

Molecular modelling: principles and applications – 2nd. Ed

Leach, Andrew R.

ISBN 978-0-582-38210-6

Contents

1. USEFUL CONCEPTS IN MOLECULAR MODELLING

- 1.1 Introduction
- 1.2 Coordinate systems
- 1.3 Potential energy surfaces
- 1.4 Molecular graphics
- 1.5 Surfaces
- 1.6 Computer hardware and software
- 1.7 Units of length and energy
- 1.8 The molecular modelling literature
- 1.9 The Internet
- 1.10 Mathematical concepts
 - 1.10.1 Series expansions
 - 1.10.2 Vectors
 - 1.10.3 Matrices, eigenvectors and eigenvalues
 - 1.10.4 Complex numbers
 - 1.10.5 Lagrange multipliers
 - 1.10.6 Multiple integrals
 - 1.10.7 Some basic elements of statistics
 - 1.10.8 The Fourier series, Fourier transform and fast-Fourier transform

2. AN INTRODUCTION TO COMPUTATIONAL QUANTUM MECHANICS

- 2.1 Introduction
 - 2.1.1 Operators
 - 2.1.2 Atomic units
 - 2.1.3 Exact solutions to the Schrödinger equation
- 2.2 One-electron atoms
- 2.3 Polyelectronic atoms and molecules
 - 2.3.1 The Born-Oppenheimer approximation
 - 2.3.2 The helium atom
 - 2.3.3 General polyelectronic systems and Slater determinants
- 2.4 Molecular orbital calculations
 - 2.4.1 Calculating the energy from the wavefunction: the hydrogen molecule
 - 2.4.2 The energy of a general polyelectronic system
 - 2.4.3 Shorthand representations of the one- and two-electron integrals
 - 2.4.4 The energy of a closed-shell system
- 2.5 The Hartree-Fock equations
 - 2.5.1 Hartree-Fock calculations for atoms and Slater's rules
 - 2.5.2 Linear combination of atomic orbitals (LCAO) in Hartree-Fock theory
 - 2.5.3 Closed-shell systems and the Roothaan-Hall equations

- 2.5.4 Solving the Roothaan-Hall equations
- 2.5.5 A simple illustration of the Roothaan-Hall approach
- 2.5.6 Application of the Hartree-Fock equations to molecular systems
- 2.6 Basis sets
 - 2.6.1 Creating a basis set
- 2.7 Calculating molecular properties using ab initio quantum mechanics
 - 2.7.1 Setting up the calculation and the choice of coordinates
 - 2.7.2 Energies, Koopman's theorem and ionisation potentials
 - 2.7.3 Calculation of electric multipoles
 - 2.7.4 The total electron density distribution and molecular orbitals
 - 2.7.5 Population analysis
 - 2.7.6 Mulliken and Löwdin population analysis
 - 2.7.7 Partitioning electron density: the theory of atoms in molecules
 - 2.7.8 Bond orders
 - 2.7.9 Electrostatic potentials
 - 2.7.10 Thermodynamic and structural properties
- 2.8 Approximate molecular orbital theories
- 2.9 Semi-empirical methods
 - 2.9.1 Zero-differential overlap
 - 2.9.2 CNDO
 - 2.9.3 INDO
 - 2.9.4 NDDO
 - 2.9.5 MINDO/3
 - 2.9.6 MNDO
 - 2.9.7 AM1
 - 2.9.8 PM3
 - 2.9.9 SAM1
 - 2.9.10 Programs for semi-empirical quantum mechanical calculations
- 2.10 Hückel theory
 - 2.10.1 Extended Hückel theory
- 2.11 Performance of semi-empirical methods
- Appendix 2.1 Some Common Acronyms Used in Computational Quantum Chemistry

3. ADVANCED AB INITIO METHODS, DENSITY FUNCTIONAL THEORY AND SOLID-STATE QUANTUM MECHANICS

- 3.1 Introduction
- 3.2 Open-shell systems
- 3.3 Electron correlation
 - 3.3.1 Configuration interaction
 - 3.3.2 Many body perturbation theory
- 3.4 Practical considerations when performing ab initio calculations
 - 3.4.1 Convergence of self-consistent field calculations
 - 3.4.2 The direct SCF method
 - 3.4.3 Calculating derivatives of the energy
 - 3.4.4 Basis set superposition error
- 3.5 Energy component analysis
 - 3.5.1 Morokuma analysis of the water dimer
- 3.6 Valence bond theories
- 3.7 Density functional theory
 - 3.7.1 Spin-polarised density functional theory
 - 3.7.2 The exchange-correlation functional

- 3.7.3 Beyond the local density approximation: gradient-corrected functionals
- 3.7.4 Hybrid Hartree-Fock/Density Functional Methods
- 3.7.5 Performance and applications of density functional theory
- 3.8 Quantum mechanical methods for studying the solid-state
 - 3.8.1 Introduction
 - 3.8.2 Band theory and orbital-based approaches
 - 3.8.3 The periodic Hartree-Fock approach to studying the solid state
 - 3.8.4 The nearly-free electron approximation
 - 3.8.5 The Fermi surface and density of states
 - 3.8.6 Density Functional Methods for studying the solid state: plane waves and pseudopotentials
 - 3.8.7 Application of solid-state quantum mechanics to the group 14 elements
- 3.9 The future role of quantum mechanics: theory and experiment working together

Appendix 3.1 Alternative Expression for a Wavefunction Satisfying Bloch's Function

4. EMPIRICAL FORCE FIELD MODELS: MOLECULAR MECHANICS

- 4.1 Introduction
 - 4.1.1 A simple molecular mechanics force field
- 4.2 Some general features of molecular mechanics force fields
- 4.3 Bond stretching
- 4.4 Angle bending
- 4.5 Torsional terms
- 4.6 Improper torsions and out-of-plane bending motions
- 4.7 Cross terms Class 1, 2 and 3 force fields
- 4.8 Introduction to non-bonded interactions
- 4.9 Electrostatic interactions
 - 4.9.1 The central multipole expansion
 - 4.9.2 Point-charge electrostatic models
 - 4.9.3 Calculating partial atomic charges
 - 4.9.4 Charges derived from the molecular electrostatic potential
 - 4.9.5 Deriving charge models for large systems
 - 4.9.6 Rapid methods for calculating atomic charges
 - 4.9.7 Beyond partial atomic charge models
 - 4.9.8 Distributed multipole models
 - 4.9.9 Using charge schemes to study aromatic-aromatic interactions
 - 4.9.10 Polarisation
 - 4.9.11 Solvent dielectric models
- 4.10 van der Waals interactions
 - 4.10.1 Dispersive interactions
 - 4.10.2 The repulsive contribution
 - 4.10.3 Modelling van der Waals interactions
 - 4.10.4 van der Waals interactions in polyatomic systems
 - 4.10.5 Reduced units
- 4.11 Many-body effects in empirical potentials
- 4.12 Effective pair potentials
- 4.13 Hydrogen bonding in molecular mechanics
- 4.14 Force field models for the simulation of liquid water
 - 4.14.1 Simple water models
 - 4.14.2 Polarisable water models
 - 4.14.3 Ab initio potentials for water
- 4.15 United atom force fields and reduced representations
 - 4.15.1 Other simplified models

- 4.16 Derivatives of the molecular mechanics energy function
- 4.17 Calculating thermodynamic properties using a force field
- 4.18 Force field parametrisation
- 4.19 Transferability of force field parameters
- 4.20 The treatment of delocalised π -systems
- 4.21 Force fields for inorganic molecules
- 4.22 Force fields for solid-state systems
 - 4.22.1 Covalent solids: zeolites
 - 4.22.2 Ionic solids
- 4.23 Empirical potentials for metals and semiconductors
- Appendix 4.1 The Interaction Between Two Drude Molecules

5. ENERGY MINIMISATION AND RELATED METHODS FOR EXPLORING THE ENERGY SURFACE

- 5.1 Introduction
 - 5.1.1 Energy minimisation: statement of the problem
 - 5.1.2 Derivatives
- 5.2 Non-derivative minimisation methods
 - 5.2.1 The simplex method
 - 5.2.2 The sequential univariate method
- 5.3 Introduction to derivative minimisation methods
- 5.4 First-order minimisation methods
 - 5.4.1 The steepest descents method
 - 5.4.2 Line search in one dimension
 - 5.4.3 Arbitrary step approach
 - 5.4.4 Conjugate gradients minimisation
- 5.5 Second derivative methods: the Newton-Raphson method
 - 5.5.1 Variants on the Newton-Raphson method
- 5.6 Quasi-Newton methods
- 5.7 Which minimisation method should I use?
 - 5.7.1 Distinguishing between minima, maxima and saddle points
 - 5.7.2 Convergence criteria
- 5.8 Applications of energy minimisation
 - 5.8.1 Normal mode analysis
 - 5.8.2 The study of intermolecular processes
- 5.9 Determination of transition structures and reaction pathways
 - 5.9.1 Methods to locate saddle points
 - 5.9.2 Reaction path following
 - 5.9.3 Transition structures and reaction pathways for large systems
 - 5.9.4 The transition structures of pericyclic reactions
- 5.10 Solid-state systems: lattice statics and lattice dynamics

6. COMPUTER SIMULATION METHODS

- 6.1 Introduction
 - 6.1.1 Time averages, ensemble averages and some historical background
 - 6.1.2 A brief description of the molecular dynamics method
 - 6.1.3 The basic elements of the Monte Carlo method
 - 6.1.4 Differences between the molecular dynamics and Monte Carlo methods

- 6.2 Calculation of simple thermodynamic properties
 - 6.2.1 Energy
 - 6.2.2 Heat capacity
 - 6.2.3 Pressure
 - 6.2.4 Temperature
 - 6.2.5 Radial distribution functions
- 6.3 Phase space
- 6.4 Practical aspects of computer simulation
 - 6.4.1 Setting up and running a simulation
 - 6.4.2 Choosing the initial configuration
- 6.5 Boundaries
 - 6.5.1 Periodic boundary conditions
 - 6.5.2 Non-periodic boundary methods
- 6.6 Monitoring the equilibration
- 6.7 Truncating the potential and the minimum image convention
 - 6.7.1 Non-bonded neighbour lists
 - 6.7.2 Group-based cutoffs
 - 6.7.3 Problems with cutoffs and how to avoid them
- 6.8 Long-range forces
 - 6.8.1 The Ewald summation method
 - 6.8.2 The reaction field and image charge methods
 - 6.8.3 The cell multipole method for non-bonded interactions
- 6.9 Analysing the results of a simulation and estimating errors
- Appendix 6.1 Basic Statistical Mechanics
- Appendix 6.2 Heat Capacity and Energy Fluctuations
- Appendix 6.3 The Real Gas Contribution to the Virial
- Appendix 6.4 Translating Particle Back into Central Box

7. MOLECULAR DYNAMICS SIMULATION METHODS

- 7.1 Introduction
- 7.2 Molecular dynamics using simple models
- 7.3 Molecular dynamics with continuous potentials
 - 7.3.1 Finite difference methods
 - 7.3.2 Predictor-corrector integration methods
 - 7.3.3 Which integration algorithm is most appropriate?
 - 7.3.4 Choosing the time step
 - 7.3.5 Multiple time step dynamics
- 7.4 Setting up and running a molecular dynamics simulation
 - 7.4.1 Calculating the temperature
- 7.5 Constraint dynamics
- 7.6 Time-dependent properties
 - 7.6.1 Correlation functions
 - 7.6.2 Orientational correlation functions
 - 7.6.3 Transport properties
- 7.7 Molecular dynamics at constant temperature and pressure
 - 7.7.1 Constant temperature dynamics
 - 7.7.2 Constant pressure dynamics
- 7.8 Incorporating solvent effects into molecular dynamics: potentials of mean force and stochastic dynamics
 - 7.8.1 Practical aspects of stochastic dynamics simulations
- 7.9 Conformational changes from molecular dynamics simulations

- 7.10 Molecular dynamics simulations of chain amphiphiles
 - 7.10.1 Simulation of lipids
 - 7.10.2 Simulations of Langmuir-Blodgett films
 - 7.10.3 Mesoscale modelling: Dissipative Particle Dynamics
- Appendix 7.1 Energy Conservation in Molecular Dynamics

8. MONTE CARLO SIMULATION METHODS

- 8.1 Introduction
- 8.2 Calculating properties by integration
- 8.3 Some theoretical background to the Metropolis method
- 8.4 Implementation of the Metropolis Monte Carlo method
 - 8.4.1 Random number generators
- 8.5 Monte Carlo simulation of molecules
 - 8.5.1 Rigid molecules
 - 8.5.2 Monte Carlo simulations of flexible molecules
- 8.6 Models used in Monte Carlo simulations of polymers
 - 8.6.1 Lattice models of polymers
 - 8.6.2 'Continuous' polymer models
- 8.7 'Biased' Monte Carlo methods
- 8.8 Tackling the problem of quasi-ergodicity: J-walking and multicanonical Monte Carlo
 - 8.8.1 J-walking
 - 8.8.2 The multicanonical Monte Carlo method
- 8.9 Monte Carlo sampling from different ensembles
 - 8.9.1 Grand canonical Monte Carlo simulations
 - 8.9.2 Grand canonical Monte Carlo simulations of adsorption processes
- 8.10 Calculating the chemical potential
- 8.11 The configurational bias Monte Carlo method
 - 8.11.1 Applications of the configurational bias Monte Carlo method
- 8.12 Simulating phase equilibria by the Gibbs ensemble Monte Carlo method
- 8.13 Monte Carlo or molecular dynamics?
- Appendix 8.1 The Marsaglia Random Number Generator

9. CONFORMATIONAL ANALYSIS

- 9.1 Introduction
- 9.2 Systematic methods for exploring conformational space
- 9.3 Model-building approaches
- 9.4 Random search methods
- 9.5 Distance geometry
 - 9.5.1 The use of distance geometry in NMR
- 9.6 Exploring conformational space using simulation methods
- 9.7 Which conformational search method should I use? A comparison of different approaches
- 9.8 Variations upon the standard methods
 - 9.8.1 The systematic unbounded multiple minimum method (SUMM)
 - 9.8.2 Low-Mode Search
- 9.9 Finding the global energy minimum: Evolutionary algorithms and simulated annealing
 - 9.9.1 Genetic and evolutionary algorithms
 - 9.9.2 Simulated annealing

9.10 Solving protein structures using restrained molecular dynamics and simulated annealing

9.10.1 X-ray crystallographic refinement

9.10.2 Molecular dynamics refinement of NMR data

9.10.3 Time-averaged NMR refinement

9.11 Structural databases

9.12 Molecular fitting

9.13 Clustering algorithms and pattern recognition techniques

9.14 Reducing the dimensionality of a data set

9.14.1 Principal components analysis

9.15 Covering conformational space: poling

9.16 A "classic" optimisation problem: predicting crystal structures

10. PROTEIN STRUCTURE PREDICTION, SEQUENCE ANALYSIS AND PROTEIN FOLDING

10.1 Introduction

10.2 Some basic principles of protein structure

10.2.1 The hydrophobic effect

10.3 First-principles methods for predicting protein structure

10.3.1 Lattice models for investigating protein structure

10.3.2 Rule-based approaches using secondary structure prediction

10.4 Introduction to comparative modelling

10.5 Sequence alignment

10.5.1 Dynamic programming and the Needleman-Wunsch algorithm

10.5.2 The Smith-Waterman algorithm

10.5.3 Heuristic search methods: FASTA and BLAST

10.5.4 Multiple sequence alignment

10.5.5 Protein structure alignment and structural databases

10.6 Constructing and evaluating a comparative model

10.7 Predicting protein structures by 'threading'

10.8 A comparison of protein structure prediction methods: CASP

10.8.1 Automated protein modelling

10.9 Protein folding and unfolding

Appendix 10.1 Some Common Abbreviations and Acronyms Used in Bioinformatics

Appendix 10.2 Some of the Most Common Sequence and Structural Databases Used in Bioinformatics

Appendix 10.3 Mutation Probability Matrix for 1 PAM

Appendix 10.4 Mutation Probability Matrix for 250 PAM

11. FOUR CHALLENGES IN MOLECULAR MODELLING: FREE ENERGIES, SOLVATION, REACTIONS AND SOLID-STATE DEFECTS

11.1 Free energy calculations

11.1.1 The difficulty of calculating free energies by computer

11.2 The calculation of free energy differences

11.2.1 Thermodynamic perturbation

11.2.2 Implementation of free energy perturbation

11.2.3 Thermodynamic integration

11.2.4 The 'slow growth' method

11.3 Applications of methods for calculating free energy differences

11.3.1 Thermodynamic cycles

- 11.3.2 Applications of the thermodynamic cycle perturbation method
- 11.3.3 The calculation of absolute free energies
- 11.4 The calculation of enthalpy and entropy differences
- 11.5 Partitioning the free energy
- 11.6 Potential pitfalls with free energy calculations
 - 11.6.1 Implementation aspects
- 11.7 Potentials of mean force
 - 11.7.1 Umbrella sampling
 - 11.7.2 Calculating the potential of mean force for flexible molecules
- 11.8 Approximate/"rapid" free energy methods
- 11.9 Continuum representations of the solvent
 - 11.9.1 Thermodynamic background
- 11.10 The electrostatic contribution to the free energy of solvation: the Born and Onsager models
 - 11.10.1 Calculating the electrostatic contribution via quantum mechanics
 - 11.10.2 Continuum models for molecular mechanics
 - 11.10.3 The Langevin dipole model
 - 11.10.4 Methods based upon the Poisson-Boltzmann equation
 - 11.10.5 Applications of finite difference Poisson-Boltzmann calculations
- 11.11 Non-electrostatic contributions to the solvation free energy
- 11.12 Very simple solvation models
- 11.13 Modelling chemical reactions
 - 11.13.1 Empirical approaches to simulating reactions
 - 11.13.2 The potential of mean force of a reaction
 - 11.13.3 Combined quantum mechanical/molecular mechanical approaches
 - 11.13.4 Ab initio molecular dynamics and the Car-Parrinello method
 - 11.13.5 Examples of ab initio molecular dynamics simulations
- 11.14 Modelling solid-state defects
 - 11.14.1 Defect studies of the high-Tc superconductor YBa₂Cu₃O_{7-x}
- Appendix 11.1 Calculating Free Energy Differences Using Thermodynamic Integration
- Appendix 11.2 Using the Slow Growth Method for Calculating Free Energy Differences
- Appendix 11.3 Expansion of Zwanzig Expression for the Free Energy Difference for the Linear Response Method

12. THE USE OF MOLECULAR MODELLING AND CHEMOINFORMATICS TO DISCOVER AND DESIGN NEW MOLECULES

- 12.1 Molecular modelling in drug discovery
- 12.2 Computer representations of molecules, chemical databases and 2D substructure searching
- 12.3 3D database searching
- 12.4 Deriving and using three-dimensional pharmacophores
 - 12.4.1 Constrained systematic search
 - 12.4.2 Ensemble distance geometry, ensemble molecular dynamics and genetic algorithms
 - 12.4.3 Clique detection methods for finding pharmacophores
 - 12.4.4 Maximum likelihood method
 - 12.4.5 Incorporating additional geometric features into a 3D pharmacophore
- 12.5 Sources of data for 3D databases
- 12.6 Molecular docking
 - 12.6.1 Scoring functions for molecular docking
- 12.7 Applications of 3D database searching and docking
- 12.8 Molecular similarity and similarity searching

- 12.9 Molecular Descriptors
 - 12.9.1 Partition coefficients
 - 12.9.2 Molar refractivity
 - 12.9.3 Topological indices
 - 12.9.4 Pharmacophore keys
 - 12.9.5 Calculating the similarity
 - 12.9.6 Similarity based on 3D properties
- 12.10 Selecting "diverse" sets of compounds
 - 12.10.1 Data manipulation
 - 12.10.2 Selection of diverse sets using cluster analysis
 - 12.10.3 Dissimilarity-based selection methods
 - 12.10.4 Partition-based methods for compound selection
- 12.11 Structure-based de novo ligand design
 - 12.11.1 Locating favourable positions of molecular fragments within a binding site
 - 12.11.2 Connecting molecular fragments in a binding site
 - 12.11.3 Structure-based design methods to design HIV-1 protease inhibitors
 - 12.11.4 Structure-based design of templates for zeolite synthesis
- 12.12 Quantitative structure-activity relationships
 - 12.12.1 Selecting the compounds for a QSAR analysis
 - 12.12.2 Deriving the QSAR equation
 - 12.12.3 Cross-validation
 - 12.12.4 Interpreting a QSAR equation
 - 12.12.5 Alternatives to multiple linear regression: discriminant analysis, neural networks and classification methods
 - 12.12.6 Principal Components Regression
- 12.13 Partial least squares
 - 12.13.1 Partial least squares and molecular field analysis
- 12.14 Combinatorial libraries
 - 12.14.1 The design of "drug-like" libraries
 - 12.14.2 Library enumeration
 - 12.14.3 Combinatorial subset selection
 - 12.14.4 The future

INDEX