

Contents

About the author	xiii
Preface	xv
Acknowledgments	xvii
1. Introduction to biomolecular simulations	1
1.1 Introduction	1
1.1.1 Simulation methods	2
1.1.2 Force field development	3
1.1.3 Software development	4
1.1.4 Perspectives	4
References	5
Section 1	
Basic knowledge	
2. Molecular mechanics and force field	13
2.1 Introduction	13
2.2 All-atom model	14
2.3 Polarizable model	15
2.4 Coarse-grained model	19
2.5 Summary	23
References	24
3. Quantum chemistry theory	31
3.1 Quantum chemistry theory	31
3.1.1 Hartree-Fock and post-HF method	32
3.1.2 Density-functional theory	37
3.1.3 Semi-empirical method	43
3.2 Conclusion	46
References	46

4. Machine learning	51
4.1 Artificial neural network	51
4.1.1 Neural network	51
4.1.2 Deep learning	53
4.2 Cluster analysis	64
4.2.1 K-means	65
4.2.2 Density-based spatial clustering of applications with noise	66
4.2.3 Spectral clustering	67
4.2.4 FDP	68
4.3 Classification	72
4.3.1 Decision tree	72
4.3.2 Classification and regressions trees	74
4.3.3 Random forest	74
4.4 Dimensionality reduction	74
4.4.1 Principal component analysis	75
4.4.2 T-SNE	77
4.4.3 Uniform manifold approximation and projection	78
References	79

Section 2 Methods and approaches

5. Monte Carlo and molecular dynamics	83
5.1 Monte Carlo and molecular dynamics	83
5.1.1 Lagrangian and Hamiltonian dynamics	83
5.1.2 Discretizing Hamilton's equations	85
5.1.3 Statistical mechanics, partition functions, and probability density functions	86
5.1.4 Monte Carlo simulations and the metropolis method	88
5.1.5 Monte Carlo simulation in isobaric-isothermal ensemble	89
5.1.6 Molecular dynamics simulations	89
5.1.7 Molecular dynamics simulation in the canonical ensemble	91
5.1.8 Molecular dynamics simulation in the isothermal-isobaric ensemble	95
6. Control and adjustment of simulation conditions	99
6.1 Solvent	99
6.1.1 Water models	99
6.1.2 Implicit solvent models	100
6.1.3 Coarse-Grained water models	100
6.1.4 All-Atom water models	101
6.2 Boundary conditions	104

6.2.1 The boundary effect	104
6.2.2 Periodic boundary conditions	104
6.2.3 Cutoff and minimum image convention	104
6.2.4 Limitation of box shape	106
6.3 Restraints and constraints	107
6.3.1 Timestep and constraints	107
6.3.2 Restraints	110
6.4 Temperature and pressure	112
6.4.1 Temperature	113
6.4.2 Pressure	115
References	116
7. Multiscale model	119
7.1 Multiscale model 张跃斌 >15页	119
7.1.1 The quantum mechanics/molecular mechanics methodology	119
7.1.2 Molecular mechanics/coarse grained methods	126
7.1.3 From bead-based models into the continuum regime	135
8. Enhanced sampling	141
8.1 Introduction	141
8.2 Collective variables-based enhanced sampling	141
8.2.1 Umbrella sampling	142
8.2.2 Metadynamics	144
8.2.3 Steered molecular dynamics	146
8.2.4 Adaptive biasing force method	147
8.3 Collective variables free enhanced sampling	148
8.3.1 Replica-exchange molecular dynamics	149
8.3.2 Accelerated molecular dynamics and adaptations	150
8.3.3 Random acceleration molecular dynamics	153
8.4 Combination of enhanced sampling methods	154
8.5 Programs and packages	154
8.6 Summary and outlook	155
Notes	155
References	156
9. Software and hardware	167
9.1 QM software	167
9.1.1 Gaussian	167
9.1.2 MOLCAS	168
9.1.3 Materials studio	169

9.1.4	SIESTA	170
9.1.5	ORCA	170
9.1.6	GAMESS	170
9.1.7	MOLPRO	171
9.2	Software for molecular dynamics simulations	172
9.2.1	AMBER	172
9.2.2	LAMMPS	172
9.2.3	NAMD	173
9.2.4	GROMACS	174
9.2.5	CHARMM	174
9.3	CPU, GPU, FPGA, and ASIC	175
9.3.1	Molecular dynamics simulations algorithms	176
9.3.2	The bottleneck of molecular dynamics simulations	178
9.3.3	Central processing units	178
9.3.4	Graphics processing units	180
9.3.5	Field programmable gate arrays	181
9.3.6	Application specific integrated circuits	183
	References	184

Section 3

Applications and case studies

10.	Protein folding	191
10.1	Introduction	191
10.2	Preparing for an laMD simulation	192
10.3	Reweighting and free energy calculation	197
10.4	Protein folding simulation	201
10.5	Summary	204
	Notes	205
	References	205
11.	RNA folding and structure prediction	207
11.1	Introduction	207
11.1.1	Graph	210
11.1.2	Ernwin and RAGTOP	210
11.1.3	RNAkb	211
11.1.4	SimRNA	211
11.1.5	SPQR	211
11.1.6	IsRNA1 and IsRNA2	212
11.1.7	AIMS_RNA_B3	212
11.1.8	RNApps	213
11.1.9	iFoldRNA	213
11.1.10	HiRe-RNA	214

11.2	Case study: RNA folding simulation based on AIMS_RNA_B3 model	214
11.2.1	Preparation of tools and materials	214
11.3	Data folder	215
11.3.1	Secondary structure predictions	215
11.3.2	Generate initial conformation using targeted molecular dynamics simulation	215
11.3.3	All atom molecular dynamics simulation	217
11.4	Trajectory converter	217
11.5	Cluster analysis	217
11.6	Packed all-atom nonbonding parameters	218
11.6.1	Fitting the coarse grain force field	218
11.7	Fitting van der Waals parameters	218
11.8	Fitting electrostatic parameters	218
11.9	Generate coarse-grained topology	219
11.10	Fitting bond parameters (bond length, angle, and dihedral)	219
11.10.1	Coarse-grained simulations	219
11.11	Cluster analyses	219
11.11.1	CG conformation back-mapping to AA conformation	220
11.11.2	Supplement of targeted MD	220
11.12	Summary	221
	References	221
12.	Enzyme catalysis	227
12.1	Background	227
12.2	Modeling details	228
12.2.1	Prepare the force field file for small organic molecule	230
12.2.2	System setting-up	230
12.2.3	Pre-equil and MD simulation	231
12.2.4	QM/MM-MD	233
12.2.5	SMD and umbrella sampling	235
12.2.6	PMF	235
12.3	Results and discussion	237
12.3.1	RMSD	237
12.3.2	Reaction path	237
12.4	Conclusion	239
	References	240
13.	Post-translational modification of proteins	243
13.1	Preparation of simulation system	244
13.2	Molecular dynamics simulations	244
13.3	Trajectories analysis	245

13.3.1	T255 glycosylation decreases the distance between N lobe and C lobe of PGK1 protein	249
13.3.2	T255 glycosylation destroys the interaction between F258 and V217 as well as A218	251
References		253
14.	Regulation of small molecule on proteins	255
14.1	Introduction	255
14.2	The analysis of molecular dynamics simulation	256
14.2.1	Protein and simulation stability analysis	257
14.2.2	Clustering analysis	257
14.2.3	Principal-component analysis	258
14.2.4	Correlations function analysis	258
14.2.5	Cross-correlation analysis	259
14.2.6	MM-PB/GBSA calculation	259
14.3	Examples of regulation of small molecules on proteins	260
14.3.1	Regulation of UDP-glucose on HuR	260
14.3.2	Regulation of 3PG on PGK1	262
14.3.3	Regulation of aKG on IKKa	266
References		269
15.	Recognition of protein with nucleic acid	273
15.1	Preparation of simulation system	274
15.2	Molecular dynamics simulations	274
15.3	Trajectories analysis	277
15.3.1	Normal analysis of cGAS protein	279
15.3.2	Normal analysis of nucleic acid	281
15.3.3	Interaction between cGAS and nucleic acid	283
References		284
16.	Dynamics and functions of membrane proteins	287
16.1	Introduction	287
16.2	Methodology and analysis	288
16.2.1	Single-bilayer system preparation	288
16.2.2	Double-bilayer system preparation	288
16.2.3	Simulation setup	290
16.3	Electric potential calculation in membrane systems	291
16.4	Ion transport and conduction	292
16.5	PMF calculations of GIRK2 in different complex states	294
16.6	Maximum conductance estimation from umbrella sampling	297
16.7	Global functional motions of GIRK2	299

Appendix	301
References	303
17. Assembly and functions of complex of multiple components	307
17.1 Preparing AlphaFold2 and AF2Complex environment	308
17.2 The assembly of the entire NDH-1L complex using AF2Complex	310
17.3 The assembly of the homotrimer of cyanobacterial photosystem I using AF2Complex	316
References	322
18. Protein small molecule design	323
18.1 Setup scaffold library	323
18.2 Software required	324
18.3 Calculations	324
18.3.1 Preparation	324
18.3.2 RifGen	327
18.3.3 Intermediate process	327
18.3.4 PatchDock	328
18.3.5 RifDock	328
18.3.6 Predictor	331
18.3.7 FastDesign	333
18.3.8 Motif grafting	334
18.3.9 Redo and final filter	336
18.4 Conclusions	337
Section 4	
Future directions and perspectives	
19. Molecule scale	341
References	346
20. System scale	351
20.1 Physical aspect	351
20.2 Time-scale aspect	354
20.3 Development and future	357
References	358
Index	363