Bioinformatics Methods

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 recogniion

Ab initio Methods

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

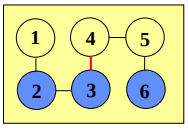
Build-up techniques

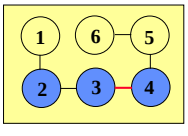
Deterministic methods
- quantum mechanics
- diffusion equations
- DFT

Stochastic searches
- Monte Carlo
- genetic algorithms

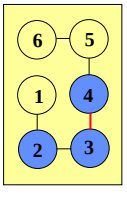
Genetic Algorithms Applications

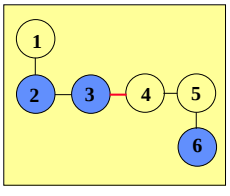
Parents
10 00 01 00 10
10 00 00 01 11

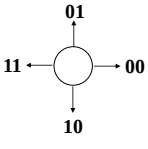




Children
10 00 10 01 11
10 00 01 00 10

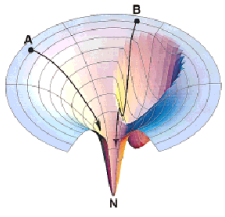






HP Lattice Models

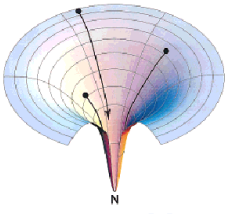
A



HP model

HH	ε < 0
HP	ε = 0
PP	ε = 0

B

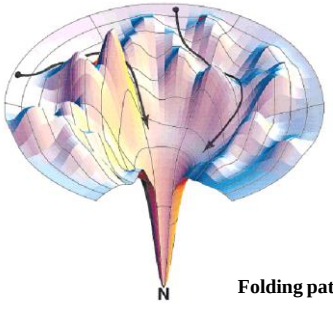


HP+ model

HH	ε < 0
HP	ε = 0
PP	ε = 0
NNC	ε > 0

Chan & Dill, 1998

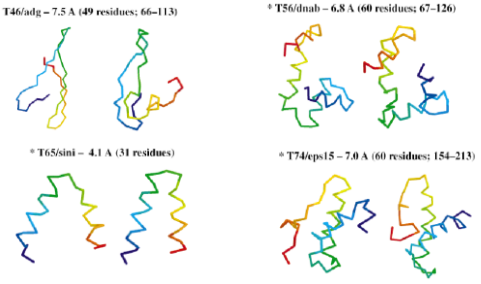
HP Lattice Models



Folding pathways

Chan & Dill, 1998

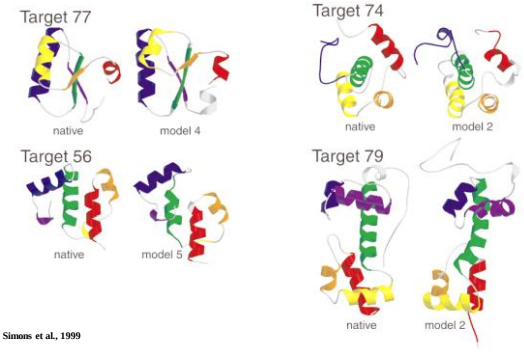
Hierarchical *ab initio* prediction



Lattice models
Knowledge-based scoring functions

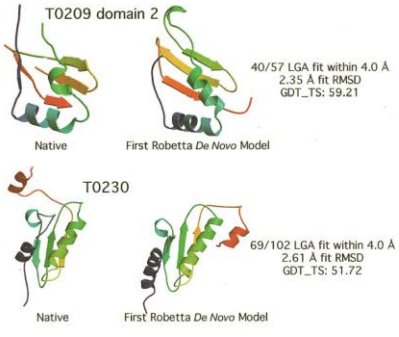
Samudrala et al., 1999

Ab initio prediction using Rosetta



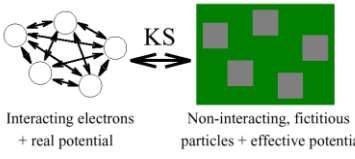
Simons et al., 1999

Ab initio prediction using Robetta

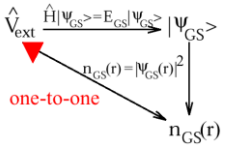


Chivian et al., 2005

Density Functional Theory

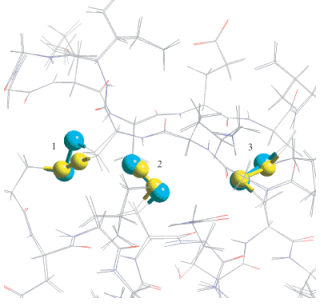


HK theorem: Each local one-particle potential corresponds to exactly one ground state density.



Adopted from Wilfried Aulbur, OSU

Density Functional Theory



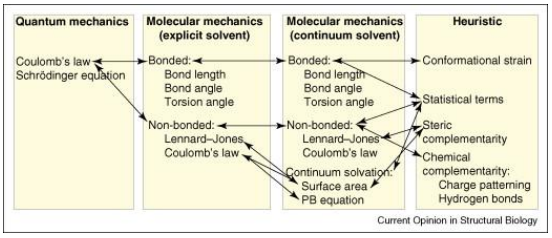
DFT optimization of NMR structure (1PNH)

Andreoni et al., 1999

Protein Modeling Methods

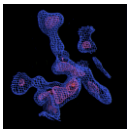
- ***Ab initio* methods:**
  solution of a protein folding problem
  search in conformational space
- **Energy-based methods:**
  energy minimization
  molecular simulation
- **Knowledge-based methods:**
  homology modeling
  fold recogniion

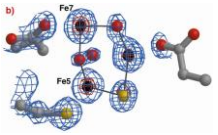
Potential Energy Functions



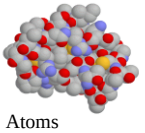
Boas & Harbury, 2007

Molecular structure representation

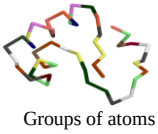




Elementary particles



Atoms



Groups of atoms

Potential Energy Function

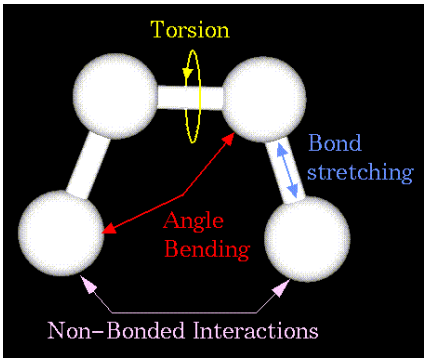
$$PEF(R) = \sum_{bonds} K_b\{b(R)-b_{eq}\}^2 + \sum_{angles} K_\theta\{\theta(R)-\theta_{eq}\}^2 +$$

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

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

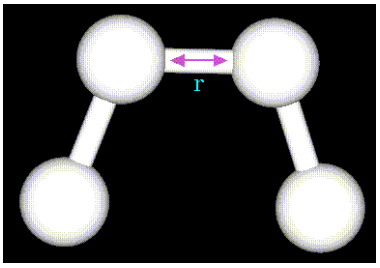
Forcefields: AMBER, CHARMM, CVF, ECEPP, GROMOS

Non-Bonded Interactions

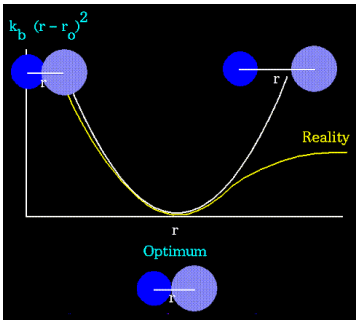


Bond length

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

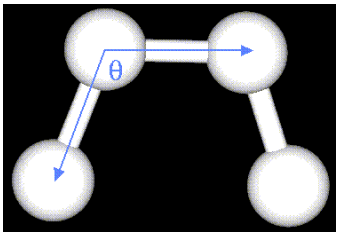


Bond length

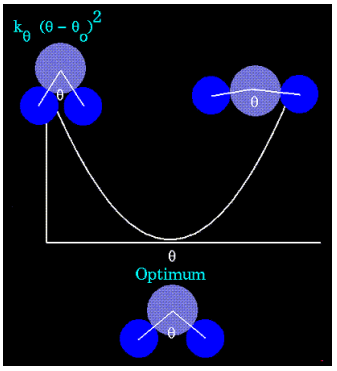


Bond angle

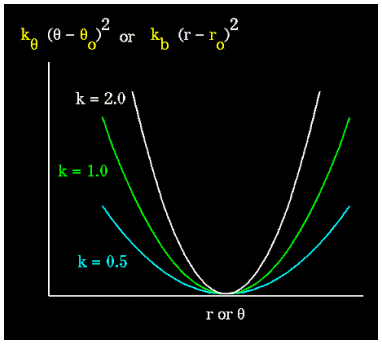
$$E = \sum_{\text{angles}} k_\theta \, (\theta - \theta_o)^2$$



Bond angle

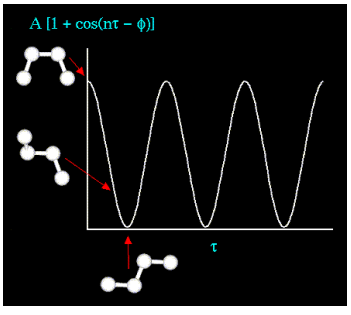


Bond length and angle (parameters)

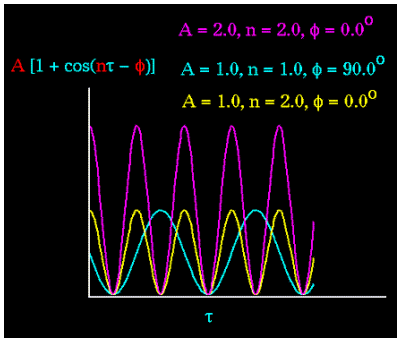


Torsional angle

$$E = \sum_{\text{torsions}} A \, [1 + \cos(n\tau - \phi)]$$



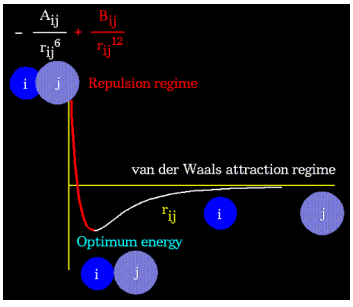
Torsional angle (parameters)



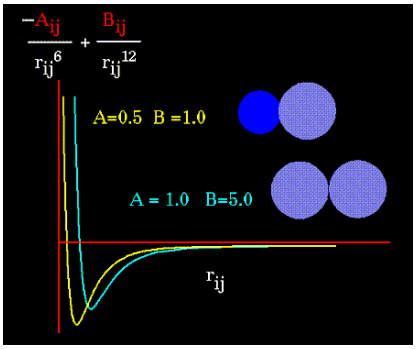
Non-bonded terms

$$E = \sum_i \sum_j \frac{-A_{ij}}{r_{ij}^6} + \frac{B_{ij}}{r_{ij}^{12}} + \sum_i \sum_j \frac{q_i \, q_j}{r_{ij}}$$

van der Waals term Electrostatic term



Non-bonded terms (parameters)



Potential Energy Function

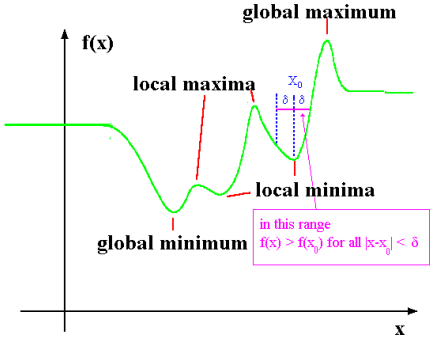
$$PEF(R) = \sum_{bonds} K_b\{b(R)-b_{eq}\}^2 + \sum_{angles} K_\theta\{\theta(R)-\theta_{eq}\}^2 +$$

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

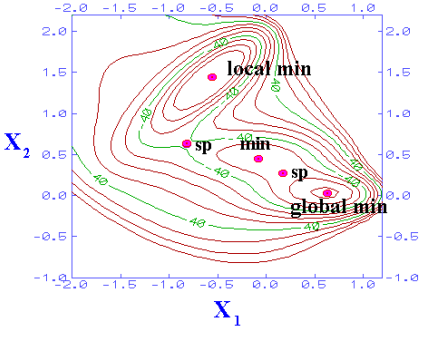
$$\sum_{\substack{non\text{-}bonded \\ atom\ pairs\ i,j}} \left[\frac{A_{ij}}{r_{ij}(R)^{12}} - \frac{B_{ij}}{r_{ij}(R)^6} + \frac{q_i q_j}{\varepsilon_r \varepsilon_0 r_{ij}(R)} \right] \quad (1)$$

Forcefields: AMBER, CHARMM, CVF, ECEPP, GROMOS

Energy Minimazation



Energy Minimazation



Molecular Dynamics

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

Molecular Dynamics

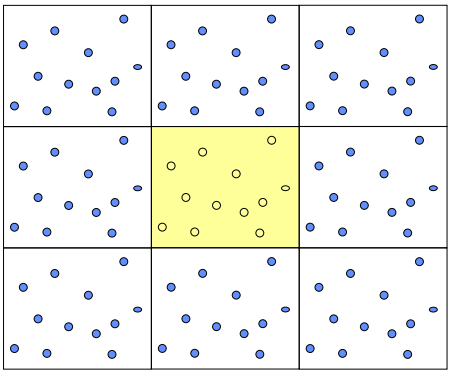
$$F_i = m_i a_i$$

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

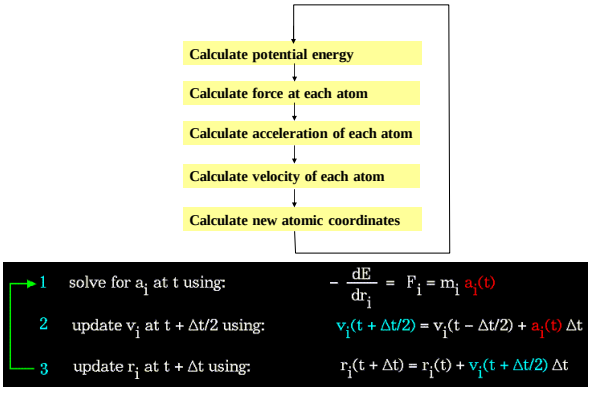
$$-dE / dr_i = F_i$$

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

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

Temperature in molecular dynamics

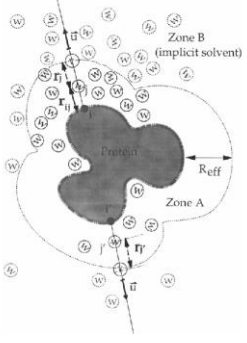
$$U_{kin} = \sum \frac{1}{2} m_i v_i^2 = \frac{3}{2} NkT$$

N – number of atoms

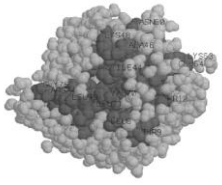
k – Boltzmann constant

T – absolute temperature

MD of proteins: Solvent model

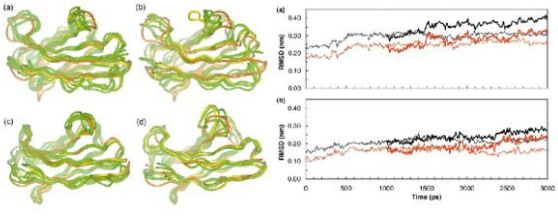


Adopted from V.Daggett (1999)



MD simulation of ubiquitin (400 ps)

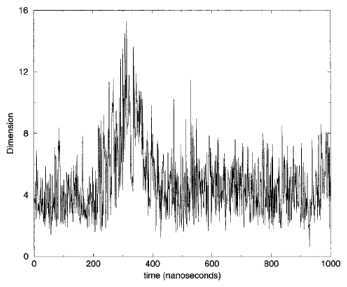
MD of proteins: mobile regions



Snapshots of V_H domain simulation at 300 and 340 K

Adopted from W.F.VanGunsteren (2001)

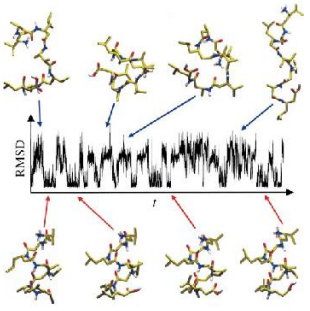
MD of proteins: long runs

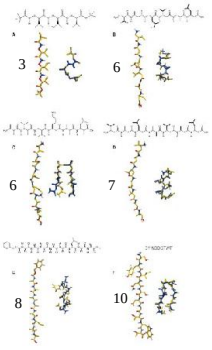


1 microsecond simulation of villin

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

MD: Reversible folding of peptides





Adopted from W.F.VanGunsteren (2001)