

Molecular Modelling

PRINCIPLES AND APPLICATIONS

Second edition

Andrew R. Leach

Glaxo Wellcome Research and Development



Harlow, England • London • New York • Boston • San Francisco • Toronto • Sydney • Singapore • Hong Kong
Tokyo • Seoul • Taipei • New Delhi • Cape Town • Madrid • Mexico City • Amsterdam • Munich • Paris • Milan

Contents

Preface to the Second Edition	xiii
Preface to the First Edition	xv
Symbols and Physical Constants	xvii
Acknowledgements	xxi
1 Useful Concepts in Molecular Modelling	1
1.1 Introduction	1
1.2 Coordinate Systems	2
1.3 Potential Energy Surfaces	4
1.4 Molecular Graphics	5
1.5 Surfaces	6
1.6 Computer Hardware and Software	8
1.7 Units of Length and Energy	9
1.8 The Molecular Modelling Literature	9
1.9 The Internet	9
1.10 Mathematical Concepts	10
Further Reading	24
References	24
2 An Introduction to Computational Quantum Mechanics	26
2.1 Introduction	26
2.2 One-electron Atoms	30
2.3 Polyelectronic Atoms and Molecules	34
2.4 Molecular Orbital Calculations	41
2.5 The Hartree-Fock Equations	51
2.6 Basis Sets	65
2.7 Calculating Molecular Properties Using <i>ab initio</i> Quantum Mechanics	74
2.8 Approximate Molecular Orbital Theories	86
2.9 Semi-empirical Methods	86
2.10 Hückel Theory	99
2.11 Performance of Semi-empirical Methods	102
Appendix 2.1 Some Common Acronyms Used in Computational Quantum Chemistry	104
Further Reading	105
References	105

3	Advanced <i>ab initio</i> Methods, Density Functional Theory and Solid-state Quantum Mechanics	108
3.1	Introduction	108
3.2	Open-shell Systems	108
3.3	Electron Correlation	110
3.4	Practical Considerations When Performing <i>ab initio</i> Calculations	117
3.5	Energy Component Analysis	122
3.6	Valence Bond Theories	124
3.7	Density Functional Theory	126
3.8	Quantum Mechanical Methods for Studying the Solid State	138
3.9	The Future Role of Quantum Mechanics: Theory and Experiment Working Together	160
	Appendix 3.1 Alternative Expression for a Wavefunction Satisfying Bloch's Function	161
	Further Reading	161
	References	162
4	Empirical Force Field Models: Molecular Mechanics	165
4.1	Introduction	165
4.2	Some General Features of Molecular Mechanics Force Fields	168
4.3	Bond Stretching	170
4.4	Angle Bending	173
4.5	Torsional Terms	173
4.6	Improper Torsions and Out-of-plane Bending Motions	176
4.7	Cross Terms: Class 1, 2 and 3 Force Fields	178
4.8	Introduction to Non-bonded Interactions	181
4.9	Electrostatic Interactions	181
4.10	Van der Waals Interactions	204
4.11	Many-body Effects in Empirical Potentials	212
4.12	Effective Pair Potentials	214
4.13	Hydrogen Bonding in Molecular Mechanics	215
4.14	Force Field Models for the Simulation of Liquid Water	216
4.15	United Atom Force Fields and Reduced Representations	221
4.16	Derivatives of the Molecular Mechanics Energy Function	225
4.17	Calculating Thermodynamic Properties Using a Force Field	226
4.18	Force Field Parametrisation	228
4.19	Transferability of Force Field Parameters	231
4.20	The Treatment of Delocalised π Systems	233
4.21	Force Fields for Inorganic Molecules	234
4.22	Force Fields for Solid-state Systems	236
4.23	Empirical Potentials for Metals and Semiconductors	240
	Appendix 4.1 The Interaction Between Two Drude Molecules	246
	Further Reading	247
	References	247

5	Energy Minimisation and Related Methods for Exploring the Energy Surface	253
5.1	Introduction	253
5.2	Non-derivative Minimisation Methods	258
5.3	Introduction to Derivative Minimisation Methods	261
5.4	First-order Minimisation Methods	262
5.5	Second Derivative Methods: The Newton-Raphson Method	267
5.6	Quasi-Newton Methods	268
5.7	Which Minimisation Method Should I Use?	270
5.8	Applications of Energy Minimisation	273
5.9	Determination of Transition Structures and Reaction Pathways	279
5.10	Solid-state Systems: Lattice Statics and Lattice Dynamics	295
	Further Reading	300
	References	301
6	Computer Simulation Methods	303
6.1	Introduction	303
6.2	Calculation of Simple Thermodynamic Properties	307
6.3	Phase Space	312
6.4	Practical Aspects of Computer Simulation	315
6.5	Boundaries	317
6.6	Monitoring the Equilibration	321
6.7	Truncating the Potential and the Minimum Image Convention	324
6.8	Long-range Forces	334
6.9	Analysing the Results of a Simulation and Estimating Errors	343
	Appendix 6.1 Basic Statistical Mechanics	347
	Appendix 6.2 Heat Capacity and Energy Fluctuations	348
	Appendix 6.3 The Real Gas Contribution to the Virial	349
	Appendix 6.4 Translating Particle Back into Central Box for Three Box Shapes	350
	Further Reading	351
	References	351
7	Molecular Dynamics Simulation Methods	353
7.1	Introduction	353
7.2	Molecular Dynamics Using Simple Models	353
7.3	Molecular Dynamics with Continuous Potentials	355
7.4	Setting up and Running a Molecular Dynamics Simulation	364
7.5	Constraint Dynamics	368
7.6	Time-dependent Properties	374
7.7	Molecular Dynamics at Constant Temperature and Pressure	382
7.8	Incorporating Solvent Effects into Molecular Dynamics: Potentials of Mean Force and Stochastic Dynamics	387
7.9	Conformational Changes from Molecular Dynamics Simulations	392
7.10	Molecular Dynamics Simulations of Chain Amphiphiles	394

Appendix 7.1	Energy Conservation in Molecular Dynamics	405
	Further Reading	406
	References	406
8	Monte Carlo Simulation Methods	410
8.1	Introduction	410
8.2	Calculating Properties by Integration	412
8.3	Some Theoretical Background to the Metropolis Method	414
8.4	Implementation of the Metropolis Monte Carlo Method	417
8.5	Monte Carlo Simulation of Molecules	420
8.6	Models Used in Monte Carlo Simulations of Polymers	423
8.7	'Biased' Monte Carlo Methods	432
8.8	Tackling the Problem of Quasi-ergodicity: J-walking and Multicanonical Monte Carlo	433
8.9	Monte Carlo Sampling from Different Ensembles	438
8.10	Calculating the Chemical Potential	442
8.11	The Configurational Bias Monte Carlo Method	443
8.12	Simulating Phase Equilibria by the Gibbs Ensemble Monte Carlo Method	450
8.13	Monte Carlo or Molecular Dynamics?	452
	Appendix 8.1 The Marsaglia Random Number Generator	453
	Further Reading	454
	References	454
9	Conformational Analysis	457
9.1	Introduction	457
9.2	Systematic Methods for Exploring Conformational Space	458
9.3	Model-building Approaches	464
9.4	Random Search Methods	465
9.5	Distance Geometry	467
9.6	Exploring Conformational Space Using Simulation Methods	475
9.7	Which Conformational Search Method Should I Use? A Comparison of Different Approaches	476
9.8	Variations on the Standard Methods	477
9.9	Finding the Global Energy Minimum: Evolutionary Algorithms and Simulated Annealing	479
9.10	Solving Protein Structures Using Restrained Molecular Dynamics and Simulated Annealing	483
9.11	Structural Databases	489
9.12	Molecular Fitting	490
9.13	Clustering Algorithms and Pattern Recognition Techniques	491
9.14	Reducing the Dimensionality of a Data Set	497
9.15	Covering Conformational Space: Poling	499
9.16	A 'Classic' Optimisation Problem: Predicting Crystal Structures	501

Further Reading	505
References	506
10 Protein Structure Prediction, Sequence Analysis and Protein Folding	509
10.1 Introduction	509
10.2 Some Basic Principles of Protein Structure	513
10.3 First-principles Methods for Predicting Protein Structure	517
10.4 Introduction to Comparative Modelling	522
10.5 Sequence Alignment	522
10.6 Constructing and Evaluating a Comparative Model	539
10.7 Predicting Protein Structures by 'Threading'	545
10.8 A Comparison of Protein Structure Prediction Methods: CASP	547
10.9 Protein Folding and Unfolding	549
Appendix 10.1 Some Common Abbreviations and Acronyms Used in Bioinformatics	553
Appendix 10.2 Some of the Most Common Sequence and Structural Databases Used in Bioinformatics	555
Appendix 10.3 Mutation Probability Matrix for 1 PAM	556
Appendix 10.4 Mutation Probability Matrix for 250 PAM	557
Further Reading	557
References	558
11 Four Challenges in Molecular Modelling: Free Energies, Solvation, Reactions and Solid-state Defects	563
11.1 Free Energy Calculations	563
11.2 The Calculation of Free Energy Differences	564
11.3 Applications of Methods for Calculating Free Energy Differences	569
11.4 The Calculation of Enthalpy and Entropy Differences	574
11.5 Partitioning the Free Energy	574
11.6 Potential Pitfalls with Free Energy Calculations	577
11.7 Potentials of Mean Force	580
11.8 Approximate/'Rapid' Free Energy Methods	585
11.9 Continuum Representations of the Solvent	592
11.10 The Electrostatic Contribution to the Free Energy of Solvation: The Born and Onsager Models	593
11.11 Non-electrostatic Contributions to the Solvation Free Energy	608
11.12 Very Simple Solvation Models	609
11.13 Modelling Chemical Reactions	610
11.14 Modelling Solid-state Defects	622
Appendix 11.1 Calculating Free Energy Differences Using Thermodynamic Integration	630
Appendix 11.2 Using the Slow Growth Method for Calculating Free Energy Differences	631

Appendix 11.3	Expansion of Zwanzig Expression for the Free Energy Difference for the Linear Response Method	631
	Further Reading	632
	References	633
12	The Use of Molecular Modelling and Chemoinformatics to Discover and Design New Molecules	640
12.1	Molecular Modelling in Drug Discovery	640
12.2	Computer Representations of Molecules, Chemical Databases and 2D Substructure Searching	642
12.3	3D Database Searching	647
12.4	Deriving and Using Three-dimensional Pharmacophores	648
12.5	Sources of Data for 3D Databases	659
12.6	Molecular Docking	661
12.7	Applications of 3D Database Searching and Docking	667
12.8	Molecular Similarity and Similarity Searching	668
12.9	Molecular Descriptors	668
12.10	Selecting 'Diverse' Sets of Compounds	680
12.11	Structure-based <i>De Novo</i> Ligand Design	687
12.12	Quantitative Structure-Activity Relationships	695
12.13	Partial Least Squares	706
12.14	Combinatorial Libraries	711
	Further Reading	719
	References	720
Index		727