2015/9/1
2015年バイオスーパーコンピューティング・ソフト講習会シリーズ

「CafeMol」粗視化分子モデル計算ソフト講習会

検崎博生
理化学研究所情報基盤センター

CafeMol (www.cafemol.org)

- Features are;
  - Various CG models
    - protein/DNA/RNA
    - multiple basin model
    - accurate CG model
  - Simulating protein-at-work
    - “switching”
- Under development
  lipid

- Developer

検崎博生、古賀信康、藤原慎司、堀直人、金田亮、李 文飛、岡崎圭一、姚 新秋、寺川剛、伊藤真志保、高田彰二
京都大学高田研究室

CafeMol 2.1 (2013/7)
source & manual released

Overview of CafeMol

- General-purpose coarse-grained (CG) biomolecular modeling and simulation software
  - Protein: 1 bead / 1 amino acid
  - Nucleic acid:

    3 beads (sugar, base, phosphate) / nucleotide
  - Lipid: ~3 beads / lipid
- Written by FORTRAN90 with MPI and Open MP
- Version 1.0 is released (only protein) (2010/12/27)
  - Version 2.0 (protein, DNA, RNA) (2012/5/31)
  - Version 2.1 (protein, DNA, RNA) (2013/7/1)

Menu

1. Biomolecular simulation and CafeMol
2. Model and applications of CafeMol
3. Recent development of CafeMol

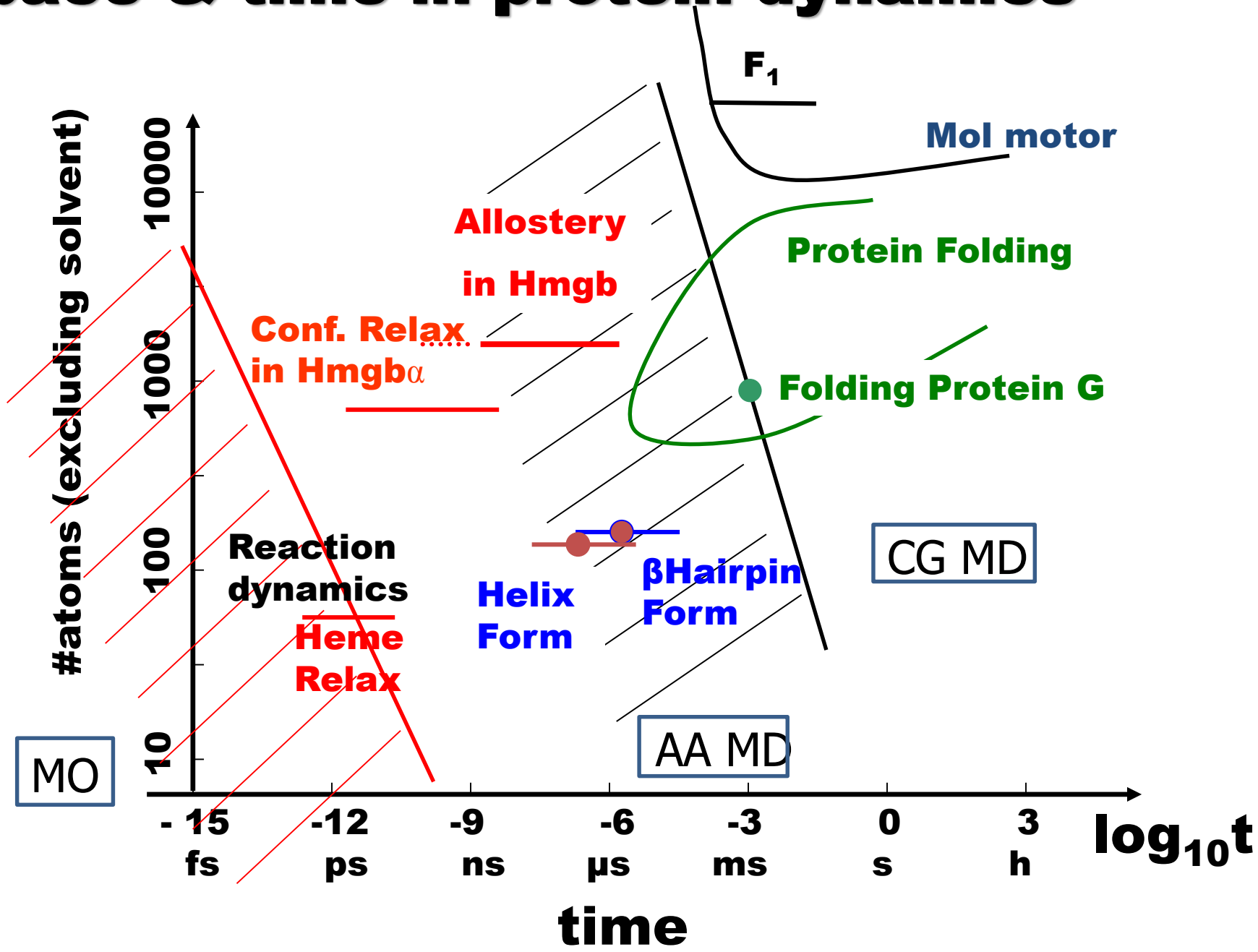
Menu

1. Biomolecular simulation and CafeMol
2. Model and applications of CafeMol
3. Recent development of CafeMol

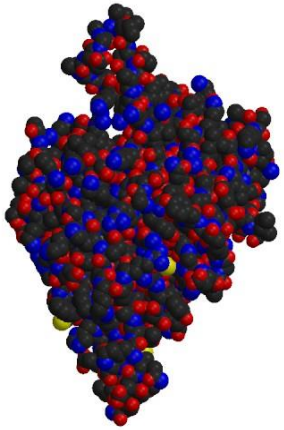
Biomolecular simulations

- Molecular orbital(MO) method
  - quantum mechanics
    - chemical reaction
- Molecular dynamics(MD) method
  - classical mechanics
    - conformational change of biomolecules
- Continuum method
  - elastic model
    - muscle contraction
  - fluid dynamics
    - blood flow

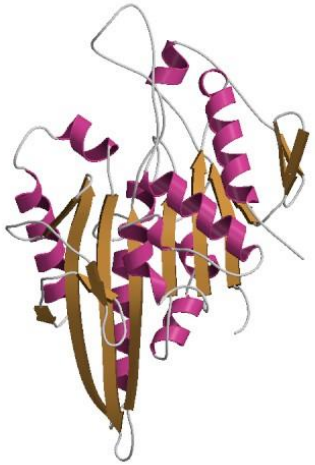
Space & time in protein dynamics



Why coarse-grained model ?



coarse-grained



Reducing the number of particles to 1/100

- 1 amino acid residue →1 particle ($C_\alpha$atom)
- No water molecule
  (included in potential energy)

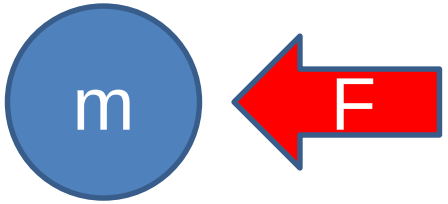
Protein folding and
functional conformational change

10^6 times speed up

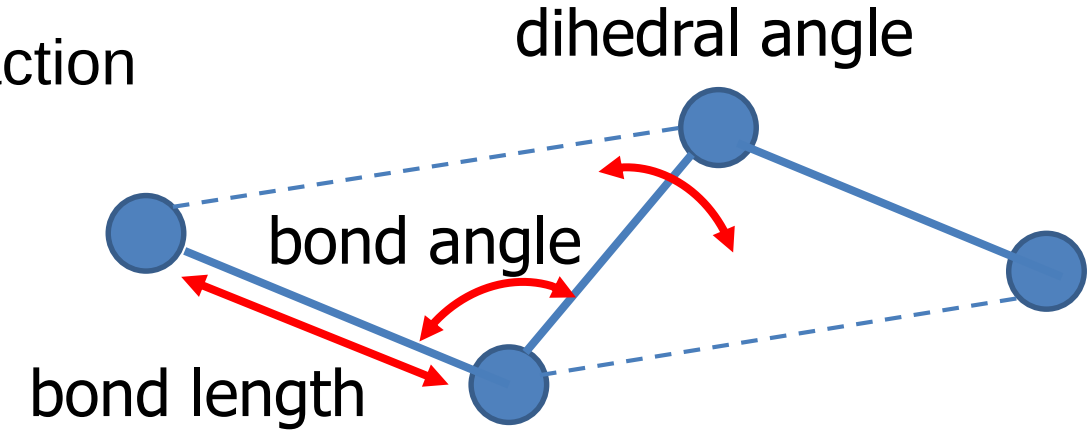
- Reducing the number of particles
- Enlarge time step
- Low friction

AA: 100PC 1week → CG: 1PC 1minute!!

Molecular dynamics



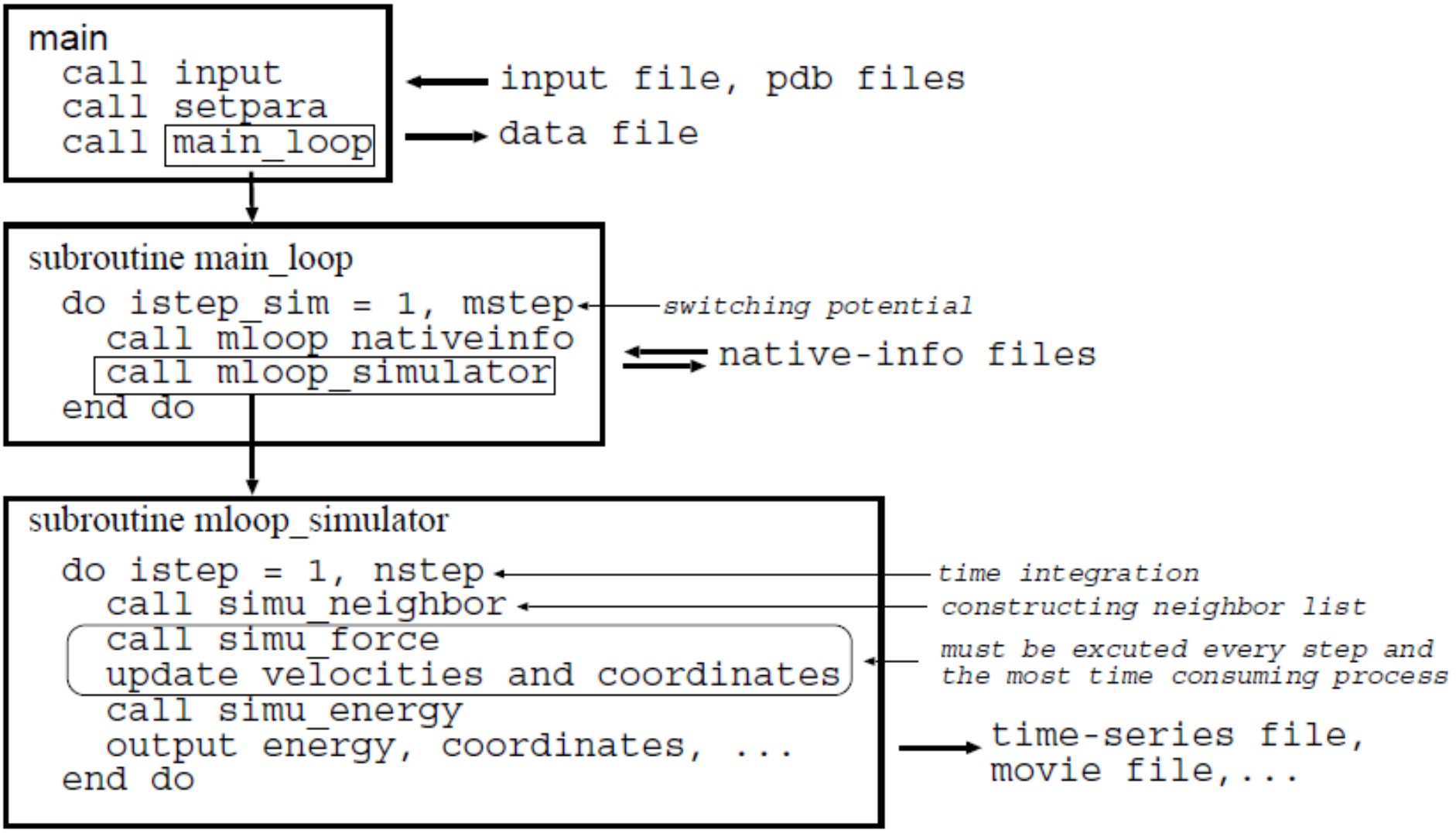
- Calculate molecular movement
  - Classical mechanics
  - Numerically integrate Newton's equation of motion
  - m(mass) x a(acceleration) = F(force)
  - interaction parameters are derived from MO (AA MD)
- Interaction
  - bond length, bond angle, dihedral angle
  - van der Waals interaction
  - electrostatic interaction



Simulation method

- Dynamics
  - Newtonian dynamics with Berendsen thermostat
  - Langevin dynamics
  - Multi-Particle Collision dynamics (MPC)
- Time integration
  - velocity Verlet algorithm
- Run mode
  - Constant temperature simulation
  - Simulated annealing
  - Auto-search of Tf
  - Replica exchange method
  - Potential “switching”
- Useful option
  - anchor, bridge, pulling, fix, box

CafeMol code



CafeMol code parallization

- Time integral(MPI+OpenMP)
  - neighboring list
    - each node calculates distances between assigned pairs and makes neighboring list
  - force, energy
    - each node calculates force or energy including their neighboring list
    - these calculations are parallelized by OpenMP
- Temperature/Hamiltonian REMD(MPI)
  - replica_1, replica_2, …. , replica_n

Units in CafeMol

- The length unit is A
  - 1A = 10^{-10}m
- The energy unit is kcal/mol
  - 1kcal/mol = 6.9478pNnm = 0.04337eV
  - $300k_BT$ = 0.6kcal/mol
- The mass unit is cafemol-mass-unit (cafemu)
  - each amino acid has the mass of 10cafemu
  - 1cafemu = 13.7amu = 2.275×10^{-26}kg
- The unit of charge is elementary-electric charge (e)
- The unit time is café time
  - 1cafe time= 1.809×10^{-13}s ~ 200fs
  - intrinsic dynamics is accelerated by coarse-graining the energy landscape

Input and output files

- Input files
  - Input file(.inp)
  - PDB file(.pdb)
  - native-info file(.ninfo)
  - parameter file(.para)
  - AICG files

red: essential files

green: frequently used

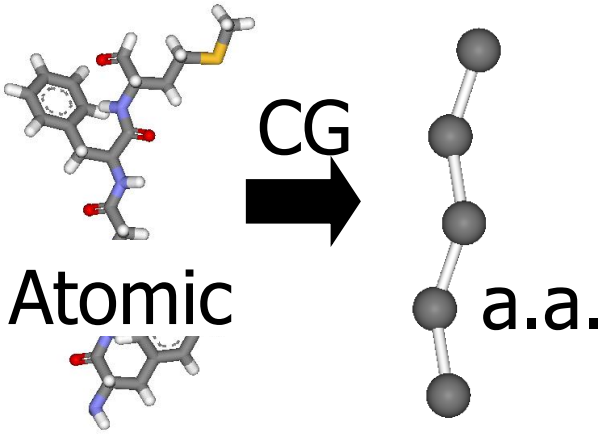
black: optional files

- Output files
  - data file(.data)
  - time-series file(.ts)
  - native-info file(.ninfo)
  - coordinate and velocity
    - PDB format(.pdb)
    - CARD format(.crd, .vdcd)
  - trajectory
    - PDB format(.movie)
    - DCD coordinate(.dcd, .vdcd)
  - PSF file(.psf)
  - restart file(.rst)
  - replica information file(.rep)

Menu

1. Biomolecular simulation and CafeMol
2. Model and applications of CafeMol
3. Recent development of CafeMol

Models and energy functions



1 beads / 1 amino acid

A. Go-like model
B. Atomic interaction based CG(AICG) model
C. Flexible local potential(FLP) model
D. AICG2+ model
E. Multiple basin model
F. DNA/RNA model
G. Elastic network model
H. Electrostatic and hydrophobic interactions
I. Explicit and implicit ligands

Go-like model

C. Clementi, H. Nymeyer, and J.N. Onuchic, J. Mol. Biol. (2000)

Based on the energy landscape theory

Structure based

$$V_{protein} = V_{local} + V_{go} + V_{ex}$$

$$V_{local} = K_b \sum_i \left(r_{i,i+1} - r_{0i,i+1}\right)^2 + K_\theta \sum_i \left(\theta_i - \theta_{0i}\right)^2$$
$$+ K_\phi^1 \sum_i \left(1 - \cos\left(\phi_i - \phi_{0i}\right)\right) + K_\phi^3 \sum_i \left(1 - \cos 3\left(\phi_i - \phi_{0i}\right)\right)$$

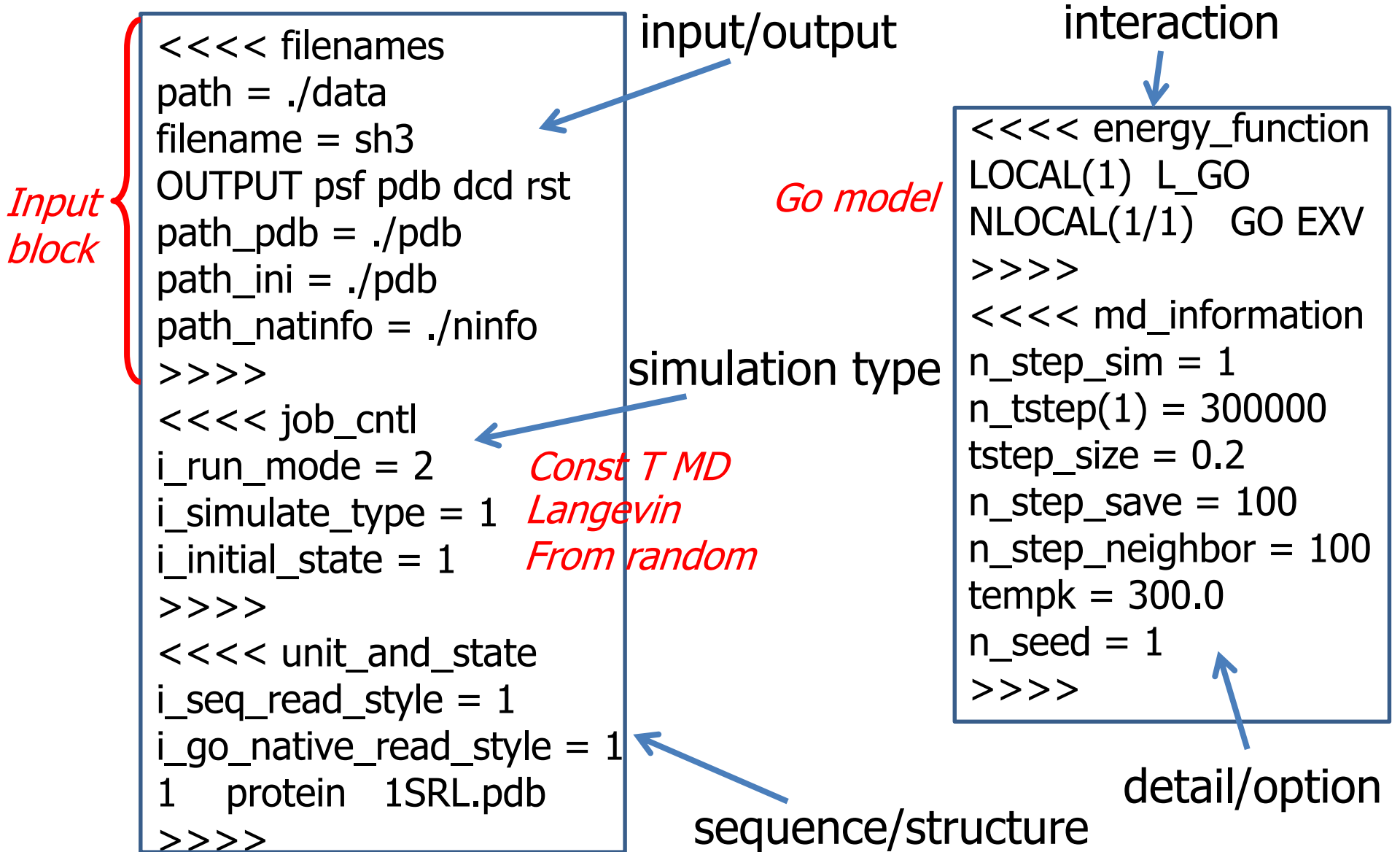
$$V_{go} = \varepsilon_{go} \sum_{i,j}^{native} \left[5\left(\frac{r_{0ij}}{r_{ij}}\right)^{12} - 6\left(\frac{r_{0ij}}{r_{ij}}\right)^{10} \right]$$

$$V_{ex} = \varepsilon_{ex} \sum_{i,j}^{nonnative} \left(\frac{\sigma}{r_{ij}}\right)^{12}$$

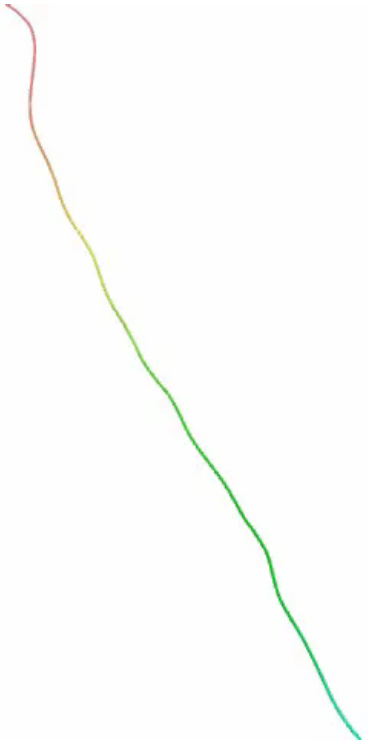
θ: bond angle
ϕ: dihedral angle
(0 means native state)

$K_b = 100\varepsilon$
$K_\theta = 20\varepsilon$
$K_\phi^1 = \varepsilon$
$K_\phi^3 = 0.5\varepsilon$
$\varepsilon_{go} = 0.18\varepsilon$
$\varepsilon_{ex} = \varepsilon$
σ= 4A
ε=1.0kcal/mol

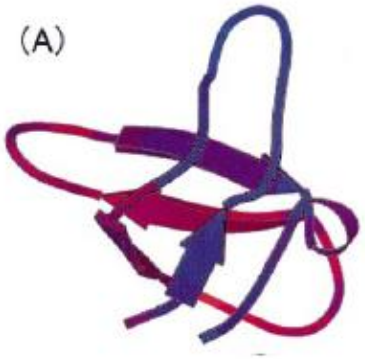
Example of input file (folding simulation of src SH3)



Folding simulation of src SH3 domain

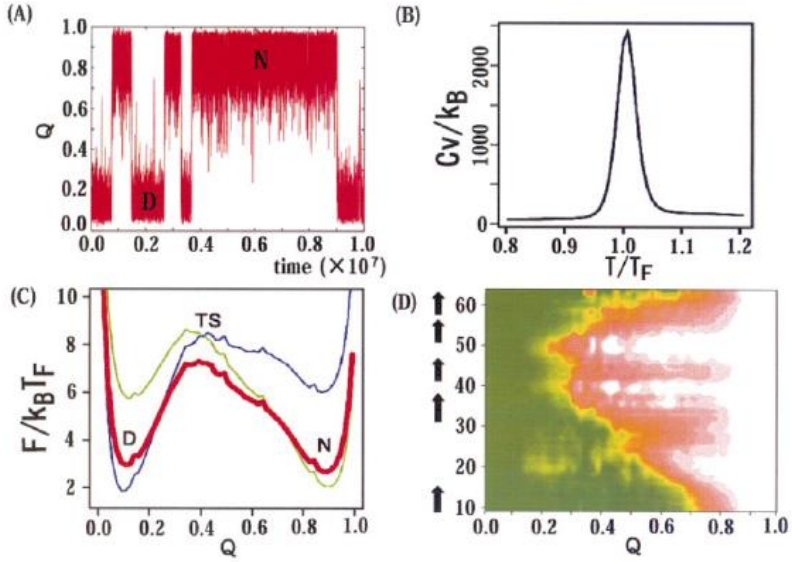


N. Koga, and S. Takada, J. Mol. Biol. (2001)

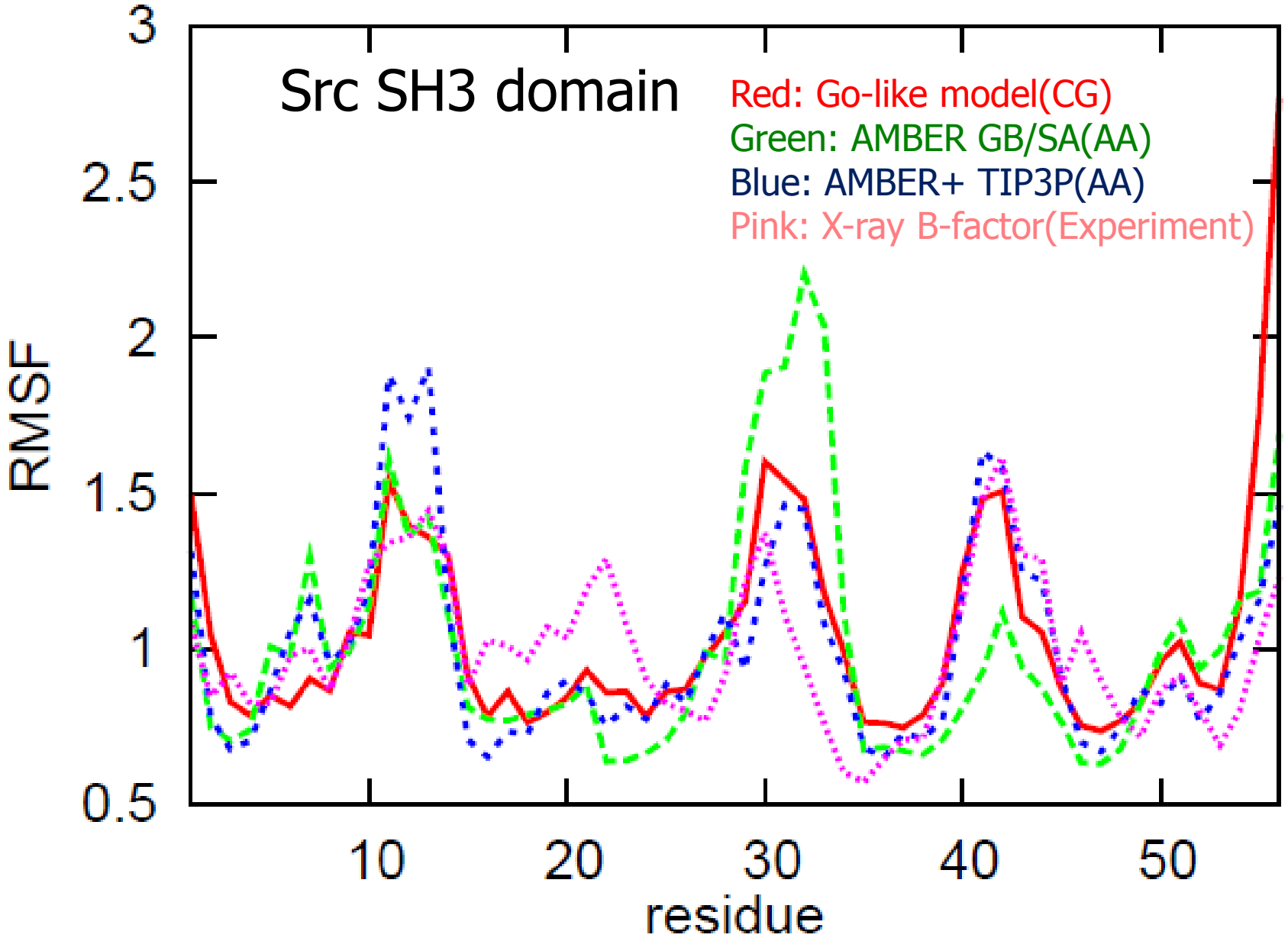


computed ϕ–values

experimental ϕ–values



Native fluctuation by Go-like model



Native-info

all-in-one style

```
<<<< native_info_sim1
NINFO(all/all) f1atp_all.ninfo
>>>>
```

one-by-one style

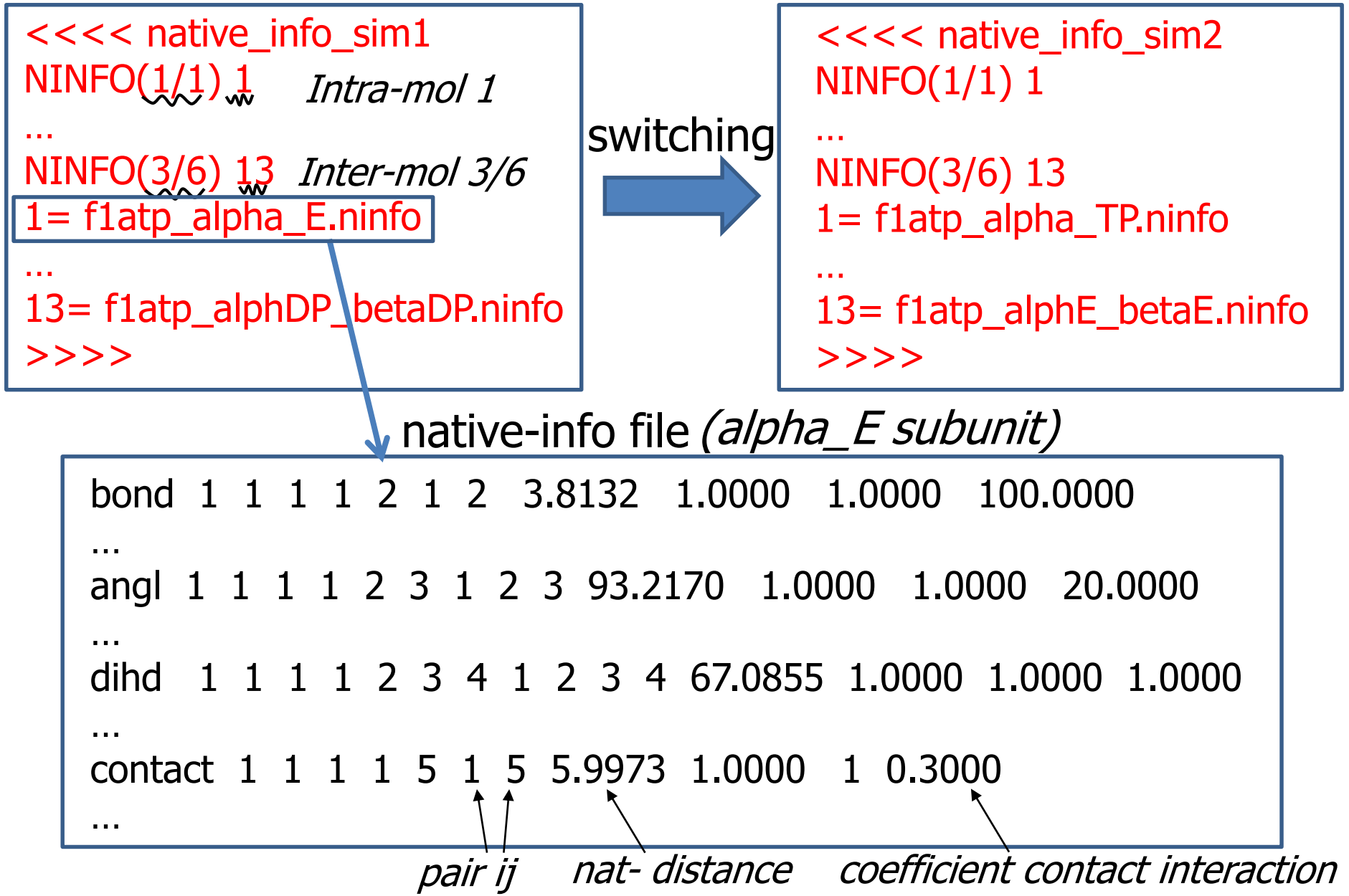
```
<<<< native_info_sim1
NINFO(1/1) 1      Intra-mol 1
…
NINFO(3/6) 13   Inter-mol 3/6
1= f1atp_alpha_E.ninfo
…
13= f1atp_alphDP_betaDP.ninfo
>>>>
```

native-info file *(alpha_E subunit)*

```
bond  1  1  1  1  2  1  2   3.8132   1.0000   1.0000   100.0000
…
angl  1  1  1  1  2  3  1  2  3  93.2170   1.0000   1.0000   20.0000
…
dihd   1  1  1  1  2  3  4  1  2  3  4  67.0855  1.0000  1.0000  1.0000
…
contact  1  1  1  1  5  1  5  5.9973  1.0000   1  0.3000
…
```

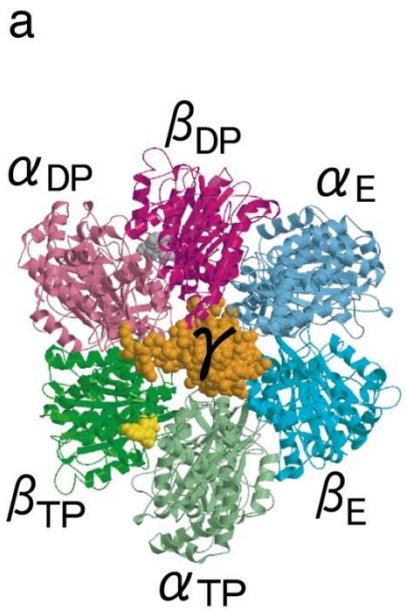
pair ij *nat- distance* *coefficient contact interaction*

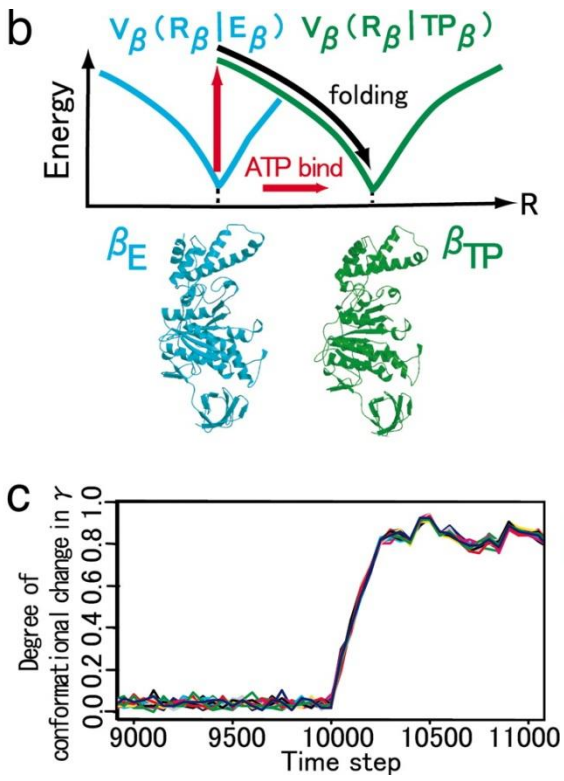
“Switching” simulation

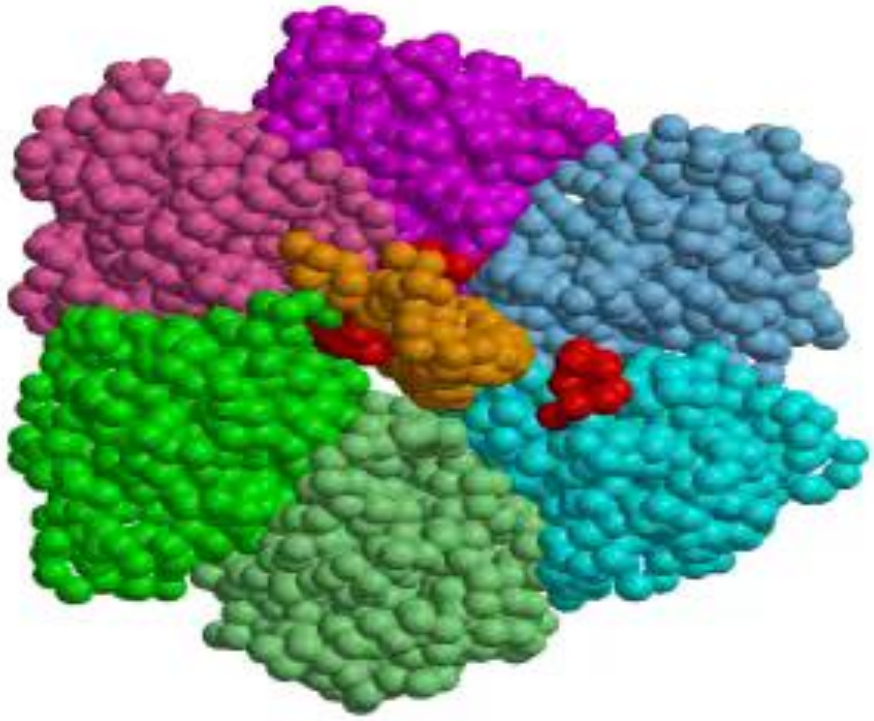


Rotation mechanism of F_1-ATPase by switching Go model

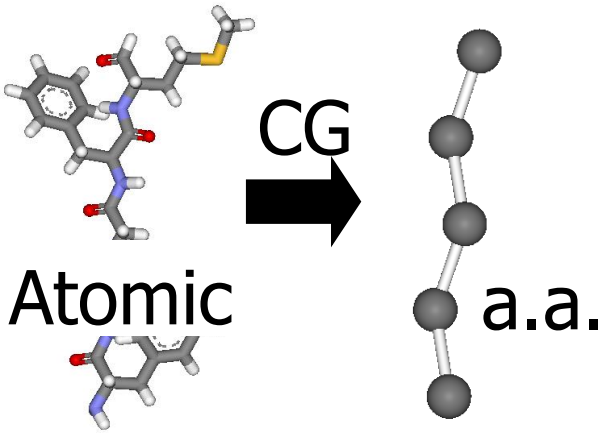
N. Koga, and S. Takada, PNAS (2006)







Atomic interaction based CG (AICG) model



W. Li, P. Wolynes, S. Takada, PNAS (2011)

$$V = \sum_i k_b^i (r^i - r_0^i)^2 + \sum_i k_a^i (\theta^i - \theta_0^i)^2 + \sum_i \{\varepsilon_{\phi,1}^i [1 - \cos(\phi^i - \phi_0^i)] + \varepsilon_{\phi,3}^i [1 - \cos 3(\phi^i - \phi_0^i)]\} + \sum_{i>j-3}^{native} \varepsilon^{ij} [5(r_0^{ij}/r^{ij})^{12} - 6(r_0^{ij}/r^{ij})^{10}] + \sum_{i>j-3}^{non-native} \varepsilon (C/r^{ij})^{12}$$

Wenfei Li

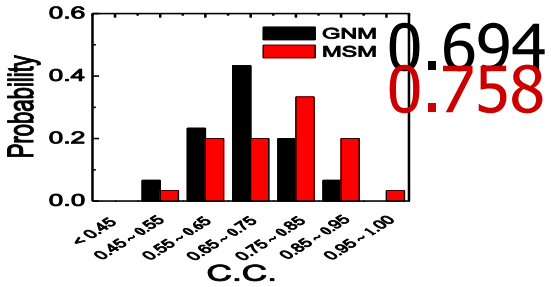
1) Contact energy ε_{ij} from pairwise all-atom (AA) energy

$$E^{IJ}(R_{IJ}) = \sum_{i \in I} \sum_{j \in J} u_{AA}(r_{ij}) \qquad u_{AA}(r) = V(r) + \Delta G_{pol}^{GB}(r) + \Delta G^{SA}(r)$$

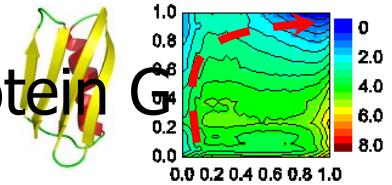
2) Coefficients fitted by AA-derived fluctuation (23 proteins)

param	K_b	K_a^G	k_a^H	k_a^E	k_a^T	k_a^C	ε_ϕ^G	ε_ϕ^H	ε_ϕ^E	ε_ϕ^T	ε_ϕ^C	ε_{nloc}
Av.	109.94	13.40	40.03	17.32	19.35	11.70	0.29	1.76	1.32	0.82	0.81	0.37

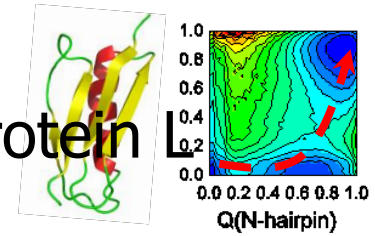
Test for fluctuation, structural change, & folding



Protein G



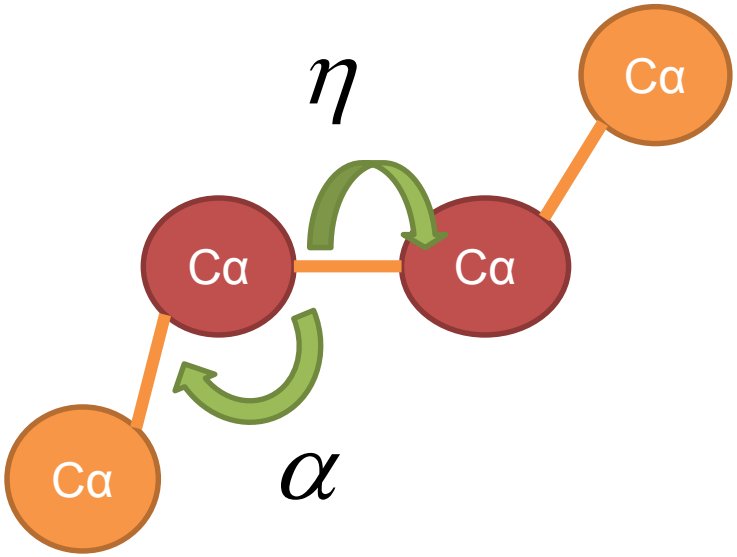
Protein L



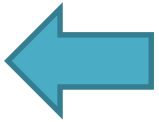
Flexible local potential(FLP) model

T. Terakawa, and S. Takada, Biophys. J. (2011)

Probability distributions to CG potential
for flexible local potentials



$P_a(\alpha)$
$P_d(\eta)$



RCSB PDB PROTEIN DATA BANK
Loop segment library

Boltzmann inversion; prob. to pot.

$$V_a = -k_B T \frac{\ln P_a(\alpha)}{\sin \alpha} \qquad V_d = -k_B T \ln P_d(\eta)$$

AICG2+ model

W. Li, T. Terakawa, W. Wang, S. Takada, PNAS (2012)

improved AICG model + FLP model

$$V = \sum_I k_b (r^I - r_0^I)^2$$

$$+ \sum_I V_a^I(\theta^I) + \sum_I V_{dih}^I(\phi^I)$$

$$+ \sum_{J=I+2} \varepsilon_{1,3}^{IJ} \exp\left(-\frac{(r^{IJ} - r_0^{IJ})^2}{2w^2}\right) + \sum_{J=I+3} \varepsilon_{1,4}^{IJ} \exp\left(-\frac{(\phi^I - \phi_0^I)^2}{2w_\phi^2}\right) +$$

$$+ \sum_{I>J+3}^{native} \varepsilon_{nloc}^{IJ} [5(r_0^{IJ} / r^{IJ})^{12} - 6(r_0^{IJ} / r^{IJ})^{10}] + \sum_{I>J+3}^{non-native} \varepsilon (C / r^{IJ})^{12}$$

Multiple-basin model for proteins

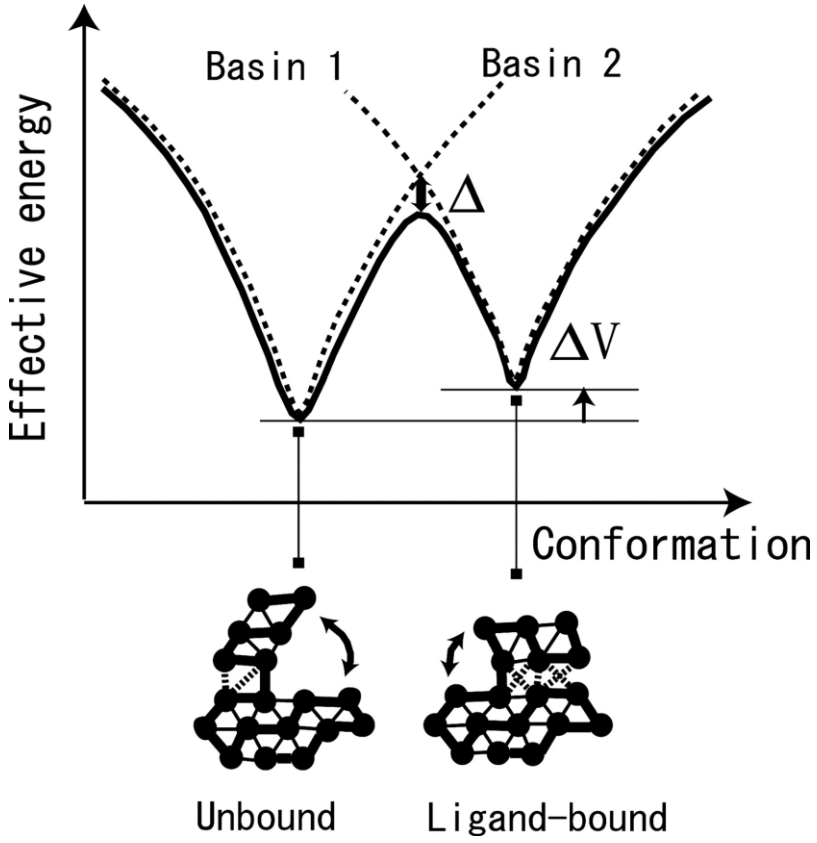
K. Okazaki, N. Koga, S. Takada, J.N. Onuchic, and P.G. Wolynes, PNAS (2006)

Use of 2 references

$$\begin{pmatrix} V(R|R_1) & \Delta \\ \Delta & V(R|R_2)+\Delta V \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = V_{MB}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$$

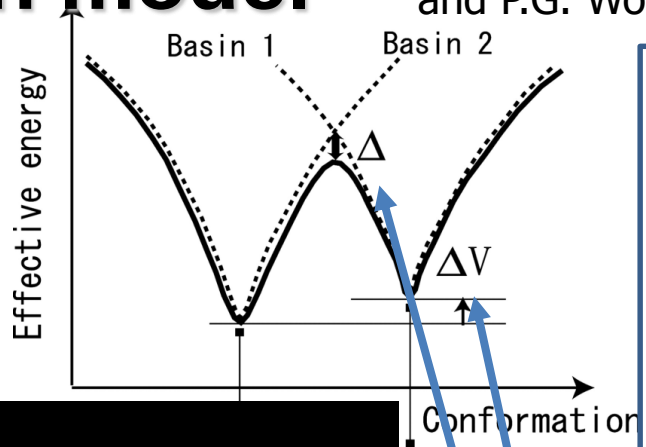
$$V_{MB} = \frac{V(R|R_1)+V(R|R_2)+\Delta V}{2} - \sqrt{\left(\frac{V(R|R_1)-V(R|R_2)-\Delta V}{2}+\Delta\right)}$$

$$\chi = \log\left(\frac{c_2}{c_1}\right)$$



Conformational change by Multi-basin model

K. Okazaki, N. Koga, S. Takada, J.N. Onuchic, and P.G. Wolynes, PNAS (2006)

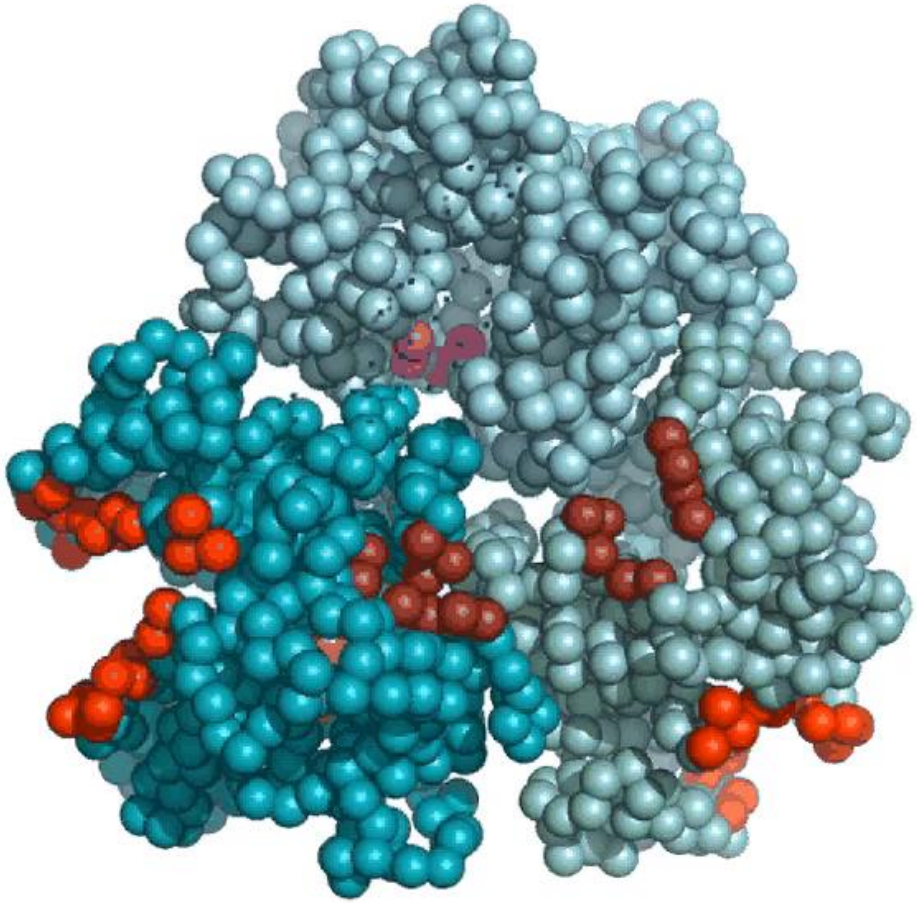


nd-bound

```
<<<< unit_and_state
i_seq_read_style = 1
i_go_native_read_style = 1
1a  protein  1GGG_2.pdb
1b  protein  1WDN_2.pdb
>>>>
<<<< energy_function
NLOCAL(1a/1a)    GO EXV
NLOCAL(1b/1b)    GO EXV
MULTIGO_SYSTEM(1a)  1a/1a
MULTIGO_SYSTEM(1b)  1b/1b
>>>>
<<<< multiple_go
bdemax_mgo = 100.0
baemax_mgo = 1.0
dihemax_mgo = 0.5
ENEGAP(1)(1)  0.0   -1.8
DELTA(1ab)    28.0
>>>>
```

Conformational change using 3-state multi-basin model

X. Yao, H. Kenzaki, S. Murakami, and S. Takada, *Nature Comm.* (2010)



Multidrug transporter AcrB

- Largely responsible for multidrug resistance
- Asymmetric homo-trimer structure

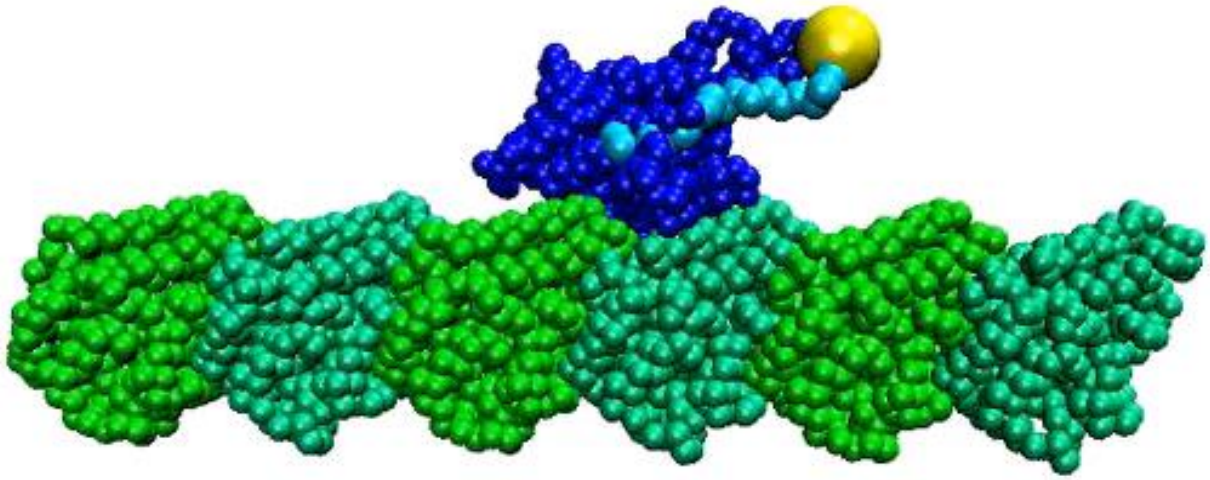
Suggesting mechanism of drug exportation and functional rotation!

Sliding movement of KIF1A

R. Kanada, T. Kuwata, H.Kenzaki, S.Takada, PLOS Comput. Biol. (2013)

1 phase: multiple-basin (T, D)
2 phase: go(D)
3 phase: multiple-basin(D, phi)
4 phase: go(phi)
5 phase: go(T)

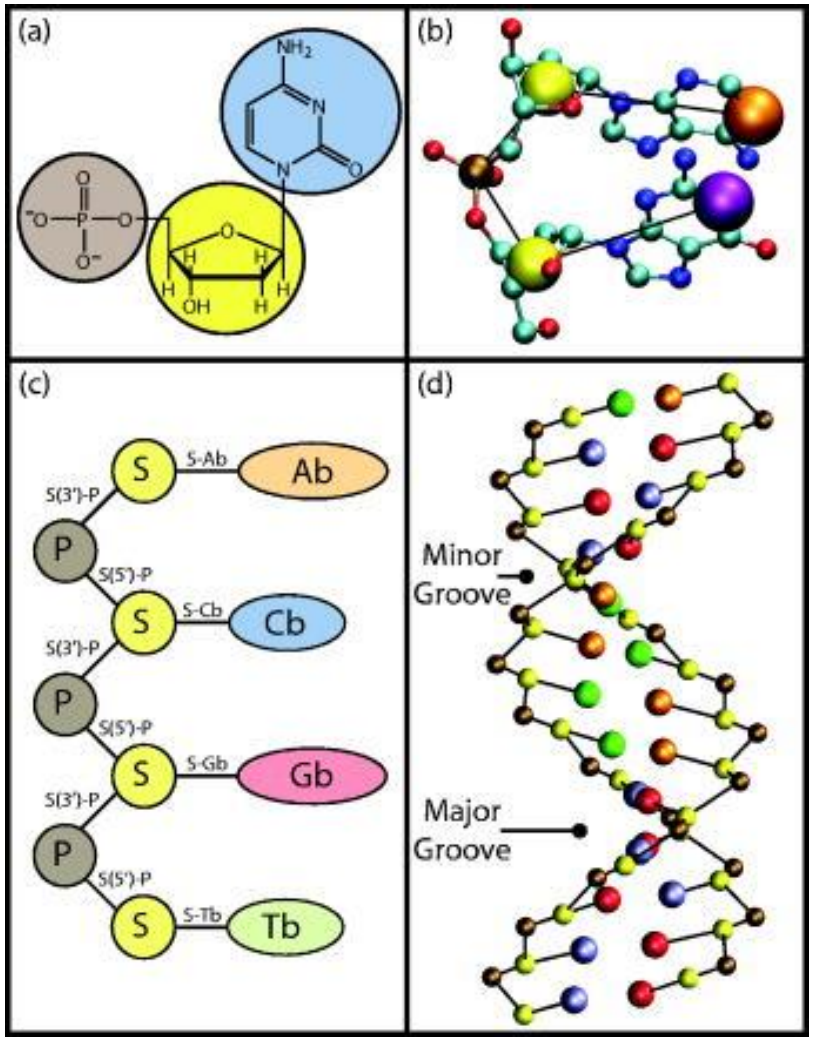
KIF1A:blue
tubulin:green
cargo:yellow



CG DNA model

T.A. Knotts IV, N.Rathore, D.C. Shwartz, and J.J. Pablo, J. Chem. Phys. (2007)

- Three interactions sites
  - Phosphate
  - Sugar
  - Base
- Reproduce various DNA behavior
  - Salt-dependent melting
  - Bubble formation
  - Mechanical properties



3SPN.1 force field

E.J. Sambrisiki, D.C. Schwartz, and J.J. de Pablo,Knotts, Biophys. J. (2009)

$$V_{dna} = V_{local} + V_{stack} + V_{bp} + V_{ex} + V_{qq} + V_{solv}$$

$$V_{local} = K_{b1}\sum_i \left(r_{i,i+1} - r_{0i,i+1}\right)^2 + K_{b2}\sum_i \left(r_{i,i+1} - r_{0i,i+1}\right)^4$$

$$+ K_\theta \sum_i \left(\theta_i - \theta_{0i}\right)^2 + K_\phi \sum_i \left(1 - \cos\left(\phi_i - \phi_{0i}\right)\right)$$

$$V_{stack} = 4\varepsilon_1 \sum_{i,j}^{Nst} \left[\left(\frac{\sigma_{0ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{0ij}}{r_{ij}}\right)^{6} \right]$$

$$V_{bp} = \sum_{i,j}^{Nbp} 4\varepsilon_{bpi} \left[5\left(\frac{r_{0ij}}{r_{ij}}\right)^{12} - 6\left(\frac{r_{0ij}}{r_{ij}}\right)^{10} \right]$$

$$V_{ex} = 4\varepsilon_1 \sum_{i,j}^{Nex} \left[\left(\frac{\sigma_0}{r_{ij}}\right)^{12} - \left(\frac{\sigma_0}{r_{ij}}\right)^{6} \right] + \varepsilon_1 \,(if\ r_{ij} < d_{cut}),$$

$$= 0\ (if\ r_{ij} > d_{cut})$$

θ: bond angle
ϕ: dihedral angle
(0 means B-type DNA)

Structre-based interaction

$K_{b1} = 1\varepsilon$
$K_{b2} = 100\varepsilon$
$K_\theta = 1400\varepsilon$
$K_\phi = 28\varepsilon$
$\varepsilon_{bpGC} = 2.532\varepsilon$
$\varepsilon_{bpAT} = 2.0\varepsilon$
ε =0.1839kcal/mol

3SPN.1 force field (electrostatic and solvation interaction)

$$V_{qq} = \sum_{i,j}^{N} \left(\frac{q_i q_j}{4\pi\varepsilon_0 \varepsilon(T,C) r_{ij}} \right) e^{-r_{ij}/\lambda_D}$$

Debye-Huckel theory

$$\varepsilon(T,C) = \varepsilon(T)a(C)$$

ε=78

$$\varepsilon(T) = 249.4 - 0.788T/K + 7.20\times10^{-4}(T/k)^2$$

$$a(C) = 1.000 - 0.2551C/M$$

$$+ 5.151\times10^{-2}(C/M)^2 - 6.889\times10^{-3}(C/M)^3$$

Debye length

$$\lambda_D = \left(\frac{\varepsilon_0 \varepsilon RT}{2N_A^2 e_q^2 I} \right)$$

$$V_{solv} = \sum_{i<j}^{N_{solv}} \varepsilon_s \left[1 - e^{-a(r_{ij} - r_s)} \right]^2 - \varepsilon_s$$

$$\varepsilon_s = \varepsilon_N A_I$$

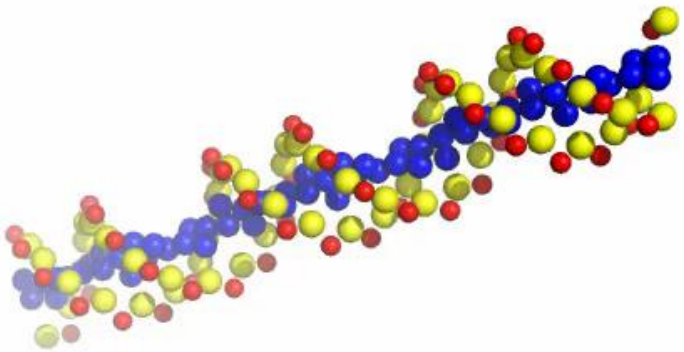
$$e_N = e_0 (1 - [1.40418 - 0.268231 N_{nt}]^{-1})$$

$$A_I = 0.474876(1 + \{0.148378 + 10.9553[Na^+]\}^{-1})$$

α^{-1} = 5.333A
r_s = 13.38A
$\varepsilon_0 = 0.504982\varepsilon$

CG DNA simulation

- 30 bp DNA duplex
- Langevin dynamics (300K)
- $[Na^+]$ = 69mM
- cutoff length $20\kappa_D$(Debye length)



```
<<<< unit_and_state
i_seq_read_style = 2
i_go_native_read_style = 3
1-2    dna     sequence
>>>>
<<<< energy_function
LOCAL(1-2)  L_BDNA
NLOCAL(1-2/1-2)  DNA ELE
>>>>
<<<< electrostatic
cutoff_ele = 20.0
ionic_strength = 0.069
diele_water = 78.0
>>>>
<<<< in_box
xbox = 120.0
ybox = 120.0
zbox = 120.0
boxsigma = 4.0
>>>>
```

DH

3SPN.1

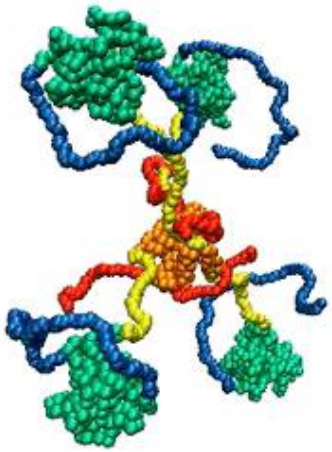
Intra mol 1,2

Inter mol 1-2

Diffusional search of p53 on DNA

T. Terakawa, H. Kenzaki, and S. Takada, JACS, (2012)

- Tumor suppressor
- Transcription factor
  1. Nonspecific diffusional search on DNA
  2. Specific binding



protein:AICG and FLP model
DNA:CG DNA model
protein-DNA:
excluded and electrostatic interaction

Core domain hopping
C-terminal domain sliding

Simulation of nucleosome

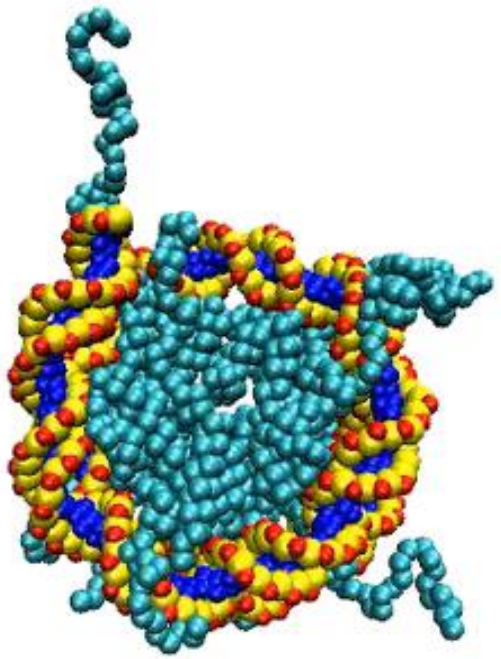
1KX5.pdb

- Electrostatic interaction + Go potential

$\varepsilon_{go}^{\text{pro-dna}} = 0.8\varepsilon_{go}^{\text{pro}}$
$[Na^+]$ = 200mM

$\varepsilon_{go}^{\text{pro-dna}}$:coefficient of protein-DNA Go potential

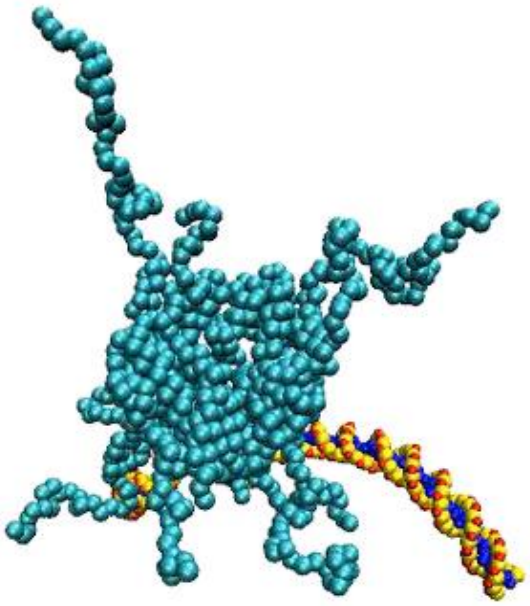
```
<<<< energy_function
LOCAL(1-2) L_BDNA
LOCAL(3-10) L_GO
NLOCAL(1-2/1-2) ELE DNA
NLOCAL(1-2/3-10)  GO EXV ELE
NLOCAL(3-10/3-10)  GO EXV
>>>>
<<<< electrostatic
cutoff_ele = 5.0
ionic_strength = 0.2
diele_water = 78.0
>>>>
```



H. Kenzaki, et al, Plos Comput. Biol. to be published

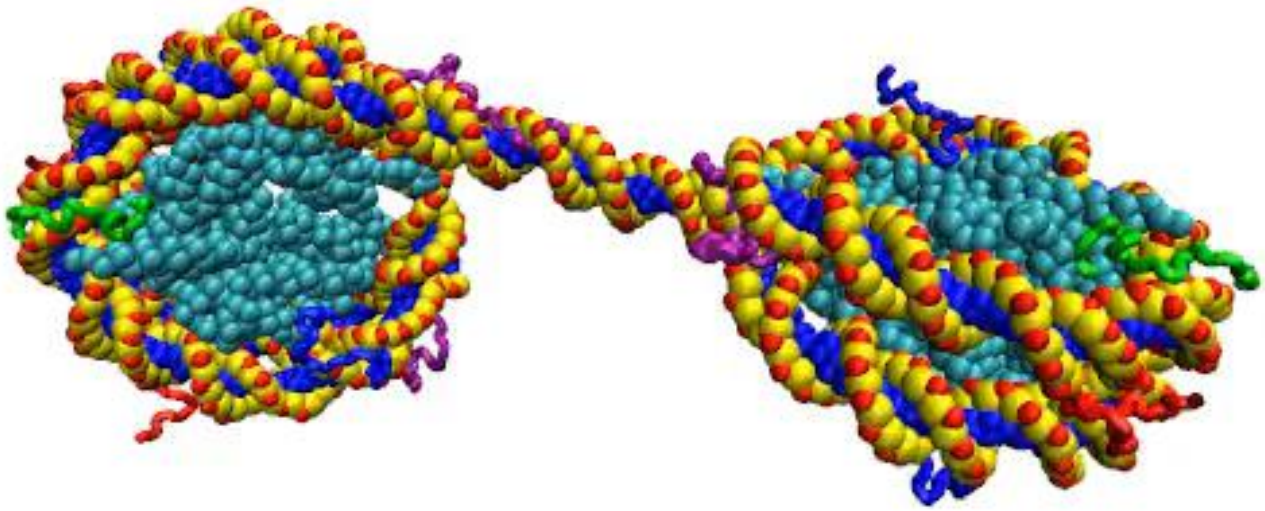
Nucleosome wrapping simulation
For demonstration

Na^+=200mM



Dinucleosome wrapping simulation
For demonstration

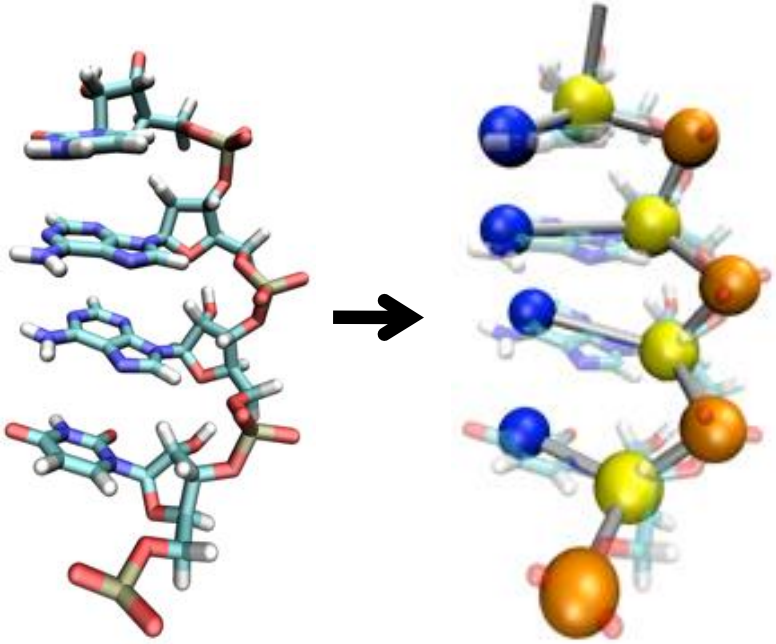
Na^+=100mM



CG RNA model

N. Hori, and S. Takada, JCTC (2012)

1 nucleotide = 3 beads



- Phosphate (phosphorus atom)
- Sugar (center of ribose ring)
- Base (pyrimidine: N1 atom)
  (purine: N3 atom)

CG RNA model (local)

$$V_{\text{total}} = V_{\text{local}} + V_{\text{stack}} + V_{\text{basepair}} + V_{\text{nonlocal}} + V_{\text{exclude}}$$

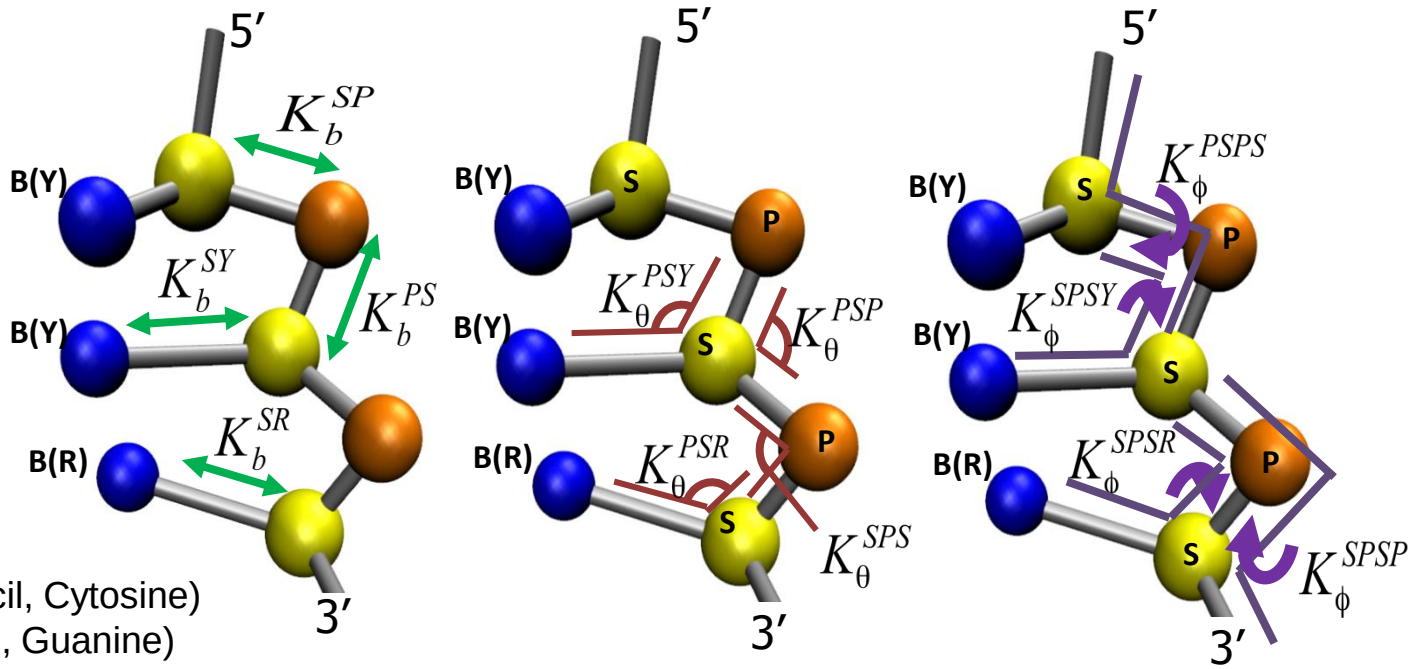
Bond length **Bond angle**

Local

$$V_{local} = \sum_{ibd\in \substack{PS,\\ SP,\\ SB}} K_b^{\eta}(r_{ibd} - r_{ibd}^{(0)})^2 + \sum_{iba\in \substack{PSP,\\ SPS,\\ SPB}} K_{\theta}^{\eta}(\theta_{iba} - \theta_{iba}^{(0)})^2$$

$$+ \sum_{idih\in \substack{PSPS,\\ SPSP,\\ SPSB}} \{K_{\phi}^{\eta}(1 - cos(\phi_{idih} - \phi_{idih}^{(0)})) + 0.5K_{\phi}^{\eta}(1 - cos3(\phi_{idih} - \phi_{idih}^{(0)}))\}$$

Dihedral angle



Base type
Y : Pyrimidine (Uracil, Cytosine)
R : Purine (Adenine, Guanine)

CG RNA model (nonlocal)

Nonlocal

Reference (X-ray structure's) value

$$V_{\text{contact}} = \sum_{i,j \in S,B}^{\text{native contact}} \varepsilon_\xi \left[5 \left(\frac{r_{ij}^0}{r_{ij}} \right)^{12} - 6 \left(\frac{r_{ij}^0}{r_{ij}} \right)^{10} \right]$$

$$\varepsilon_\xi = \varepsilon_{BP3}$$

$$\varepsilon_\xi = \varepsilon_{BP2}$$

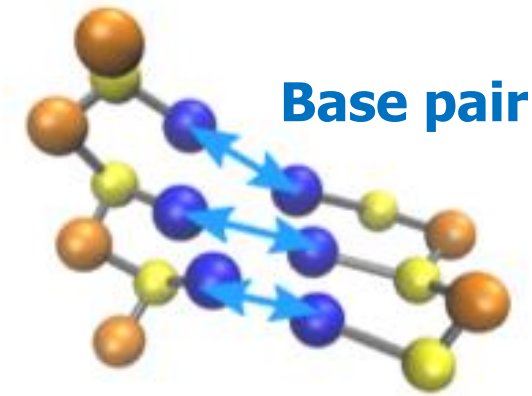
Base pair

$$\varepsilon_\xi = \varepsilon_{\text{ST}}$$

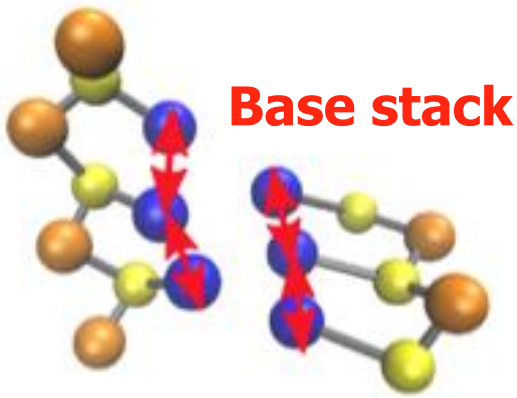
Base stack

$$\varepsilon_\xi = \varepsilon_{con}^{\eta}$$

Contact (Other)



Base pair



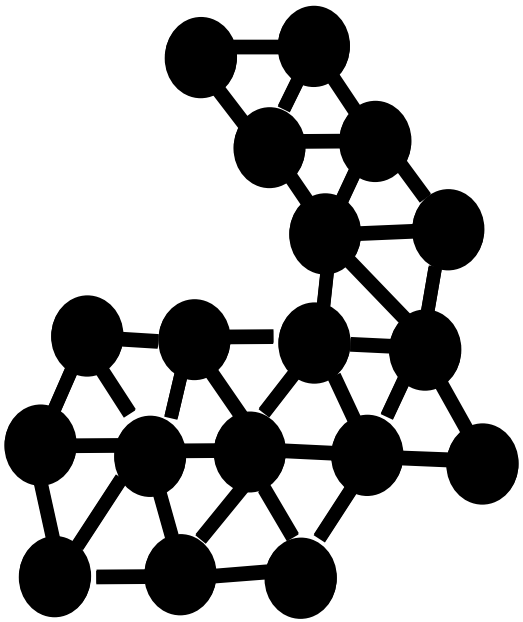
Base stack

$$V_{exclude} = \sum_{i,j}^{\text{non-native}} \varepsilon_{ex} \left(\frac{\sigma}{r_{i,j}} \right)^{12}$$

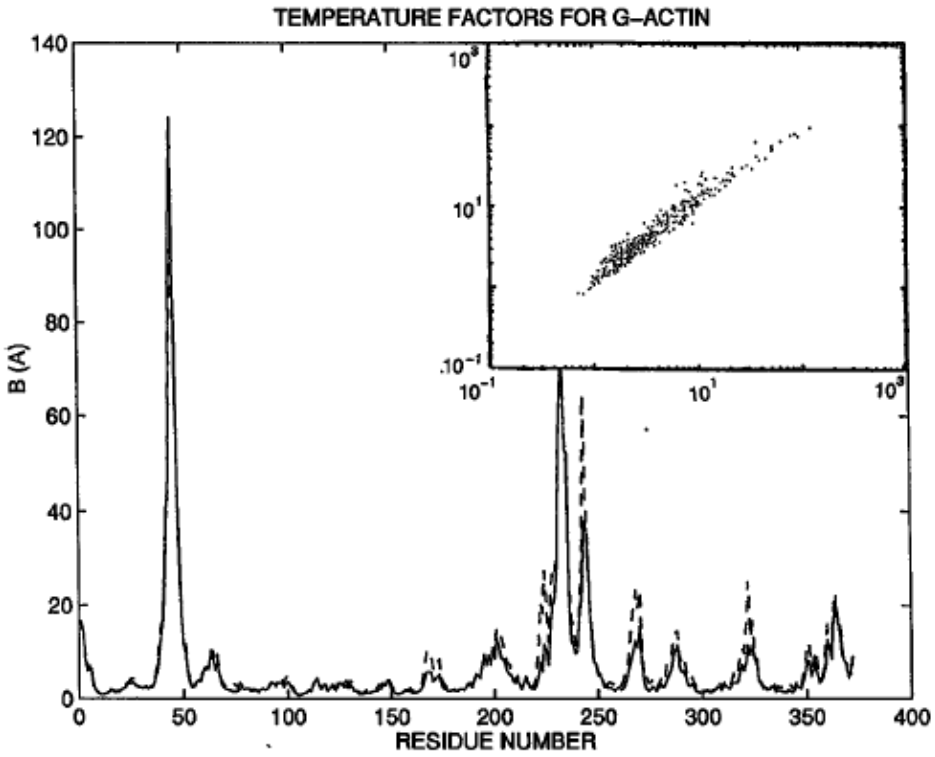
Excluded volume

Elastic network model

$$E = \sum_{ij,s.t.r_{ij}^0 < r_c} K(r_{ij} - r_{ij}^0)^2$$



Atomic fluctuation reproduced by ENM (Tirion96)



Electrostatic and hydrophobic interactions

Debye-Huckel form for electrostatics

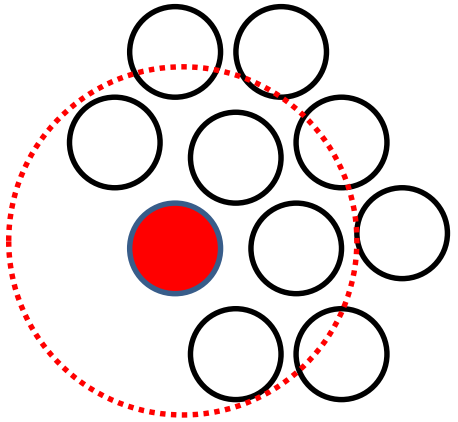
$$V_{\text{ele}} = \sum_{i<j}^{N} \frac{q_i q_j}{4\pi\epsilon_0\epsilon_k r_{ij}} e^{-r_{ij}/\kappa_D}$$

Debye length

$$\kappa_D = \left(\frac{\varepsilon_0 \varepsilon RT}{2N_A^2 e_q^2 I} \right)$$

HP interactions analogous to ASA

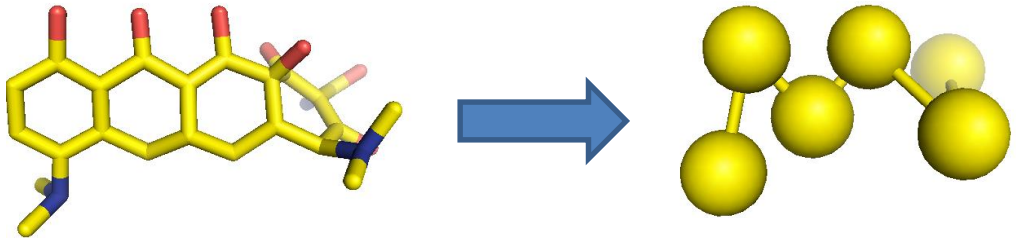
$$V_{\text{HP}} = -c_{\text{HP}} \sum_{i \in \text{HP}} \epsilon_{\text{HP},A(i)} S_{\text{HP}}(\rho_i)$$



Count coordination number
for each hydrophobic particle

Explicit and Implicit ligands

Explicit ligand: as a rigid molecule



Implicit ligand:
MD-MC scheme with ligand-mediated contact

$$V_{protein} \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} V_{protein} + V_{imp-lig}$$

$$V_{imp-lig} = \sum_{\substack{ligand-mediated \\ contact-pairs}} -c_{lig}\varepsilon_{go} \exp\left[-\frac{(r_{ij}/r_{0ij}-1)^2}{2(\sigma/r_{0ij})^2}\right]$$

Useful option1

- redefine parameters

dfcontact = 4.5 definition of native contact(default 6.5A)

cdist_rep12 = 6.0 reference distance in excluded interaction(default 4.0A)

rneighbor_dist = 20.0 truncation distance for neighbor list(default 24.0A)

fric_const = 0.02 friction constant(default 0.25)

- delete interaction

DEL_LGO(mp_{ini}-mp_{las})

DEL_GO($mp_{ini,i}$-$mp_{las,i}$/$mp_{ini,j}$-$mp_{las,j}$)

- box: fix some particles or units

xbox = boxsize_x

ybox = boxsize_y

zbox = boxsize_z

boxsigma = sigma

$$V_{box} = 0 \qquad (d > 3\sigma),$$

$$= k_{box}\left(\frac{\sigma}{d}\right)^{12} \qquad (0.5\sigma < d < 3\sigma),$$

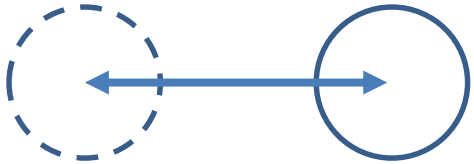
$$= k_{box}\left(\frac{\sigma}{0.5\sigma}\right)^{12}\left(1+12\,\frac{0.5\sigma - d}{0.5\sigma}\right) \quad (d < 0.5\sigma)$$

Useful option2

- anchor: particle i constrain to some position

ANCH i k_i l_0 x_0 y_0 z_0

$$V_{anchor} = k_i (r_{i0} - l_0)^2 \ (r_{i0} > l_0),$$
$$= 0 \ (r_{i0} < l_0)$$



- bridge: bind I and j particles by a harmonic spring

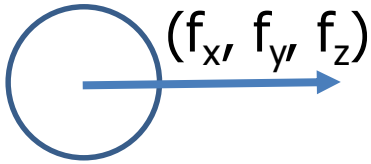
BRIDGE i j k_{ij} l_0

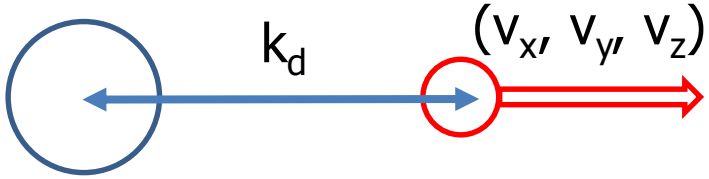
$$V_{bridge} = k_{ij} (r_{ij} - l_0)^2 \ (r_{ij} > l_0),$$
$$= 0 \ (r_{ij} < l_0)$$

- pulling: particle I is pulled by constant force or constant velocity

PULL_CF i f_x f_y f_z

PULL_CV i k_d v_x v_y v_z x_0 y_0 z_0





- fix: fix some particles or units

FIX_UNIT($unit_{ini}$-$unit_{las}$)

FIX_MP(mp_{ini}-mp_{las})

Menu

1. Biomolecular simulation and CafeMol
2. Model and applications of CafeMol
3. Recent development of CafeMol

Recent development of CafeMol

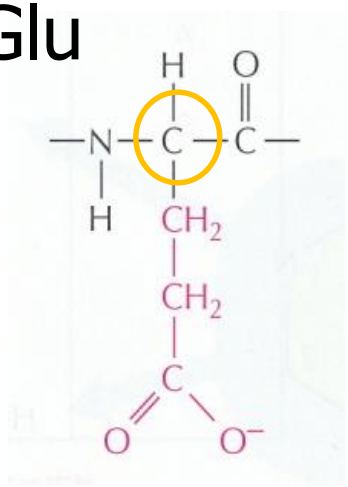
- Partial charge
  - RESPAC by Terakawa
- CG GBSA
  - CG Accessible Surface Area (ASA) by Li
  - CG Generalized Born (GB) by Chang
- Multicanonical simulation by Ito
- New DNA models
  - 3SPN.2 by Terakawa and Tan
  - 3SPN.2C by Tan
- Refined hydrophobic interaction by Tanaka and Kanada
- Implicit hydrodynamics by Hori
- 2-beads model by Li

Determination of partial charge in Cα CG model

Terakawa

JCTC 2013

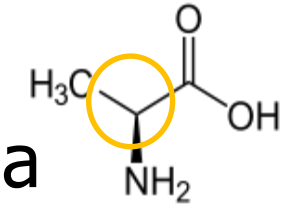
Glu



Conventional **partial charge**

-1 -0.3

Ala



0 -0.7

protein specific partial charge

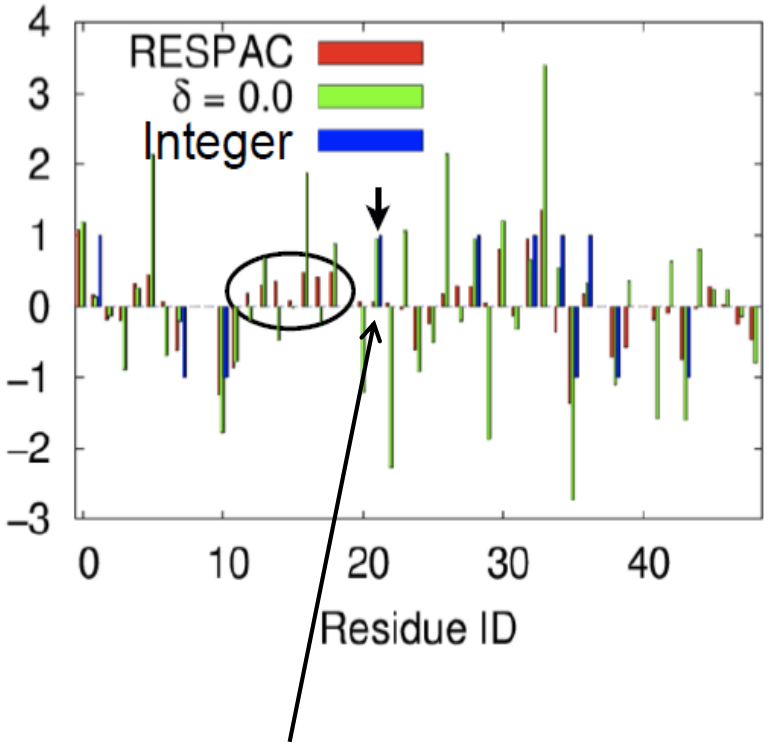
RESP-like procedure from Atom to CG (RESPAC)

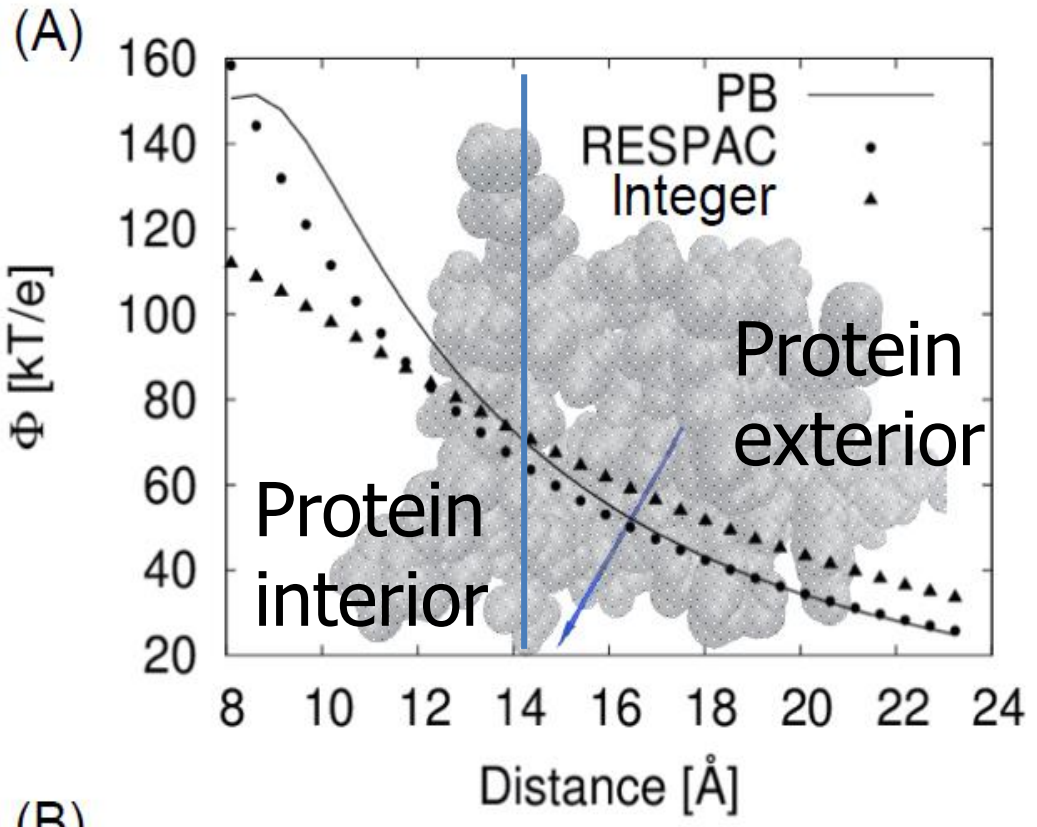
Minimize χ w.r.t. q's

$$\chi_{RESPAC}(\phi_{ref}^{PB},\vec{q}) = \int_{\Omega} d^3\vec{r}\,[\phi_{ref}^{PB} - \sum_i q_i \frac{exp(-\kappa\,|\vec{r}-\vec{r}_i|)}{\varepsilon\,|\vec{r}-\vec{r}_i|}]^2 + \delta\sum_i (q_i - q_{0i})^2 + \lambda(q_{tot} - \sum_i q_i)^2$$

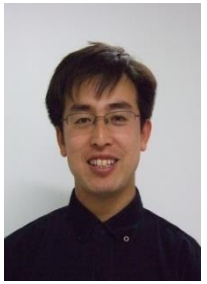
Determination of partial charge in Cα CG model

Lac repressor





CG Accessible Surface Area（ASA）

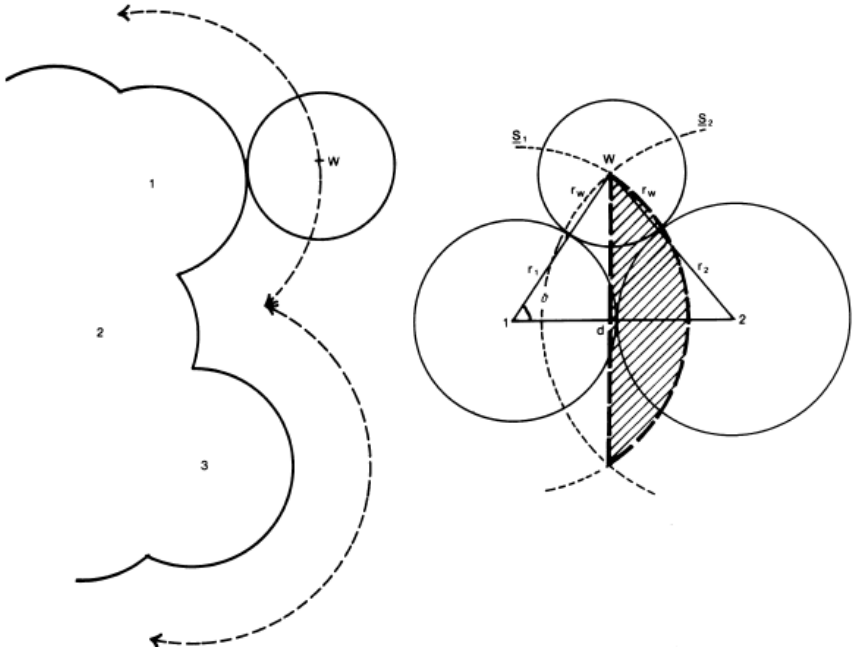


Li

POPS: Fraternali & Cavallo, NAR 2002,

Based on Wodak & Jenin, PNAS(1980)

$$SASA = \sum_i A_i$$



$$A_i(r^N) = S_i \prod_{\substack{j=1 \\ j \neq i}}^{N} \left(1 - \frac{p_i p_{ij} b_{ij}\left(r_{ij}\right)}{S_i} \right)$$

$$b_{ij} = 0; \quad \text{when } r_{ij} \geq R_i + R_j + 2R_{solv}$$

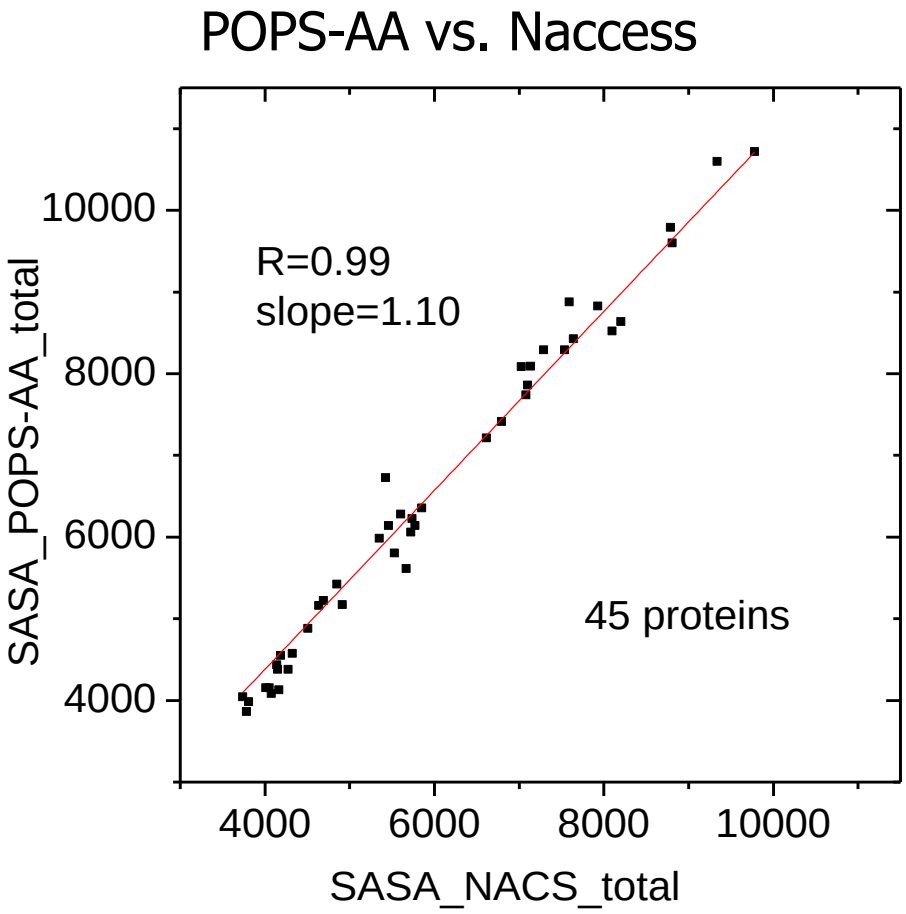
$$b_{ij} = \pi \left(R_i + R_{solv}\right)\left(R_i + R_j + 2R_{solv} - r_{ij}\right)$$

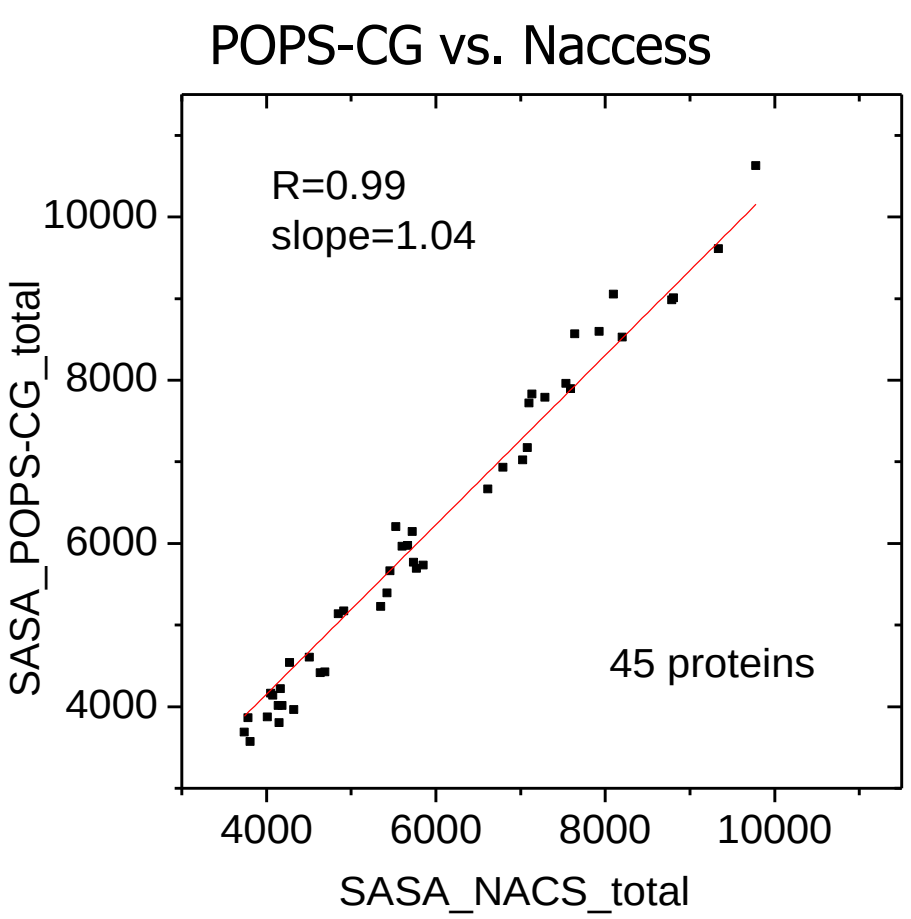
$$\left(1 + (R_j - R_i) r_{ij}^{-1}\right); \text{when } r_{ij} < R_i + R_j + 2R_{solv}$$

$$S_i = 4\pi (R_i + R_{solv})^2$$

CG Accessible Surface Area（ASA）

Total SASA by POPS and Naccess



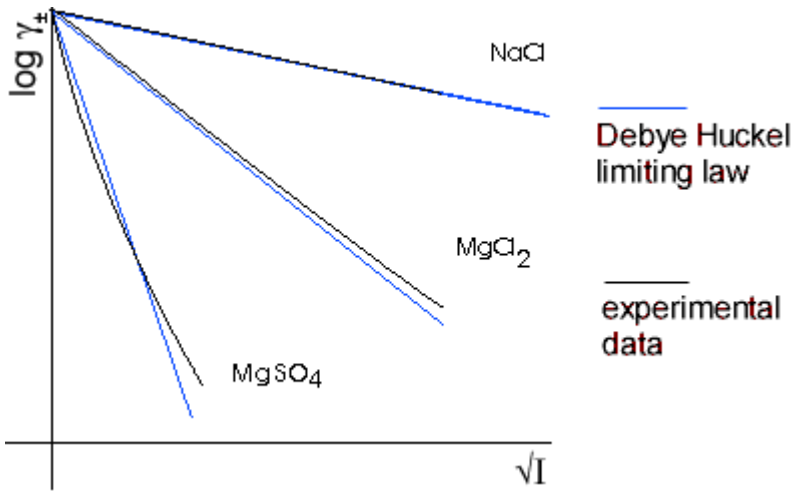


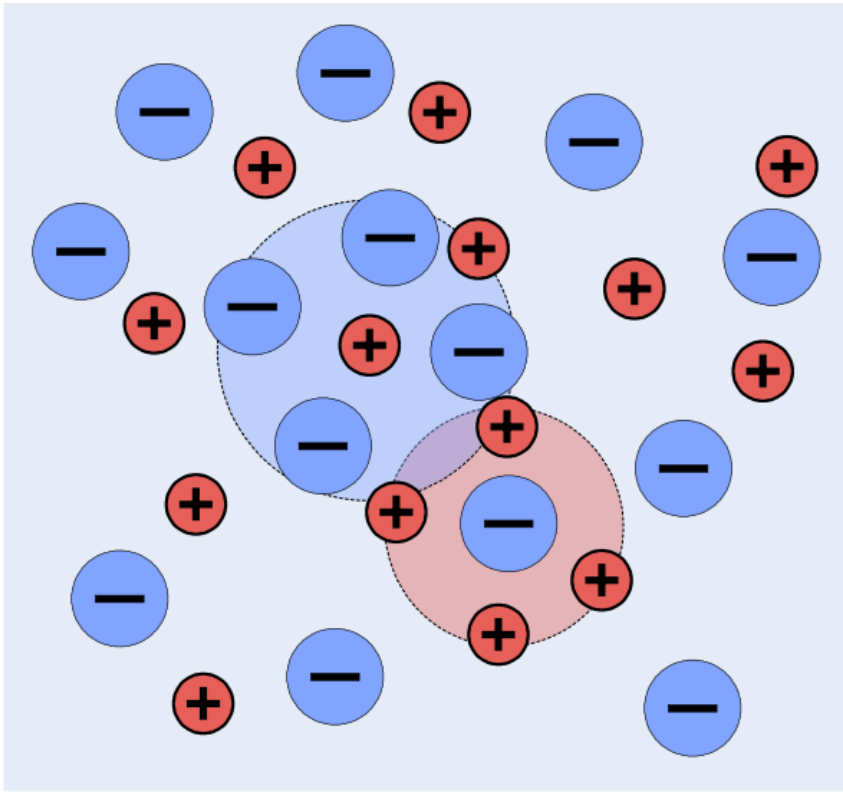
Debye-Huckel model (ionic solution)

$$V_{ele} = \sum_{I<J} \frac{q_I q_J}{4\pi\varepsilon_0\varepsilon_k r^{IJ}} e^{-r^{IJ}/\kappa_D},$$

$$\kappa_D = \left(\frac{\varepsilon_0\varepsilon_k k_B T}{2N_A e^2 I}\right)^{0.5},$$

Ionic strength





CG Generalized Born Model

Still's formulae for GB

$$PMF_{elec} = \frac{1}{2}\sum_{i \neq j}\frac{q_i q_j}{\varepsilon_{in} r_{ij}} - \frac{1}{2}\sum_{i,j}\left(\frac{1}{\varepsilon_{in}} - \frac{1}{\varepsilon_{w}}\right)\frac{q_i q_j}{f_{GB}(r_{ij})}$$

$$f_{GB}(r_{ij}) = \left[r_{ij}^2 + R_i R_j \exp\left(-r_{ij}^2 / 4 R_i R_j\right)\right]^{\frac{1}{2}}$$

Born radius *Ri* is defined by

$$R_i = -166 / G_{pol,i}$$

$$\frac{1}{R_i} = \frac{1}{R_{vdW,i}} - \int_{in, r > R_{vdW,i}} \frac{1}{r_i^4} dV$$

Based on Dominy-Brooks formulation of GB

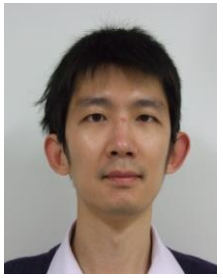
$$G_{\mathrm{pol},i} = \left(1 - \frac{1}{\epsilon}\right)\left[\frac{1}{\lambda}\left(\frac{-166}{R_{\mathrm{vdW},i}}\right) + P_1\left(\frac{166}{R_{\mathrm{vdW},i}^2}\right) + \sum_j^{\mathrm{bond}} \frac{P_2 V_j}{r_{ij}^4} + \sum_j^{\mathrm{angle}} \frac{P_3 V_j}{r_{ij}^4} + \sum_j^{\mathrm{nonbond}} \frac{P_4 V_j}{r_{ij}^4}\mathrm{CCF}\right]$$

where CCF =

$$\begin{cases} 1.0; & \left(\frac{r_{ij}}{R_{\mathrm{vdW},i} + R_{\mathrm{vdW},j}}\right)^2 > \frac{1}{P_5} \\ \left\{0.5\left[1.0 - \cos\left(\left(\frac{r_{ij}}{R_{\mathrm{vdW},i} + R_{\mathrm{vdW},j}}\right)^2 P_5 \pi\right)\right]\right\}^2; & \left(\frac{r_{ij}}{R_{\mathrm{vdW},i} + R_{\mathrm{vdW},j}}\right)^2 \leq \frac{1}{P_5} \end{cases}$$

CG Generalized Born Model

via multiscale fitting of Born self energy

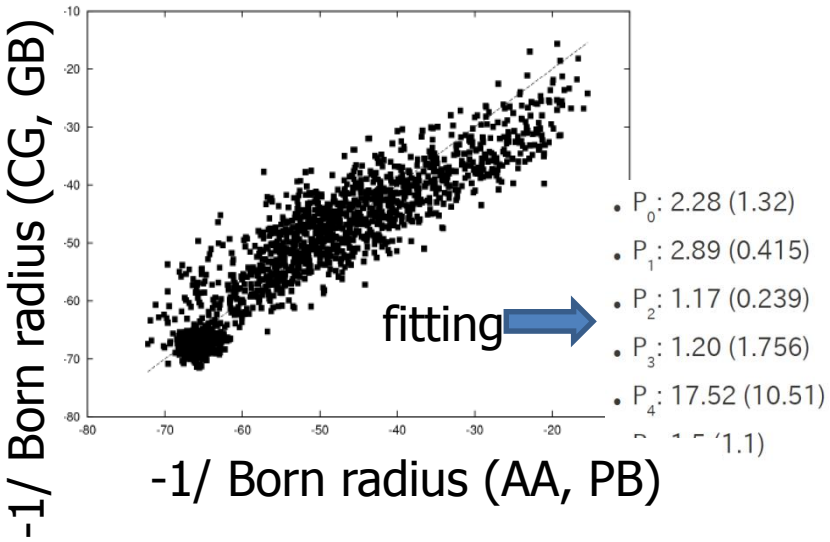


Le Chang

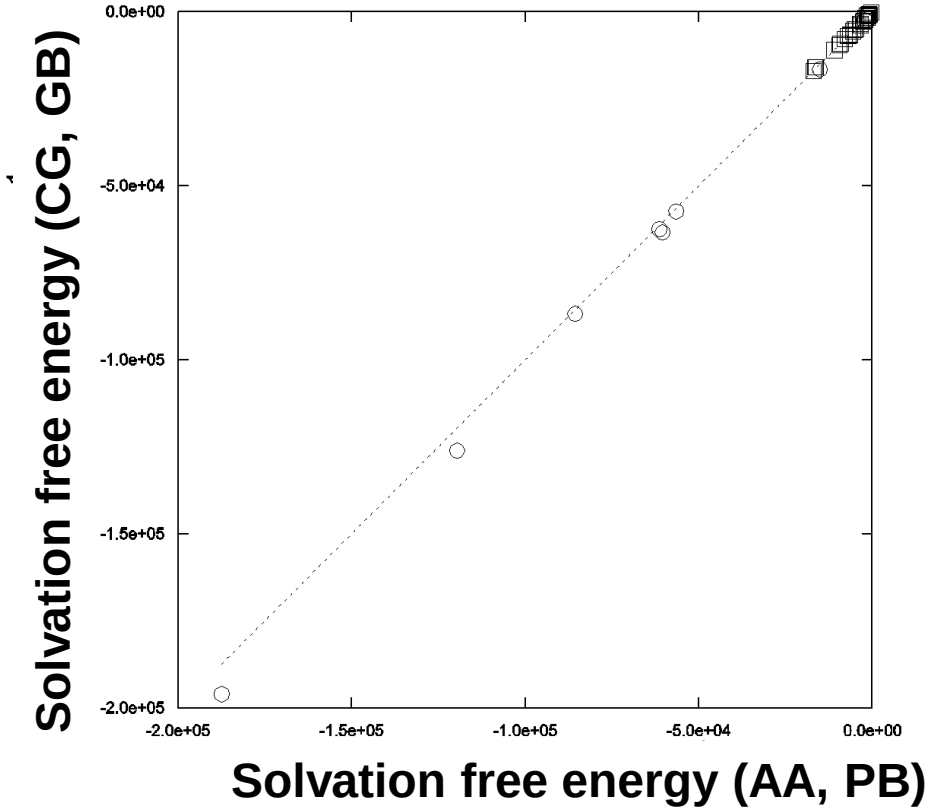
We fit self-energy of point charge
based on Dominy-Brooks formulation of GB

$$G_{pol,i}^{CG,GB} \sim G_{pol,i}^{AA,PB}$$

Parameter optimization



Test against AA PB



Acknowledgement

CafeMol development has been supported by Research and Development of the Next-Generation Integrated Simulation of Living Matter, a part of the Development and Use of the Next-Generation Supercomputer Project of the Ministry of Education, Culture, Sports, Science and Technology.