computing the degree of monotonicity M	199
	References	201
	Exercises	207
7	Case study of biological models	209
7.1	Segment polarity network models	209
7.1.1	Boolean network model	210

7.1.2	Signal transduction network model	210
7.2	ABA-induced stomatal closure network	211
7.3	Epidermal growth factor receptor signaling network	212
7.4	<i>C. elegans</i> metabolic network	214
7.5	Network for T cell survival and death in large granular lymphocyte leukemia	215
References		216
	Exercises	217
Glossary		219
Index		221