

**Protein/Ligand Docking**

Introduction to Structural Bioinformatics

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (1)

---

---

---

---

---

---

---

---

**Docking Categories**

- ❑ Protein/Protein Docking
- ❑ Protein/Ligand Docking
- ❑ Protein/{DNA,RNA} Docking
  - ❑ Quite difficult due to flexibility of DNA/RNA
  - ❑ Little work available!

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (2)

---

---

---

---

---

---

---

---

**Protein/Protein Docking**

- ❑ Challenge: start with unbound coordinates for components and compute the coordinates for the complex
- ❑ Often used starting point: individual parts from complex
- ❑ Induced fit!

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (3)

---

---

---

---

---

---

---

---

**Protein/Protein Docking [2]**

- ❑ Most approaches use 'rigid-body' assumption
- ❑ Need:
  - ❑ geometric surface model
  - ❑ heuristic cost function
  - ❑ Geometric/combinatorial matching algorithm

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (4)

---

---

---

---

---

---

---

---

**Geometric Models**

- ❑ Want sparse but information rich surface representation
- ❑ Examples:
  - ❑ Connolly surface derived
  - ❑ Grid-based representations
  - ❑ Spherical harmonics

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (5)

---

---

---

---

---

---

---

---

**Heuristic Cost Function**

- ❑ Usually some form of geometric fit
- ❑ Highly insufficient
- ❑ Chemical interaction terms:
  - ❑ H-bonds, salt bridges
  - ❑ Lipophilic interactions
  - ❑ Solvation free energy
- ❑ Errors often same order of magnitude as individual energy contributions

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (6)

---

---

---

---

---

---

---

---

### Protein/Ligand Docking

- ❑ Main difference: flexibility of small molecule taken into account
- ❑ Used in database screening, rational drug design...
- ❑ Complementary surfaces are much smaller and less discriminating
- ❑ Need to model flexibility and weak interactions!

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (7)

---

---

---

---

---

---

---

---

### Algorithms

- ❑ Widely used: DOCK
  - ❑ Fill active/binding site with cluster of overlapping spheres
  - ❑ Match sphere centers of cluster with similar clusters representing the ligand
  - ❑ Rank predicted complexes by contact score or potential energy

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (8)

---

---

---

---

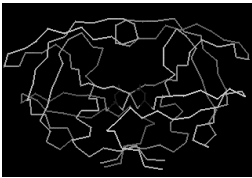
---

---

---

---

### DOCK Algorithm



Start with crystal coordinates (e.g. HIV-1 Protease)  
Example uses protein/ligand illustration!

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (9)

---

---

---

---

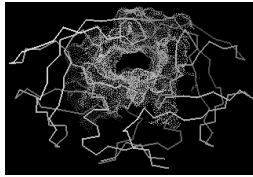
---

---

---

---

### DOCK Algorithm [2]



Generate Molecular Surface for receptor (active site only)

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (10)

---

---

---

---

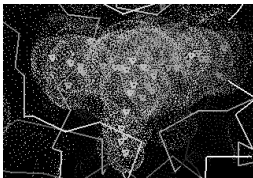
---

---

---

---

### DOCK Algorithm [3]



- Generate spheres to fill the active site
- Centers of spheres become location for potential locations for ligand atoms

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (11)

---

---

---

---

---

---

---

---

### DOCK Algorithm [4]

- Matching
  - sphere centers are matched to ligand atoms
  - Typically in the order of 10,000 different orientations for each ligand are analysed
- Scoring

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (12)

---

---

---

---

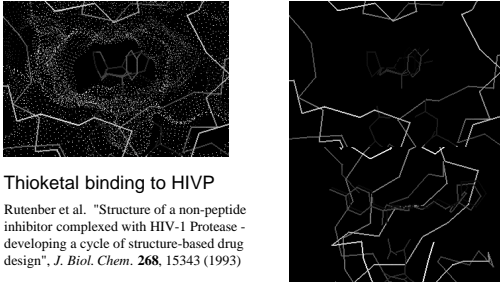
---

---

---

---

### DOCK Example Result



**Thiokeetal binding to HIVP**  
Rutenber et al. "Structure of a non-peptide inhibitor complexed with HIV-1 Protease - developing a cycle of structure-based drug design", *J. Biol. Chem.* **268**, 15343 (1993)

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (13)

---

---

---

---

---

---

---

---

### Others Algorithms: FlexX

- <http://cartan.gmd.de/flexx/>
- Uses incremental docking algorithm
- Model:
  - Torsional angles w/ 1-12 settings
  - Ring conformations w/ 1-6 settings
  - H-bonds, salt bridges
  - Lipophilic interactions
  - Discretization of interaction surfaces
- Online server available!!

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (14)

---

---

---

---

---

---

---

---

### Yet Other Algorithms

- FTDOCK
  - <http://www.icnet.uk/bmm/software.html>
  - Implement algorithm by Fourier correlation [Katschalski-Katzir et al., PNAS 89: 2195 (1992)]
  - Uses fast fourier transformations

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (15)

---

---

---

---

---

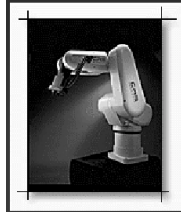
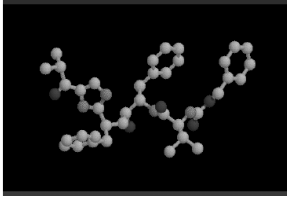
---

---

---

### Yet Other Algorithms [2]

- Brutlag/Singh: robotic path planning



Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (16)

---

---

---

---

---

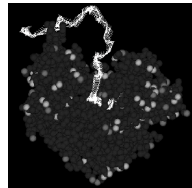
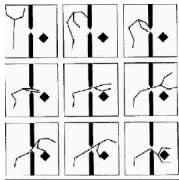
---

---

---

### Yet Other Algorithms [3]

- Brutlag/Singh: robotic path planning [2]
- "articulated robot" v. flexible ligand



Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (17)

---

---

---

---

---

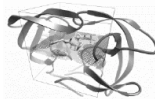
---

---

---

### Yet Other Algorithms [4]

- Olson: AutoDock
  - [www.scripps.edu/pub/olson-web/doc/autodock/](http://www.scripps.edu/pub/olson-web/doc/autodock/)
  - Ligand is docked to a set of grids
  - Free-energy scoring function based on a linear regression analysis, the AMBER force field, and a large set of diverse, known inhibition constants



Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (18)

---

---

---

---

---

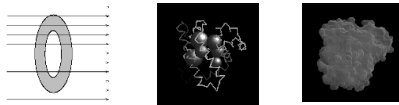
---

---

---

### Yet Other Algorithms [5]

- Prisant: Ray casting
  - <http://www.chem.duke.edu/research/prisant/protein/protein.html>



Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (19)

---

---

---

---

---

---

---

---

### Ligand Database: ReliBase

- Database of ligands and small molecules
- Query via Text, Sequence, SMILES, 2D/3D diagrams
- 3D visualization applet hooks into Rasmol
- URL: <http://relibase.ebi.ac.uk/>

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (20)

---

---

---

---

---

---

---

---

### Molecular Mechanics & Dynamics

Introduction to Structural Bioinformatics

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (21)

---

---

---

---

---

---

---

---

**Outline**

- Applications of Computed Structure
- Molecular Mechanics
  - Equations
  - Minimization
  - Limitations
- Molecular Dynamics
  - Algorithms
  - Applications
  - Free energy simulations
- MD Examples:
  - MD simulation of coiled coils
  - Relative binding constants for HIV protease inhibitors

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (22)

---

---

---

---

---

---

---

---

**Why Compute Structure?**

- HAVE: information about composition  
WANT: two (connectivity) or three dimensional structure
- HAVE: structure and physico-chemical properties  
WANT: rationalization of their relationship
- HAVE: information about the relationship  
WANT: predict new structures/properties

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (23)

---

---

---

---

---

---

---

---

**Central Postulate**

Structure  $\equiv$  Energy

- Any geometrical system has an associated Energy determined by its coordinates
- A fundamental description includes a multidimensional potential energy surface
- Minima on this surface correspond to stable configurations

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (24)

---

---

---

---

---

---

---

---

**Computed Structure**

- ❑ General Problem: Locate the energy minima starting from arbitrarily chosen configuration
- ❑ Need: mathematical description of structure/energy relationship
- ❑ Structure: Cartesian or internal coordinates

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (25)

---

---

---

---

---

---

---

---

**Outline**

- ❑ Applications of Computed Structure ✓
- ❑ Molecular Mechanics
  - ❑ Equations
  - ❑ Minimization
  - ❑ Limitations
- ❑ Molecular Dynamics
  - ❑ Algorithms
  - ❑ Applications
  - ❑ Free energy simulations
- ❑ MD Examples:
  - ❑ MD simulation of coiled coils
  - ❑ Relative binding constants for HIV protease inhibitors

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (26)

---

---

---

---

---

---

---

---

**Molecular Mechanics**

- ❑ Definition: A mathematical model to predict accurate structures and energies of molecules
- ❑ Basic Concept: optimize the geometry of a molecule (computed bonds and bond angles) to adopt “natural” meanings and values

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (27)

---

---

---

---

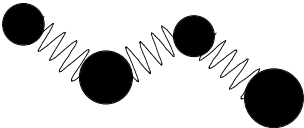
---

---

---

---

### MM View of Molecules



Molecules: masses joined by springs

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (28)

---

---

---

---

---

---

---

---

### MM: Simple Chemical View

- Molecules consist of atoms
- Atoms are bonded to each other according to the rules of valence (bonded terms)
- Molecules are flexible (torsional terms)
- Non-bonded atoms interact with each other (non-bonded terms)

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (29)

---

---

---

---

---

---

---

---

### MM Formula

$$E = \sum E_{STR} + \sum E_{BEND} + \sum E_{OOP} + \sum E_{TOR} + \sum E_{VDW} + \sum E_{ELE}$$

STR = stretching; BEND = bending; OOP = out of plane bending  
 TOR = torsional; VDW = van der Waals; ELE = electrostatic

The MM energy expression presents the relationship between Chemical structure and its energy.

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (30)

---

---

---

---

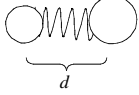
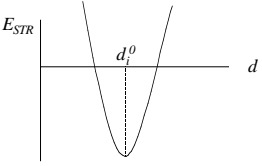
---

---

---

---

### Bond Stretching Energy

$$E_{STR} = \sum_{i=1}^n \frac{1}{2} k^{d_i} (d_i - d_i^0)^2$$

$d_i$  = length of  $i^{th}$  bond  
 $d_i^0$  = equilibrium length of  $i^{th}$  bond  
 $k^{d_i}$  = bond stretching force constant

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (31)

---

---

---

---

---

---


---

---

---

---

### Angle Bending Energy



$$E_{BEND} = \sum_{i=1}^n \frac{1}{2} k^{\Theta_i} (\Theta_i - \Theta_i^0)^2$$

$\Theta_i$  = angle between two adjacent bonds  
 $\Theta_i^0$  = equilibrium value for the  $i^{th}$  angle  
 $k^{\Theta_i}$  = angle bending force constant

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (32)

---

---

---

---

---

---

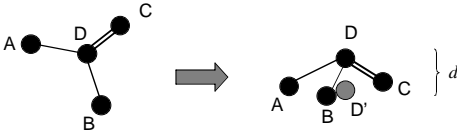
---

---

---

---

### Out of Plane Bending Energy



$$E_{OOP} = \sum_{i=1}^n \frac{1}{2} k_i^{OOP} d_i^2$$

$d_i$  = distance between center atom and plane  
 $k^{OOP}$  = out of plane bending constant

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (33)

---

---

---

---

---

---

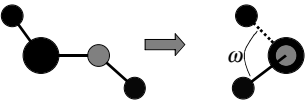
---

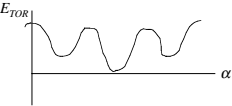
---

---

---

### Torsional Energy



$$E_{TOR} = \frac{V_1}{2}(1 + \cos \omega) + \frac{V_2}{2}(1 - 2 \cos \omega) + \frac{V_3}{2}(1 + 3 \cos \omega)$$


$V_1, V_2, V_3 = \text{torsional barriers}$   
 $\alpha = \text{torsion angle}$

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (34)

---

---

---

---

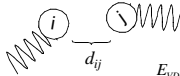
---

---

---

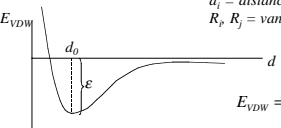
---

### van der Waals Energy



$$E_{VDW} = \sum_{i=1}^n \sum_{j=i+1}^n \epsilon \left( \left[ \frac{d_{ij}}{(R_i + R_j)} \right]^{-12} - 2 \left[ \frac{d_{ij}}{(R_i + R_j)} \right]^{-6} \right)$$

$\epsilon = \text{van der Waals constant}$   
 $d_{ij} = \text{distance between atoms } i \text{ and } j$   
 $R_i, R_j = \text{van der Waals radius of atoms } i \text{ and } j$



$$E_{VDW} = \epsilon \left[ \left( \frac{d_0}{d} \right)^{12} - 2 \left( \frac{d_0}{d} \right)^6 \right]$$

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (35)

---

---

---

---

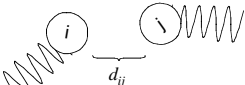
---

---

---

---

### Electrostatic Energy



$$E_{ELE} = \sum_{i=1}^n \sum_{j>i} \frac{Q_i Q_j}{4\pi\epsilon d_{ij}}$$

$\epsilon = \text{dielectric constant}$   
 $d_{ij} = \text{distance between atoms } i \text{ and } j$   
 $Q_i, Q_j = \text{net atomic charge at atoms } i \text{ and } j$

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (36)

---

---

---

---

---

---

---

---

**How to find a stable structure**

- Potential energy surface = multidimensional surface describing the energy of a molecule in terms of the nuclear positions
- Algorithm:
  - deform (twist, bend, stretch, pull...)
  - calculate energy changes
  - mathematically optimize (steepest descent, conjugate gradient...)
- Result: nearest (local) energy minimum

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (37)

---

---

---

---

---

---

---

---

**How to find a global minimum**

- Multi minimum problem: minimization only proceeds to lower energy and thus only finds minimum closest to starting point
- Need to start from many independent (random) locations
- Algorithms include Monte-Carlo simulations e.g.

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (38)

---

---

---

---

---

---

---

---

**MM Optimization**

$$E = \sum E_{STR} + \sum E_{BEND} + \sum E_{OOP} + \sum E_{TOR} + \sum E_{VDW} + \sum E_{ELE}$$

- Structure and Energy are equivalent
- Optimize Geometry & minimize Energy (don't mix them up!!)
- Task: find stable conformers starting from arbitrary geometry
- Molecular properties associated with ground state: heat of formation, vibrational spectra, shape (somewhat)

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (39)

---

---

---

---

---

---

---

---

### MM Optimization [2]

- ❑ Iterative geometry optimization always aims at lowering potential energy
- ❑ Minimum energy depends on starting geometry
- ❑ No known GENERAL method to find global minimum; only local minima closest to starting geometry are found

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (40)

---

---

---

---

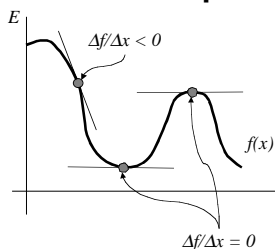
---

---

---

---

### MM Optimization [3]



Numerically, finding energy minima involves identification of point(s) on the potential energy surface where the first derivatives are zero

This is true for local maxima as well!!!

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (41)

---

---

---

---

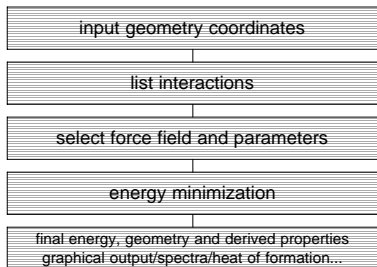
---

---

---

---

### MM Minimization Scheme



Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (42)

---

---

---

---

---

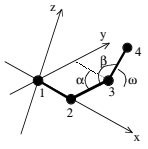
---

---

---

### Cartesian/Internal Coordinates?

☐ Need 3N coordinates for Cartesian v.  
3N-6 for internal:



Internal:

1. Atom: center of coord. System
2. Atom: on one of the axes  
-> one length
3. Atom: in the plane of that and 2nd axes  
-> one length & one angle ( $\alpha$ )
4. Atom: somewhere  
-> one length, one angle ( $\beta$ ) & one torsional angle ( $\omega$ )

Cartesian coordinates are easier w/ forcefield terms!

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (43)

---

---

---

---

---

---

---

---

### MM Limitations

☐ MM will not:

- ☐ Provide parameters
- ☐ Calculate "natural" MM values (bonds, angles, charges ...)
- ☐ Handle unstable structures (e.g. transition states)
- ☐ Simulate structure modifications (chemical reactions)
- ☐ Account for conformation dependent charges
- ☐ Generally account for effects of re-distribution of electrons around nuclei.

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (44)

---

---

---

---

---

---

---

---

### MM Cans & Cannots

- ☐ Describe energy as a function of coordinates (i.e. enthalpy)
- ☐ Allow atomic motion (minimization) but always downhill on the PE surface
- ☐ Describe conformations and conformers

- ☐ Describe energy as a function of time and temperature (i.e. entropy and free energy)
- ☐ Allow atoms move uphill the PE surface
- ☐ Describe conformational states (more realistic!)
- ☐ Include important parameters of "molecular life": temperature, atomic mass, time

**Related to molecular motion!!**

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (45)

---

---

---

---

---

---

---

---

**Outline**

- Applications of Computed Structure
- Molecular Mechanics
  - Equations
  - Minimization
  - Limitations
- Molecular Dynamics
  - Algorithms
  - Applications
  - Free energy simulations
- MD Examples:
  - MD simulation of coiled coils
  - Relative binding constants for HIV protease inhibitors

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (46)

---

---

---

---

---

---

---

---

**Molecular Dynamics**

- Definition: computational technique that studies the motions of atoms and molecules that occur due to their interactions
- Calculates fluctuations of atomic positions as a function of time
- Obtains macroscopic properties of a system from microscopic interactions
- "Snapshot": set of atomic coordinates at a certain moment during the simulation
- "Trajectory": set of consecutive snapshots

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (47)

---

---

---

---

---

---

---

---

**Motions in "real life"**

- Kinetic properties of liquids and solids (also cellular membranes!)
- Reaction mechanisms
- Solubilities
- Thermodynamic properties
- Binding
- Folding of proteins, polymers, nucleic acids
- MD: extension of classical MM modeling by application of Newtonian laws

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (48)

---

---

---

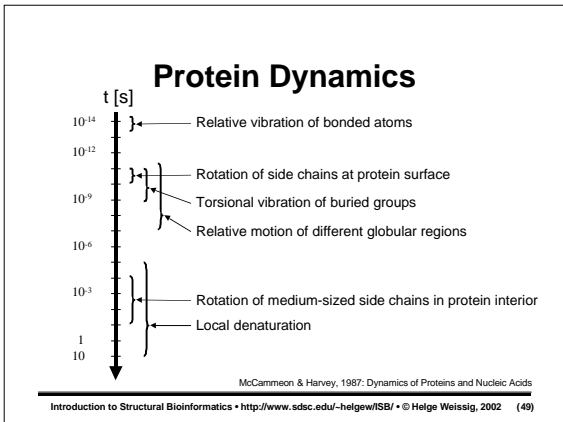
---

---

---

---

---




---

---

---

---

---

---

---

---

- ### MD Success Applications
- ❑ Structure: position, distance & orientation  
    -> NMR & X-ray data refinement!!!!
  - ❑ Mobility: B-factors, occupancy
  - ❑ Dynamics: vibrational frequencies, relaxation rates, diffusion
  - ❑ Thermodynamics: density, free energy, viscosity, conductance
- Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (50)

---

---

---

---

---

---

---

---

### Basic Equations

- ❑ New position of atom i:  $r_i(t + \Delta t) = r_i(t) + v_i \Delta t$
- ❑ Velocity of atom i:  $v_i(t) = v_i(t - \Delta t) + a_i \Delta t$
- ❑ Acceleration of atom i:  $a_i = \frac{F_i}{m_i}$
- ❑ Force at atom i:  $F_i = \partial / \partial r_i E(r_1 \dots r_n)$   
 $E(r_1 \dots r_n) = \text{MM potential energy}$

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (51)

---

---

---

---

---

---

---

---

### Temperature in MD

□ Kinetic energy:  $U_{kin} = \sum \frac{1}{2} m_i v_i^2 = \frac{3}{2} nkT$

*n* = number of atoms  
*k* = Boltzmann constant  
*T* = absolute temperature

- Temperature is maintained by scaling of velocities at each MD step
- Typically 300K

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (52)

---

---

---

---

---

---

---

---

### Temperature in MD [2]

- Simulated annealing runs at higher T (e.g. 1000K) to enhance conformational sampling
- Periodic cooling is used to settle into local minimum energy conformation
- Most effective way for conformational sampling of large biopolymer systems

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (53)

---

---

---

---

---

---

---

---

### MD Ensembles

- Constant Energy and volume (NVE)
- Constant temperature and volume (NVT)
- Constant temperature and pressure (NPT)

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (54)

---

---

---

---

---

---

---

---

### MD Programs

- ❑ AMBER: Weiner & Kollman, J. Comp. Chem. 2: 287 (1981)
- ❑ CHARMM Bråbøks et al., J. Comp. Chem. 4: 187 (1983)
- ❑ GROMOS: van Gunsteren et al., PNAS 80: 4315 (1983)
- ❑ DISCOVER: Hagler et al., J. Am. Chem. Soc. 96: 5319 (1974)
- ❑ CEDAR: Carson & Hermans, in "Molecular Dynamics and Protein Structure", ed. J. Hermans (1985)

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (55)

---

---

---

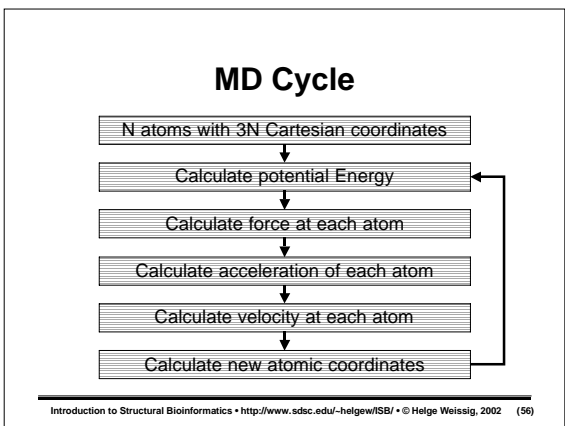
---

---

---

---

---



---

---

---

---

---

---

---

---

### MD Time Steps

- ❑ Rule of thumb: use 1/10th of highest frequency motion (i.e. bond vibration,  $10^{-14}$  s)
- ❑ Resulting time steps:  
$$\Delta t = 10^{-15} \text{ sec} = 1 \text{ fs} = 0.001 \text{ ps}$$
- ❑ Fixed bond lengths (SHAKE method) allows increase to 2 fs

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (57)

---

---

---

---

---

---

---

---

**Typical MD Protocol**

- Describe the system
- Describe MD features
- Equilibrate the system
- Run simulation
- Analysis

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (58)

---

---

---

---

---

---

---

---

**Typical MD Protocol [2]**

- Describe the system
  - Define atomic positions
  - Define system topology (connectivity)
  - Define potential function (parameter file)
  - Add missing atoms (H, hetero atoms...)
  - Optimize geometry
  - Add solvent, counter ions ...

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (59)

---

---

---

---

---

---

---

---

**Typical MD Protocol [3]**

- Describe MD features
  - Time step
  - # of time steps (length of simulation)
  - Frequency of non-bonded list update
  - Conditions (T, V/P)
  - Initial velocities
  - Special considerations (PBC...)

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (60)

---

---

---

---

---

---

---

---

**Typical MD Protocol [4]**

- Equilibrate the system
  - Equilibrate solvent with solute
  - Define # of heating and equilibration steps
  - Monitor system properties (PE, KE, T, P) during heating and equilibration for convergence

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (61)

---

---

---

---

---

---

---

---

**Typical MD Protocol [5]**

- Run simulation
  - Specify save frequency for coordinates and velocities
  - Monitor evolution of KE, PE, T, P
  - Monitor desired parameters (RMSD, important distances...)

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (62)

---

---

---

---

---

---

---

---

**Typical MD Protocol [6]**

- Analysis
  - Structure stability
  - Important molecular motions
  - Average structure
  - thermodynamics

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (63)

---

---

---

---

---

---

---

---

### MD limitations

- ❑ Very small time steps (1-2 fs)
- ❑ Total simulation time 300 - 500 ps
- ❑ Limited exploration of conformational space
- ❑ Accumulation of numerical errors during simulation run

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (64)

---

---

---

---

---

---

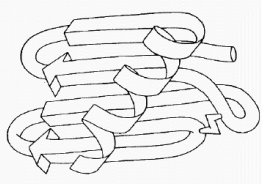
---

---

### MD limitations [2]

- ❑ In general: MD modeling is good at predicting unrealistic structures

Incredulase



Richardson & Richardson: "Some design principles: Betabellin" in "Protein Engineering" (Oxender & Fox eds., 1987)

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (65)

---

---

---

---

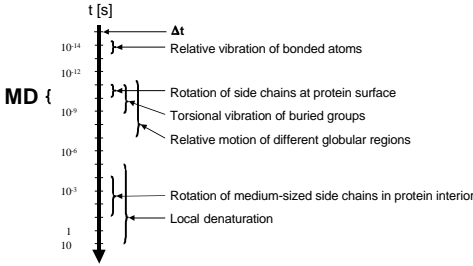
---

---

---

---

### MD in Protein Dynamics



Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (66)

---

---

---

---

---

---

---

---

**Outline**

- Applications of Computed Structure ✓
- Molecular Mechanics ✓
  - Equations
  - Minimization
  - Limitations
- Molecular Dynamics ✓
  - Algorithms
  - Applications
  - Free energy simulations
- MD Examples:
  - MD simulation of coiled coils
  - Relative binding constants for HIV protease inhibitors

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (67)

---

---

---

---

---

---

---

---

**Alpha Helices**

- Each of the first four and last four residues forms only one intra-helical bond
- Helices often begin and end with polar residues capable of making side chain/backbone H-bonds
- Residues separated by two or three other residues are located on the same side of a  $\alpha$ -helix (4-3 repeat)

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (68)

---

---

---

---

---

---

---

---

**Coiled Coils**

- Coiled Coils: dimers of orientational parallel and in-register  $\alpha$ -helices
- Primary structure of coiled coils contain characteristic 4-3 hydrophobic repeats  $(**h**h)_n$  ( $n > 4$  in naturally occurring coiled coils)
- Coiled coils are examples of simplest folded protein structure

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (69)

---

---

---

---

---

---

---

---



### HIV Protease Inhibitors

- Importance:
  - Biological and pharmaceutical significance
  - Abundance of structural information for both free enzyme and enzyme/inhibitor complexes
- Task: calculation of relative binding constants

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (73)

---

---

---

---

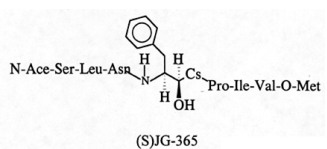
---

---

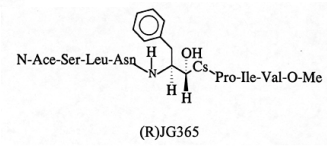
---

---

### Structure of HIVP Inhibitors



(S)JG-365



(R)JG365

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (74)

---

---

---

---

---

---

---

---

### Thermodynamic Cycle

$$\begin{array}{ccc}
 \text{HIVP} + (\text{S})\text{JG-365} & \xleftarrow{\Delta G^\circ_{12}} & \text{HIVP}/(\text{S})\text{JG-365} \\
 \uparrow \Delta G^\circ_{13} & & \uparrow \Delta G^\circ_{24} \\
 \text{HIVP} + (\text{R})\text{JG-365} & \xleftarrow{\Delta G^\circ_{34}} & \text{HIVP}/(\text{R})\text{JG-365}
 \end{array}$$

- For the entire cycle:
 
$$\Delta\Delta G^\circ = \Delta G^\circ_{12} + \Delta G^\circ_{24} - \Delta G^\circ_{34} - \Delta G^\circ_{13} = 0$$
- So:  $\Delta G^\circ_{12} - \Delta G^\circ_{34} = \Delta G^\circ_{13} - \Delta G^\circ_{24}$
- This difference (relative binding constants) is calculated with MD

---

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (75)

---

---

---

---

---

---

---

---

### Computer Alchemy

- In a course of MD simulations, gradually change one molecule ("wild" type) to another ("mutant" type)
- Calculate associated free energy difference between the "wild" type and "mutant" type states

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (76)

---

---

---

---

---

---

---

---

### Computer Alchemy [2]

- To calculate free energy, the potential energy of the system is made dependent on an artificial parameter  $\lambda$ .
- $\lambda$  is the relative contribution of one molecule and goes from 1 to 0 during the MD simulation:

Thus:  $E = E_{common} + \lambda E_{molecule 1} + (1 - \lambda) E_{molecule 2}$

- In practice  $\Delta\lambda \approx 1/n$  where  $n = \int_0^1 \frac{\partial E}{\partial \lambda} d\lambda \approx \sum_i \frac{\partial E}{\partial \lambda_i} \Delta\lambda_i$  is the total simulation time

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (77)

---

---

---

---

---

---

---

---

### MD Results

Theory	Experiment
□ Tropsha & Hermans (1992): -2.9 Kcal/mol	□ Rich et al. (1991): -2.6 Kcal/mol
□ Ferguson et al. (1991): -2.8 Kcal/mol	

Introduction to Structural Bioinformatics • <http://www.sdsc.edu/~helgew/ISB/> • © Helge Weissig, 2002 (78)

---

---

---

---

---

---

---

---