BINF 731

Protein Structure Analysis

Iosif Vaisman

2002

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 recogniion

Ab initio Methods

Simplified models
- simplified alphabet (HP)
- simplified representation (lattice)

Build-up techniques

Deterministic methods
- quantum mechanics
- diffusion equations
- DFT

Stochastic searches
- Monte Carlo
- genetic algorithms

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 recogniion

Molecular structure representation

Elementary particles

Atoms

Groups of atoms

Potential Energy Function

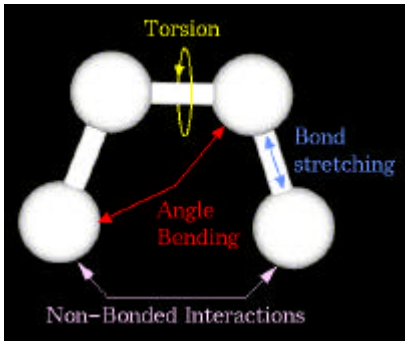
$$PEF(R) = \sum_{bonds} K_b\{b(R)-b_{eq}\}^2 + \sum_{angles} K_\theta\{\theta(R)-\theta_{eq}\}^2 +$$

$$\sum_{dihedrals} \frac{K_\phi}{2}\{1 + \cos[n\phi(R) - \gamma]\} +$$

$$\sum_{\substack{non\text{-}bonded\\ atom\ pairs\ i,j}} \left[\frac{A_{ij}}{r_{ij}(R)^{12}} - \frac{B_{ij}}{r_{ij}(R)^{6}} + \frac{q_i q_j}{\varepsilon_r \varepsilon_0 r_{ij}(R)} \right] \quad (1)$$

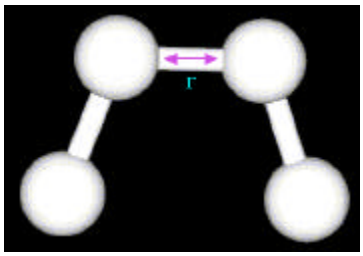
Forcefields: AMBER, CHARMM, CVF, ECEPP, GROMOS

Non-Bonded Interactions

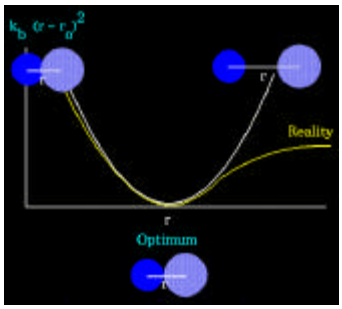


Bond length

$$E = \sum_{bonds} k_b (r - r_o)^2$$

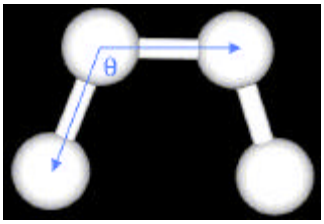


Bond length

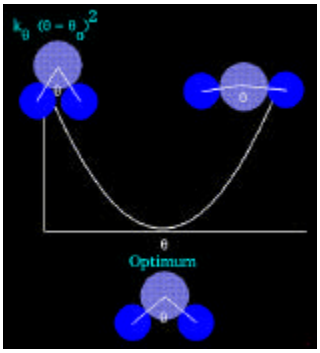


Bond angle

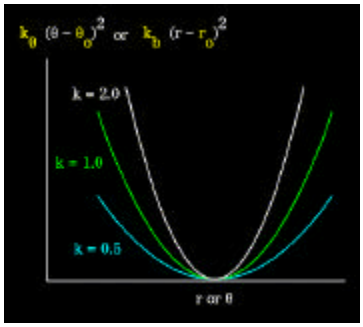
$$E = \sum_{angles\ \theta} k_\theta (\theta - \theta_o)^2$$



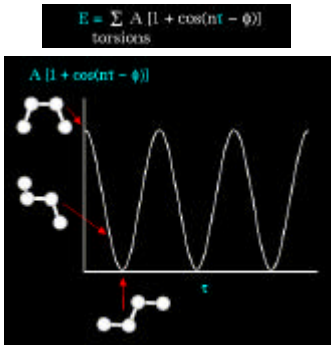
Bond angle



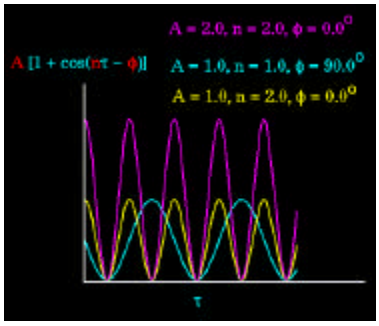
Bond length and angle (parameters)



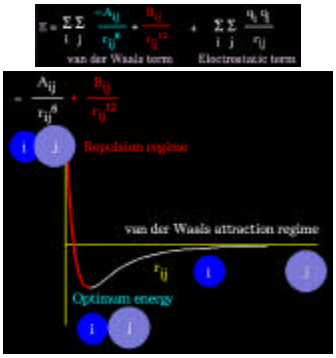
Torsional angle



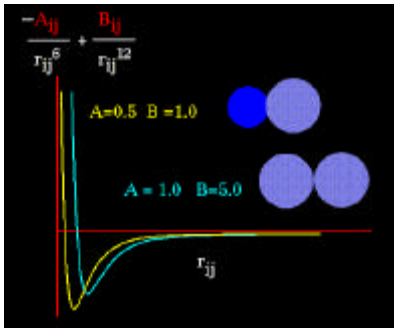
Torsional angle (parameters)



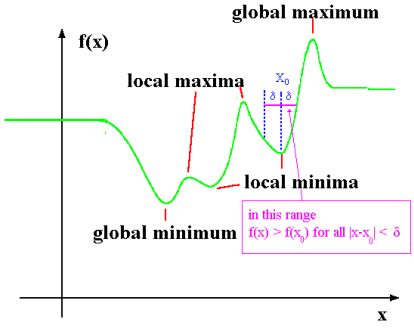
Non-bonded terms



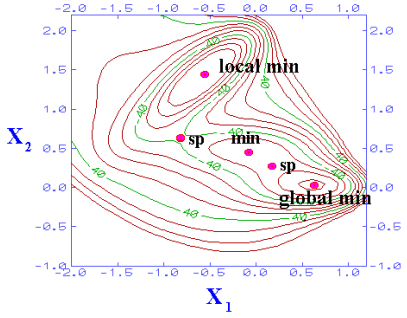
Non-bonded terms (parameters)



Energy Minimazation



Energy Minimazation



Molecular Dynamics

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

Molecular Dynamics

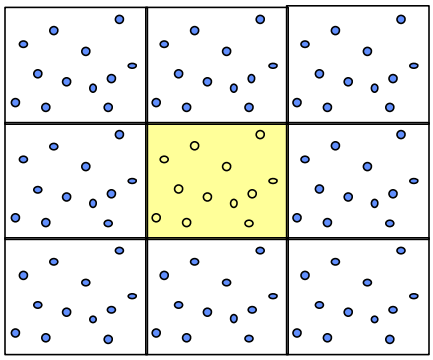
$$F_i = m_i a_i$$

$$a_i = dv_i / dt$$
$$v_i = dr_i / dt$$

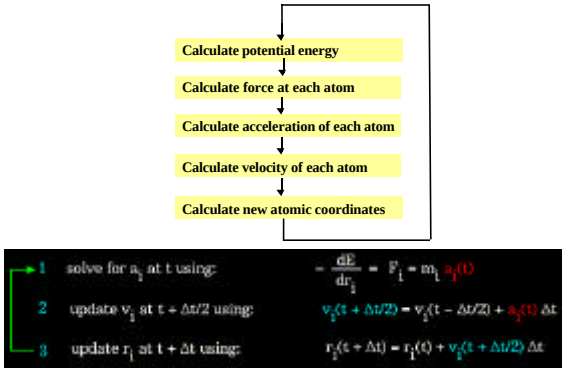
$$- dE / dr_i = F_i$$

$$- dE / dr_i = m_i \, d^2 r_i / dt^2$$

Periodic Boundary Conditions



MD cycle and integration algorithm



Time scales

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

MD Ensemble

Microcanonical (N,V,E)

Canonical (N,V,T)

Isothermal-isobaric (N,P,T)

Temperature in molecular dynamics

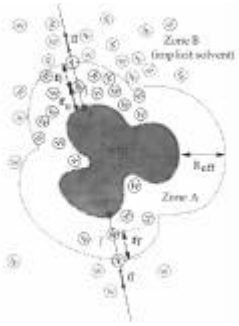
$$U_{kin} = \sum \frac{1}{2} m_i v_i^2 = \frac{3}{2} NkT$$

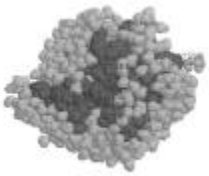
N – number of atoms

k – Boltzmann constant

T – absolute temperature

MD of proteins: Solvent model

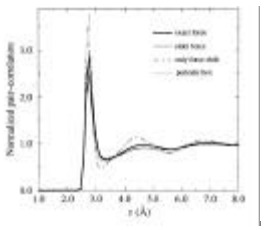




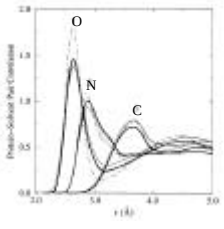
MD simulation of ubiquitin
(400 ps)

Adopted from V.Daggett (1999)

MD of proteins:
radial distribution functions



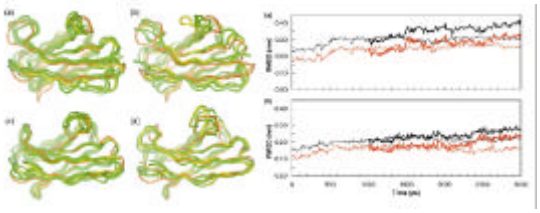
Water O-O rdf



Protein-water rdf

Adopted from V.Daggett (1999)

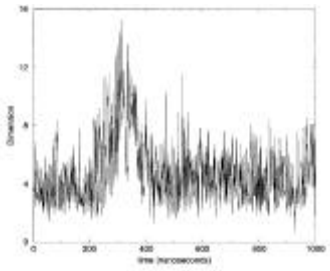
MD of proteins: mobile regions



Snapshots of V_H domain simulation
at 300 and 340 K

Adopted from W.F.VanGunsteren (2001)

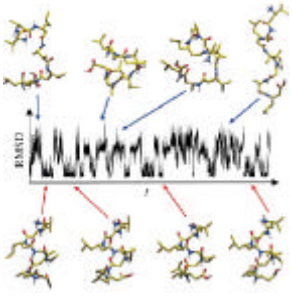
MD of proteins: long runs

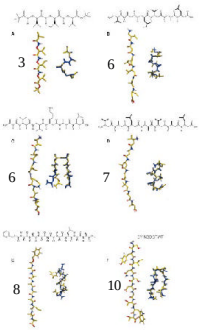


1 microsecond
simulation of
villin

Adopted from I.D.Kuntz and P.Kollman (2001)

MD: Reversible folding of peptides





Adopted from W.F.VanGunsteren (2001)