Introduction to Bioinformatics

Iosif Vaisman

Email: ivaisman@gmu.edu

Protein Modeling Methods

• *Ab initio methods*: solution of a protein folding problem
  search in conformational space

• *Energy-based methods*: energy minimization
  molecular simulation

• *Knowledge-based methods*: homology modeling
  fold recognition

Molecular structure representation

Potential Energy Function

\[
\text{PEF}(R) = \sum_{\text{bonds}} K_b (b(R)-b)^2 + \sum_{\text{angles}} K_o (\theta(R)-\theta)^2 + \\
\sum_{\text{dihedrals}} K_n (1 + \cos[n\theta(R) - \gamma]) + \\
\sum_{\text{non-bonded atom pairs } i,j} \left[ \frac{A_{ij}}{r_{ij}^n} - \frac{B_{ij}}{r_{ij}^p} + \frac{q_i q_j}{\epsilon_{ij} r_{ij}(R)} \right]
\]

Forcefields: AMBER, CHARMM, CVF, ECEPP, GROMOS

Non-Bonded Interactions

Bond length
Bond length

\[ E = \sum_{\text{angles} \theta} k_\theta (\theta - \theta_0)^2 \]

Bond angle

Torsional angle

Torsional angle (parameters)
Non-bonded terms

Energy Minimization

Molecular Dynamics

- Model system
- Initial conditions
- Boundary conditions
- Integration algorithm
- Constraints
- Ensemble
- Results

Non-bonded terms (parameters)

Energy Minimization

Molecular Dynamics

\[ F_i = m_i \ a_i \]
\[ a_i = \frac{d v_i}{d t} \]
\[ v_i = \frac{d r_i}{d t} \]
\[ - \frac{d E}{d r_i} = F_i \]
\[ - \frac{d E}{d r_i} = m_i \ 2 \frac{d^2 r_i}{d t^2} \]
Periodic Boundary Conditions

MD cycle and integration algorithm

1. Solve for $\dot{q}$ at $t$ using: $\dot{q} = F / m(q)$
2. Update $y_i$ at $t + \Delta t/2$ using: $y_i(t + \Delta t/2) = y_i(t - \Delta t/2) + \dot{y}_i(t) \Delta t$
3. Update $r_i$ at $t + \Delta t$ using: $r_i(t + \Delta t) = r_i(t) + \dot{r}_i(t + \Delta t/2) \Delta t$

Calculate potential energy
Calculate force at each atom
Calculate acceleration of each atom
Calculate velocity of each atom
Calculate new atomic coordinates

Motion Characteristic
- Relative vibration of bonded atoms: $10^{-14}$
- Rotation of side chains at protein surface: $10^{-11}$ to $10^{-10}$
- Torsional libration of buried groups: $10^{-11}$ to $10^{-9}$
- Relative motion of different globular regions: $10^{-11}$ to $10^{-7}$
- Rotation of medium-sized side chains in protein interior: $10^{-4}$ to 1
- Local denaturation: $10^{-5}$ to 10

Temperature in molecular dynamics

$$U_{\text{kin}} = \frac{1}{2} \sum m_i v_i^2 = \frac{3}{2} N k T$$

$N$ – number of atoms
$k$ – Boltzmann constant
$T$ – absolute temperature

MD Ensemble

Microcanonical $(N,V,E)$
Canonical $(N,V,T)$
Isothermal-isobaric $(N,P,T)$

MD of proteins: Solvent model

MD simulation of ubiquitin (400 ps)

Adapted from V. Daggett (1999)
MD of proteins: radial distribution functions

Water O-O rdf

Protein-water rdf

Adopted from V. Daggett (1999)

MD of proteins: mobile regions

Snapshots of V$_D$ domain simulation at 300 and 340 K

Adopted from W. F. van Gunsteren (2001)

MD of proteins: long runs

1 microsecond simulation of villin


MD: Reversible folding of peptides

Adopted from W. F. van Gunsteren (2001)