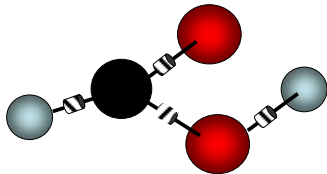


Drug-Receptor Binding

1. Computational Chemistry
 - Molecular Mechanics
 - Quantum Mechanics
 - Molecular Dynamics etc.
2. Docking
3. Force-field Based 3D-QSAR
4. Free-energy Perturbations

Molecular Mechanics I



- the molecule is represented by a set of atoms connected by springs
- bond lengths, angles, and dihedral angles in relaxed molecules have 'optimal' magnitudes (averages of those found in most molecules)

Molecular Mechanics III

- the overall energy of a system (one molecule or several interacting molecules) is calculated as the sum of all contributions
- the energy is a function of bond lengths, angles, torsion angles, and distances between atoms
- this function can be minimized - corresponding bond lengths, angles, torsion angles, and distances characterize optimal geometry of the system

Computational Chemistry Methods

Approximately, in the order of increasing precision and cost:

- classical mechanics, a.k.a. molecular mechanics, a.k.a. force fields
- quantum mechanics
 - semi-empirical models
 - Hartree-Fock models
 - density functional models
 - Moeller-Plesset models

Molecular Mechanics II

- any deviation from the equilibrium state increases the energy according to simple functions
 - for bond lengths $E_b = K_b (r - r_0)^2 / 2$
 - for bond angles $E_\theta = K_\theta (\theta - \theta_0)^2 / 2$
 - for dihedrals $E_\phi = K_\phi (1 + \cos(n\phi - \delta))^2 / 2$
 - in aromatic rings $E_\xi = K_\xi (\xi - \xi_0)^2 / 2$

- electrostatic and van der Waals interactions

$$E_{ES} = \frac{C_{12}}{r_{ij}^{12}} - \frac{C_6}{r_{ij}^6} + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}$$

Lennard-Jones potential Coulomb law

Molecular Mechanics IV

- force field – form of the functions and values of the parameters
- many force fields available, developed for
 - organic molecules (MM2, MM3)
 - proteins (CHARMM, OPLS, Amber)
 - ligand-protein interactions (MMFF4)
- specific developments
 - polarizability
 - directional hydrogen bonds (Vedani...)
 - organometallic complexes (SIBFA...)

Quantum Mechanics I

- interactions between electrons and nuclei are described
- molecular geometry in terms of minimum energy arrangements of nuclei
- background – Schrödinger equation
- exact solution – available just for hydrogen – defines wavefunctions (s, p, d... orbitals)
- the square of the wavefunction defines electron density – measured in x-ray experiments

Hartree-Fock Methods (ab initio)

- the approximations result in Roothaan-Hall equations
- solved iteratively until self-consistency
- problem – the electrons are treated as independent; their movement causes more repulsion than is actually present
- electron correlation – coupling of movements of electrons
- methods have been developed to account for electron correlation – additional cost

Density Functional Models I

- based on the Hohenberg-Kohn theorem:
“The minimal energy of a collection of electrons under the influence of an external (Coulombic) field is a unique ‘functional’ (a function of a function) of the electron density.”
- the energy includes many of the same components as the Hartree-Fock energy, but provides explicit account of electron correlation in the form from the exact (numerical) solution of a many-electron gas of uniform density

Quantum Mechanics II

To solve Schrödinger equation (SE) for systems that are more complex than hydrogen, approximations are needed:

- Born-Oppenheimer approximation – electron moves much faster than nuclei; nuclei fixed and SE only for electrons
- Hartree-Fock approximation – multi-electron wavefunction expressed as the product of single-electron wavefunctions
- LCAO approximation – molecular orbitals are expressed as a linear combinations of atomic orbitals (prescribed basis functions)

Hartree-Fock Methods: Semi-empirical

Hartree-Fock models with approximations:

- consider valence electrons only
- the basis set reduced to minimal representation
- parameterizations are based on reproducing a wide variety of experimental data, including
 - equilibrium geometries
 - heats of formation
 - dipole moments
 - ionization potentials
- frequently used models AM1 and PM3 incorporate essentially the same approximations but differ in their parameterization

Density Functional Models II

- different density functional models available, the names are formed of the last name initials of authors
- the simplest method
 - SVWN (Slater, Vosko, Wilk, Nusair)
- other methods
 - BP (Becke, Perdew)
 - BLYP (Becke, Lee, Yang, Parr)
 - B3LYP (the same authors)
- for similar cost as Hartree-Fock methods, they provide better descriptions

Moeller-Plesset Models

- account for many-electron effects
- based on a perturbation expansion of the energy
 - the first level: Hartree-Fock energy
 - the second level: MP2
 - higher level MP3, MP4 are impractical
- most expensive method

Selecting a Method

task	MM	Semi-emp.	Hartree-Fock	Density Functional	LMP 2
geometry (organic)	F-G	G	G	G	G
geometry (trans. metal)	P	F	F	G	F-G
transition-state geom.	N/A	F-G	G	G	G
conformation	G	P	F-G	G	G
thermochemistry	N/A	P	F-G	G	G
cost	VL	L	M	M	H

Rating: **G** – good, **F** – fair, **P** – poor

Cost: VL – very low, L – low, M – moderate, H – high

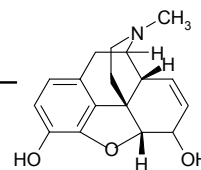
Computational Time II

Group	Method/Basis Set	Energy	Geometry
MM	MMFF	(a)	(a)
semi-empirical	AM1	(a)	0.01
Hartree-Fock	HF/3-21G	1(b)	17
	HF/6-31G*	2.5	31
	HF/6-31G**	2.9	35
density functional	BLYP/6-31G*	2.7	32
	BLYP/6-31G**	3.0	35
	B3LYP/6-31G*	3.8	42
	B3LYP/6-31G**	4.3	53
Moeller-Plesset	LMP2/6-31G*	16	-
	LMP2/6-31G**	15	-

Basis Sets

- functions describing molecular orbitals in
 - Hartree-Fock,
 - Density Functional, and
 - Moller-Plesset models
- Gaussian-type functions used most frequently
 - a polynomial in the Cartesian coordinates (x,y,z) followed by an exponential in r^2
- the coefficients determined by the fit to exponential Slater-type orbitals (STO)

Computational Time I



- used molecule – morphine
- computational demands depend on
 - the method
 - basis set
- times are relative to the Hartree-Fock method with the 3-21G basis set
- symbols
 - (a) too short to measure
 - (b) standard

Molecular Dynamics I

- simulates dynamics of a molecular system of known composition by solving the Newton equation

$$F = ma \quad a = dv / dt = d^2r / dt^2$$

- to calculate the forces F , an energy function is needed
- essentially, any method can be used depending upon the size and expected changes of the system and available computational resources
- in drug design, force fields are used as energy functions because the analyzed systems are large (drug-receptor complexes, bilayers)

Molecular Dynamics II

- the set of equations is integrated numerically

$$v(t_{n+1/2}) = v(t_{n-1/2}) + \frac{F(t_n)}{m} \Delta t$$

$$r(t_{n+1}) = r(t_n) + v(t_{n+1/2}) \Delta t$$

- initial velocities are assigned to reflect temperature of the system
- the time steps must be rather short – picoseconds
- many cycles needed to simulate 1-nanosecond event

Monte Carlo Simulation

- this method does not calculate forces to determine the motion of the system
- instead, the motion is generated by random jumps (conformations are crossing the barriers without feeling them)
- only the overall energy is calculated
- no time-dependent quantities can be derived, just equilibrium (thermodynamic) properties
- the state of the system is based on Boltzmann distribution of energies

Drug Design for Known Receptors

- receptor structure can be obtained
 - experimentally (x-ray, NMR)
 - from sequence, by homology modeling or threading
- for a receptor of known structure, the ligand binding can be examined by a variety of methods differing in quality and speed
 - docking - approximate and fast
 - de novo design - intermediate
 - force-field based 3D-QSAR - intermediate
 - binding free energy calculations – precise & slow
 - free energy perturbation – very slow, LR - faster

Langevin Dynamics

- based on a stochastic differential equation

$$F = ma - \gamma v + R$$

- two terms were added to the Newton equation
 - friction term (proportional to velocity)
 - random forces (kicks of the solvent molecules if they are not explicitly represented)
- the time steps can be longer and the system coarser than in MD

Estimating Ligand-Receptor Interactions

information about the receptor		
known		unknown structure, no template
	sequence+template	
structure	comparative models	receptor-site models
molecular simulations MD, MC free energy perturbation, LR		
docking force-field based 3D-QSAR		

Docking I

- in silico* screening tool meant to provide rapid selection of structures that will bind to a receptor
- two interrelated aspects
 - docking the structure into the receptor cavity
 - predicting the binding energy using a scoring function
- the first method – DOCK (Kuntz, 1982)
 - both receptor and ligand are treated as rigid

Docking II

- flexible ligand in rigid binding site – the problem becomes combinatorially demanding
- various techniques used to sample the vast space of possible solutions (FRED is exhaustive)
 - fast shape matching (DOCK, Eudock)
 - incremental construction (FlexX, Hammerhead)
 - TABU search (ProLead, SFDock)
 - simulated annealing (AutoDock 2.4)
 - genetic algorithms (GOLD, Gambler)
 - Lamarckian genetic algorithms (AutoDock 3.0)
 - Monte Carlo simulations (MCDock, Dockvision, QXP)
 - distance geometry (Dockit)

de novo Design

- instead of taking pre-generated structures and evaluate them for binding, a new structure can be 'grown' inside the binding site
- approaches are similar
 - LUDI (Bohm, Accelrys)
 - BUILDER (UCSF, Kuntz)
 - CombiBUILD (UCSF, Kuntz)
 - SMOG (Small Molecular Growth)
 - deWitte, Harvard
 - SPROUT (SimBioSys)

Ensemble-Based Approaches

- **Free-energy Perturbation Approach**
 - the most precise and most expensive approach to calculate the binding energy
 - based on thermodynamic cycles and gradual morphing of known ligand into a new ligand
- **Partitioning methods** - to become practical, several approximations were introduced
 - the overall energy can be written as the sum of electrostatic, van der Waals, and desolvation contributions (characterized through SASA – solvent-accessible surface area)
 - ensemble averages describe the situation well
 - the methods are under development

Docking III

Future directions:

- flexible receptors so that induced fit can be incorporated
- solvent effects
- improvement in scoring functions
- coordination interactions with metals for metalloproteins (incorporated in the latest version of FlexX)

Force-field Based 3D-QSAR

- the drug-receptor interaction energy is calculated using appropriate force field
- the receptor is dissected in fragments (amino acids or larger fragments)
- the binding energies for individual fragments are weighted using MLR or PLS to account for the effects like variations in dielectric constants...
- two approaches published
 - FF-3D-QSAR (Anton Hopfinger, Chicago)
 - COMBINE (Rebecca Wade, Heidelberg)

Linear Response Methods

- LR – one of the most frequently used (software Liaison)
- Authors Aquist, Jorgensen (1994-1995)
$$\Delta G_b = \alpha \times \Delta \langle E_{vdW} \rangle + \beta \times \Delta \langle E_{el} \rangle + \gamma \times \Delta \langle SASA \rangle + \kappa$$
 - Δ difference between bound and free ligands
 - $\langle \rangle$ denotes ensemble averages of energies obtained by force-field-based MD simulation
- not suitable for bonds that are not described well by force fields (e.g. coordination bonds)
 - the two leftmost terms can be replaced by QM energy of the time-averaged structure